

Mapping the process of engagement with digital health interventions: an exploratory multiple case study

Volume 1 of 1

Madison Milne-Ives, BAS, MSc

A thesis submitted for the Degree of Doctor of Philosophy (PhD)

Translational and Clinical Research Institute

Faculty of Medical Sciences

Newcastle University

September 2024

Abstract

Digital health is increasingly important for healthcare delivery, but its impact is frequently limited by a lack of user engagement. Engagement is established as a multifaceted construct, including affective, cognitive, and behavioural components, but there is a lack of clarity in the conceptualisation and measurement of these components and a gap in knowledge around the relationships between the different components. To address this gap, this thesis will examine how affective, cognitive, and behavioural components of engagement interact with each other over the process of engagement with different types of digital health interventions and how this could improve the design of beneficial digital tools.

The thesis applied an exploratory multiple case study approach examining three mixed-methods evaluations of digital health interventions. Each case was examined individually to develop themes relating to engagement, usability, and acceptability and to generate a mapping of the relationships between the different components of engagement. Pattern matching, explanation building, and cross-case synthesis techniques were used to compare the individual case findings and examine similarities and differences in the patterns of engagement. A thematic synthesis explored key factors influencing engagement across the interventions and how contextual factors could shape their relevance.

This thesis makes an original contribution by developing a framework comprising: definitions of the components of engagement, a model mapping their relationships over the process of engagement, and recommendations for how to support them in digital health interventions. The framework highlights the close and reciprocal associations between the three different components and the importance of considering all components when evaluating and designing interventions. While this framework is still preliminary, it extends previous research by providing a structure to guide intervention developers and researchers aiming to better understand digital health engagement and how particular barriers can be addressed to improve behavioural and clinical outcomes.

Acknowledgements

First, I would like to thank my primary supervisor Professor Edward Meinert, who knew I would end up here when I joined his lab five years ago, even when I did not. Ed, you have been an incredible mentor to me in research and in my career; thank you for pushing me with the confidence that I could rise to the challenge (and for helping me learn that writing is much more efficient if you don't worry too much about having a perfect first draft!).

To my other supervisors, Professor Jackie Andrade and Dr Sophie Homer, thank you for helping me keep one foot in psychology as I delved into the world of digital health. Our conversations about psychological concepts in different contexts always reminded me of why I love the puzzle of applying theory and synthesising disciplines. I would also like to acknowledge the funder of my research, the Workforce Training and Education Directorate, NHS England (formerly Health Education England) and particularly Em Rahman, who has been incredibly patient and supportive as I navigated the winding route of this project.

There have been so many colleagues and friends who have supported me throughout this journey. To my fellow DEPTH AI lab members and the external colleagues who collaborated on these studies – it has been a pleasure to work with you all. To my friends from Somerville – you have been an inspiring example that a PhD is achievable and brought so much lightness and laughter into my years in England. To my netball girls – joining the Oxford Royals gave me so much joy and getting out of the house for games and socialising (sorry, training!) helped me hang on to my sanity. Finally, to my fellowship from Guelph – your friendship is a gift that I cherish, untarnished by the distance between us.

Gram – I wish I could have called you up to tell you when I finished my final draft, but I know you were so proud. To the rest of my family, I'm so grateful to have you in my life.

Tom – of all the challenges of two partners both doing PhDs, writing my acknowledgments after reading yours is the hardest. You said that you could not have done this without my constant love and support, and I can't find any way to say how I feel better than that.

To my parents, I love you both more than words can express. Mom, your curiosity, determination, and resilience inspire me to take on new challenges and keep persisting when they get hard. Dad, the countless hours you spent with me reading, listening for the story in songs, and exploring ideas (not to mention going through essays line-by-line to learn the art of a well-shaped phrase) are the root of who I am and how I think. Thank you.

Table of Contents

Abstract	i
Acknowledgements	ii
Table of Contents	iii
List of Figures	xi
List of Tables	xiii
Abbreviations	xv
Relevant publications	xix
Chapter 1: Introduction	1
1.1 Chapter Overview	1
1.2 Engagement with Digital Health - What's the Problem?	3
1.3 Theoretical Background	6
1.3.1 Summary of engagement as a concept	6
1.3.2 Key definitions	7
1.4 Rationale	9
1.5 Overview of Case Studies' Clinical Needs	10
1.5.1 Childhood obesity	11
1.5.2 Mental health	12
1.5.3 Post-surgical follow-up	14
1.5.4 Importance of engagement for patient benefit in these clinical cases	15
1.6 Research Question, Aim, and Objectives	15
1.6.1 Development of the research question	15
1.6.2 Aims and objectives	17
1.7 Structure of Thesis	17
Chapter 2: Systematic Review	20
2.1 Chapter Overview	20
2.2 Background	21
2.3 Methods	22

2.3.1 Study design	23
2.3.2 Eligibility criteria	23
2.3.3 Search strategy	24
2.3.4 Screening and article selection	25
2.3.5 Data extraction	25
2.3.6 Data analysis	26
2.3.7 Risk of bias and quality assessment	27
2.3.8 Systematic review update	27
2.4 Results	27
2.4.1 Included studies	27
2.4.2 Study characteristics	28
2.4.3 App characteristics	29
2.4.4 Engagement measures	37
2.4.5 Associations between Behaviour Change Techniques and engagement	41
2.4.6 Exploratory analysis of engagement components	45
2.4.7 Risk of bias and quality assessment	46
2.5 Discussion	47
2.5.1 Principal findings	48
2.5.2 Comparisons with existing literature	48
2.5.3 Strengths and limitations	48
2.5.4 Future research and recommendations	50
2.5.5 Conclusions	51
Chapter 3: Conceptualisation and Measurement of Engagement	53
3.1 Chapter Overview	53
3.2 Origin of this Study	54
3.3 Issues with the Current Measurement of Engagement	55
3.4 Methods	57
3.4.1 Scale selection	57

3.4.2 Analysis of included scales	57
3.5 Results of Scale Comparison	58
3.5.1 Selection of measures for comparison	58
3.5.2 Summary of included multifaceted measures of engagement	60
3.5.3 Comparison of scales	63
3.6 Theoretical Analysis of the DBCI-ES and TWEETS	65
3.6.1 Overall comparison of the engagement scales	65
3.6.2 How many components of engagement are there?	65
3.6.3 Definitions for this thesis	71
3.7 Evidence for Potential Patterns of Engagement Processes	72
3.8 Hypotheses Regarding the Process of Engagement	77
3.9 Implications for Broader Research	80
3.9.1 Recommendations for scale selection and future measure development	80
3.9.2 Strengths and limitations	82
3.9.3 Conclusions	83
Chapter 4: Methods	85
4.1 Chapter Overview	85
4.2 Research Design	87
4.2.1 Choosing a research design	87
4.2.2 Exploratory multiple case study design	89
4.3 Theoretical Frameworks	89
4.4 Case Selection	91
4.4.1 Rationale for case selection	91
4.4.2 Detailed description of interventions	93
4.5 Individual Case Study Methods	95
4.5.1 Case Study 1: NoObesity	96
4.5.2 Case Study 2: Wysa	100
4.5.3 Case Study 3: Dora	106

4.6 Case Study Analytic Approaches	112
4.6.1 Selection of case study analytic approaches	112
4.6.2 Thesis-specific within-case analysis	114
4.6.3 Cross-case analysis	115
Chapter 5: Within-Case Analyses	117
5.1 Chapter Overview	117
5.2 Case Study 1: Barriers and Facilitators to Engagement with and Impact of a Childhood Obesity App	119
5.2.1 Overview	119
5.2.2 Results	120
5.2.3 Influence of Case Study 1 on the thesis	131
5.2.4 Mapping of relationship between engagement components	133
5.2.5 Recommendations in the context of existing evidence	136
5.2.6 Strengths, limitations, and future research	141
5.2.7 Conclusions	142
5.3 Case Study 2: Engagement and Efficacy of a Conversational Agent App for Mental Health Support	143
5.3.1 Overview	144
5.3.2 Results	144
5.3.3 Influence of previous cases studies on Case Study 2	149
5.3.4 Mapping of relationship between engagement components	150
5.3.5 Recommendations in the context of existing evidence	153
5.3.6 Strengths and limitations	155
5.3.7 Conclusions	156
5.4 Case Study 3: Safety and Acceptability of an Autonomous Conversational Agent for Post-Surgical Follow-Up	157
5.4.1 Overview	157
5.4.2 Results	158
5.4.3 Influence of Case Study 3 on the thesis	163

5.4.4 Mapping of relationship between engagement components	164
5.4.5 Strengths and limitations	167
5.4.6 Recommendations in the context of the literature	168
5.4.7 Conclusions	170
5.5 Summary of Within-Case Analyses	171
Chapter 6: Cross-Case Synthesis	174
6.1 Chapter Overview	174
6.2 Summary of Individual Case Findings	175
6.2.1 Case Study 1: NoObesity	175
6.2.2 Case Study 2: Wysa	176
6.2.3 Case Study 3: Dora	176
6.2.4 Within-case mappings of engagement components	177
6.3 Comparability of Cases	178
6.4 Pattern-Matching of Mappings of Engagement Components	180
6.4.1 Overview	180
6.4.2 Pattern-matching analysis in relation to hypotheses	183
6.4.3 Potential explanations for differences between hypothesised and empirical mappings	189
6.5 Thematic Synthesis of Factors Influencing Engagement	191
6.5.1 Overview	191
6.5.2 Theme 1 - connection with the intervention	192
6.5.3 Theme 2 - attention, interest, and incentive to use	194
6.5.4 Theme 3 - usability for the target population	196
6.5.5 Theme 4 - confidence in intervention	197
6.6 Discussion	198
6.6.1 Principal findings	198
6.6.2 Proposed model of how engagement components are related	200
6.6.3 Strengths and limitations	202

6.6.4 Conclusion	203
Chapter 7: Discussion, Strengths and Limitations, Implications, and Conclusions	205
7.1 Chapter Overview	205
7.2 Principal Findings	206
7.3 Strengths and Limitations of the Thesis	208
7.3.1 Terminology and definitions issues	208
7.3.2 Value of case study research	210
7.3.3 Case selection and comparability	211
7.3.4 Generalisability	212
7.3.5 Validity and reliability	212
7.3.6 Other methodological issues	214
7.4 Research, Practical, and Clinical implications	216
7.4.1 Research implications	216
7.4.2 Practical implications: recommendations for digital health intervention development	220
7.4.3 Clinical implications	226
7.5 Reflections and Lessons Learnt	228
7.6 Conclusions	231
References	235
Appendices	290
Appendix 2-A. PRISMA checklist	290
Appendix 2-B. Search results by database (12 July 2022)	293
Appendix 2-C. EndNote screening	296
Appendix 2-D. Engagement components codebook	297
Appendix 2-E. Updated search results by database (15 January 2024)	298
Appendix 2-F. Revised EndNote screening for systematic review update	301
Appendix 2-G. Study characteristics	302
Appendix 2-H. Analysis of potential associations between BCTs and engagement and the evidence for the association	308

Appendix 2-I. MMAT quality assessment	330
Appendix 3-A. DBCI Engagement Scale	335
Appendix 3-B. Twente Engagement with Ehealth Technologies Scale (TWEETS)	336
Appendix 4-A. Case Study 1: Functional overview of NoObesity version 1 apps	337
Appendix 4-B. Case Study 1: Alignment of NASSS and RE-AIM frameworks	339
Appendix 4-C. Case Study 1: SRQR checklist	342
Appendix 4-D. Case Study 1: TREND checklist	345
Appendix 4-E. Case Study 1: SSI guide and study questionnaires	348
Appendix 4-F. Case Study 1: COM-B and TDF coding framework	357
Appendix 4-G. Case Study 2: CONSORT-AI Checklist	361
Appendix 4-H. Case Study 2: SSI topic guide	365
Appendix 4-I. Case Study 2: Initial codebook used in preliminary qualitative coding	367
Appendix 4-J. Case Study 3: DECIDE-AI reporting guideline	368
Appendix 4-K. Case Study 3: SRQR checklist	371
Appendix 4-L. Case Study 3: Dora’s symptom assessments with examples of significant and insignificant symptoms	374
Appendix 4-M. Case Study 3: Dora and clinician decision matrix for the patient care management outcome conveyed at the end of the follow-up assessment	375
Appendix 4-N. Case Study 3: Criteria for a patient to ‘fail’ the Dora call	376
Appendix 4-O. Case Study 3: Post-call survey	377
Appendix 4-P. Case Study 3: SSI guide	382
Appendix 5-A. Case Study 1: Thematic framework	385
Appendix 5-B. Case Study 1: Thematic framework mapping factors to COM-B	389
Appendix 5-C. Case Study 1: Thematic framework mapping factors to engagement components and patterns of engagement	395
Appendix 5-D. Case Study 2: Thematic framework	404
Appendix 5-E. Case Study 2: Initial independent thematic maps and collaborative alignment	408

Appendix 5-F. Case Study 2: Thematic framework mapping factors to engagement components and patterns of engagement	409
Appendix 5-G. Case Study 3: Number of participants involved over the study stages per site	414
Appendix 5-H. Case Study 3: Independent thematic framework	415
Appendix 5-I. Case Study 3: Thematic framework mapping factors to engagement components and patterns of engagement	417
Appendix 6-A. Thematic synthesis framework	424
Appendix 7-A. Positionality and reflexivity statement	432

List of Figures

Figure 1-1. Summary of key theoretical concepts of engagement with digital health

Figure 1-2. Thesis structure diagram

Figure 2-1. PRISMA flow diagram

Figure 3-1. Cognitive triangle tool in Cognitive Behavioural Therapy

Figure 3-2. Simplified diagrams demonstrating proposed processes of engagement adapted from studies and theories

Figure 3-3. Proposed model of the process of engagement with a DHI and its target behaviour

Figure 4-1. Procedures of a multiple case study approach

Figure 4-2. NoObesity version 1 app features

Figure 4-3. Diagram of Wysa's functionality [provided by Wysa Ltd.]

Figure 4-4. Flow diagram of Dora's call with patient

Figure 4-5. Case Study 1 logic diagram

Figure 4-6. Case Study 2 Wisconsin logic diagram

Figure 4-7. Case Study 2 study flow diagram

Figure 4-8. Case Study 2 data collection sequence diagram

Figure 4-9. Case Study 3 logic diagram

Figure 4-10. Case Study 3 study flow diagram

Figure 5-1. Case study 1 setting amongst the overall project

Figure 5-2. Case study 1 participant flow diagram

Figure 5-3. Summary of factors generated in the thematic analysis and how they were associated with engagement

Figure 5-4. Summary of key factors for engagement mapped to the COM-B components

Figure 5-5. Summary of factors generated in the thematic analysis associated with the app's perceived impact

Figure 5-6. Summary of key factors for perceived impact mapped to the COM-B components

Figure 5-7. Mapping of engagement components derived from NoObesity SSIs

Figure 5-8. Case study 2 setting amongst the overall project

Figure 5-9. Case study 2 participant flow diagram

Figure 5-10. Thematic map of qualitative interviews to determine acceptability of and engagement with Wysa

Figure 5-11. Mapping of engagement components derived from Wysa SSIs

Figure 5-12. Case study 3 setting amongst the overall project

Figure 5-13. Case study 3 participant flow diagram

Figure 5-14. Thematic map of qualitative interviews to determine acceptability of Dora calls

Figure 5-15. Mapping of engagement components derived from Dora SSIs

Figure 6-1. Combined mapping of NoObesity, Wysa, and Dora

Figure 6-2. Summary diagram of main themes generated from thematic synthesis

Figure 6-3. Thematic map of factors influencing micro behavioural engagement

Figure 6-4. Proposed model of relationships over the process of affective, cognitive, and behavioural engagement with digital health interventions

Figure 7-1. Framework of how various factors can influence the relationships between different components of engagement

List of Tables

Table 1-1. Definitions of key terms and concepts

Table 2-1. PICOS framework

Table 2-2. Search terms

Table 2-3. Summary of app characteristics

Table 2-4. Behavioural theories used in apps targeting particular health behaviours

Table 2-5. Summary of engagement measures used in the studies

Table 2-6. Analysis of potential associations between BCTs and engagement identified in the studies

Table 2-7. Risk of Bias 2 assessment of randomised studies

Table 3-1. Criteria for scale selection

Table 3-2. Assessment of scales used to measure engagement with DHIs

Table 3-3. Comparison of DBCI-ES and TWEETS

Table 3-4. Example application of CBT ‘cognitive triangle’

Table 4-1. Theoretical frameworks included in the case studies

Table 4-2. Characteristics of the DHIs

Table 4-3. Case study outcome results reported in this thesis

Table 4-4. Case Study 1 quantitative secondary outcome measures

Table 4-5. Case Study 2 outcome measurement

Table 4-6. Case Study 3 inclusion and exclusion criteria

Table 4-7. Case Study 3 outcome measurement

Table 5-1. Participant characteristics of the sample who completed the final quantitative survey (n=35) and interviewed sample (n=15)

Table 5-2. Participant ratings of the app and its features’ perceived impact (n=35)

Table 5-3. Summary of implications for implementing similar interventions in the future

Table 5-4. Engagement with Wysa app

Table 5-5. Participant demographic information

Table 6-1. Pattern matching of theoretical and empirical mappings of engagement components

Table 6-2. Summary of hypotheses and supporting evidence

Table 7-1. List of features and considerations for incorporating them

Abbreviations

A&E	Accident and Emergency (UK emergency department)
ADHD	Attention Deficit Hyperactivity Disorder
AI	Artificial Intelligence
AIM	Affective-Integration-Motivation framework
AIM-ACT	Affective-Integration-Motivation and Attention-Context-Translation framework
APA	American Psychological Association
API	Application Programming Interface
BCT	Behaviour Change Technique
BCW	Behaviour Change Wheel
CA	Conversational Agent
CBT	Cognitive Behavioural Therapy
CDC	Centers for Disease Control and Prevention
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CNWL	Central North West London NHS Foundation Trust
COM-B	Capability, Opportunity, Motivation - Behaviour model
CSQ	Client Service Questionnaire
DBCI	Digital Behaviour Change Intervention
DBCI-ES	Digital Behaviour Change Intervention Engagement Scale
DECIDE-AI	Developmental and Exploratory Clinical Investigations of DEcision support systems driven by Artificial Intelligence
DHI	Digital Health Intervention
EDC	Electronic Data Capture system
EFA	Exploratory Factor Analysis
EHR	Electronic Health Record
EQR	Enlight Quality Ratings

F2F	Face-to-face
GAD-7	Generalised Anxiety Disorder Assessment
HAPA	Health Action Process Approach
HCP	Healthcare professional
HEE	Health Education England (now South East School of Public Health, Workforce Training and Education Directorate, NHS England)
IAPT	Improving Access to Psychological Therapies (now ‘Talking Therapies’)
IMB	Information, Motivation, and Behavioral Skills model
IMI	Intrinsic Motivation Inventory
IQR	Interquartile range
MAC	Meaning and Attentional Components model
MAUQ	mHealth App Usability Questionnaire
MeSH	Medical Subject Headings
MMAT	Mixed-Methods Appraisal Tool
M-PAC	Multiprocess Action Control framework
NASSS	Non-adoption, Abandonment, Scale-up, Spread and Sustainability framework
NHS	National Health Service
NIHR	National Institute for Health and Care Research
NPS	Net Promoter Score
ORBIT	Obesity-Related Behavioral Intervention Trials model
OxIS	Oxford Internet Surveys
PhD	Doctor of Philosophy
PHQ-9	Patient Health Questionnaire (measure of depression severity)
PICOS	Population, Intervention, Comparator, Outcomes, Study
PPI	Patient and Public Involvement

PRIME theory of motivation	Plans, Evaluations, Motives, Impulses, Responses
PRISMA(-P)	Preferred Reporting Items for Systematic Reviews and Meta-analysis (Protocols)
PROSPERO	International prospective register of systematic reviews
RCT	Randomised Controlled Trial
RE-AIM	Reach Effectiveness Adoption Implementation Maintenance framework
RoB	Risk of Bias
ROBINS-I	Risk Of Bias In Non-randomized Studies of Interventions
SCT	Social Cognitive Theory
SD	Standard Deviation
SDT	Self-Determination Theory
SMS	Short Messaging Service (text message)
SPIRIT	Standard Protocol Items: Recommendations for Interventional Trials
SRQR	Standards for Reporting Qualitative Research
SSI	Semi-structured interview
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUS	System Usability Scale
TDF	Theoretical Domains Framework
TFA	Theoretical Framework of Acceptability
TPB	Theory of Planned Behaviour
TREND	Transparent Reporting of Evaluations with Nonrandomized Designs
TUQ	Telehealth Usability Questionnaire
TWEETS	TWente Engagement with Ehealth Technologies Scale
UEQ	User Experience Questionnaire
UI	User interface
UK	United Kingdom

uMARS	User Mobile App Rating Scale
UMC	Unexpected Management Changes
US	United States
USD	United States Dollar
USE	Usefulness, Satisfaction and Ease of Use questionnaire
UTAUT	Unified Theory of Acceptance and Use of Technology
UX	User experience

Relevant Publications

The following publications have informed this thesis. Sections of the thesis have been previously published or submitted for publication in these papers; the overview section for each chapter highlights which publications have been incorporated in that chapter. References and author contributions for each of the papers that were used in the preparation of this thesis are listed below.

1. **Milne-Ives M**, Lam C, Van Velthoven MH, Meinert E. Mobile Apps for Health Behavior Change: Protocol for a Systematic Review. *JMIR Research Protocols*. 2020 Jan 30; 9(1):e16931. doi:[10.2196/16931](https://doi.org/10.2196/16931)

Authors' contributions: EM and MMI conceived the study topic and designed the review protocol. MMI prepared the first draft of the protocol with revisions from CL, MVV, and EM.

2. **Milne-Ives M**, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile Apps for Health Behavior Change in Physical Activity, Diet, Drug and Alcohol Use, and Mental Health: Systematic Review. *JMIR mHealth and uHealth*. 2020 Mar 18; 8(3), e17046. doi:[10.2196/17046](https://doi.org/10.2196/17046)

Authors' contributions: MMI and CL screened the studies, MMI conducted the data extraction and risk of bias assessment, which were validated by CC. The systematic review was written by MMI with revisions from MV and EM.

3. de Pennington N, Mole G, Lim E, **Milne-Ives M**, Normando E, Xue K, Meinert E. Safety and Acceptability of a Natural Language Artificial Intelligence Assistant to Deliver Clinical Follow-up to Cataract Surgery Patients: Proposal. *JMIR Research Protocols*. 2021 Jul 28; 10(7):e27227. doi:[10.2196%2F27227](https://doi.org/10.2196/27227)

Authors' contributions: NP conceived the study topic and designed and drafted the proposal. GM, EL, MMI, EN, and KX contributed to drafting and revising the proposal. Final revision was conducted by EM.

4. **Milne-Ives M**, Homer S, Andrade J, Meinert E. Associations Between Behavior Change Techniques and Engagement With Mobile Health Apps: Protocol for a Systematic Review. *JMIR Research Protocols*. 2022 Mar 29; 11(3):e35172. doi:[10.2196/35172](https://doi.org/10.2196/35172)

Authors' contributions: The protocol was conceived and written by MMI with revisions from EM, SH, and JA.

5. **Milne-Ives M**, Homer S, Andrade J, Meinert E. Potential associations between behavior change techniques and engagement with mobile health apps: a systematic review. *Frontiers in Psychology*. 2023 Sep 18; 14:1227443. doi:[10.3389/fpsyg.2023.1227443](https://doi.org/10.3389/fpsyg.2023.1227443)

Authors' contributions: MMI conceived the topic, drafted the protocol, conducted the searches, screening, data extraction and analysis, and wrote the first draft of the paper. SH and JA contributed revisions and EM performed the final review and supervised the execution of the review.

6. **Milne-Ives M**, Rahman E, Bradwell H, Baines R, Boey T, Potter A, Lawrence W, van Velthoven MH, Meinert E. Barriers and facilitators to parents' engagement with and perceived impact of a childhood obesity app: A mixed-methods study. *PLOS Digital Health*. 2024 Mar 27; 3(3):e0000481. doi:[10.1371/journal.pdig.0000481](https://doi.org/10.1371/journal.pdig.0000481)

Authors' contributions: MMI collected data, carried out the initial analyses, drafted the initial manuscript, and critically reviewed and revised the manuscript. ER, AP, WL, and MvV conceptualised and designed the study and critically reviewed and revised the manuscript. HB, RB, and TB, analysed data and critically reviewed and revised the manuscript. EM conceptualised and designed the study, supervised data collection and analysis, and critically reviewed and revised the manuscript.

7. **Milne-Ives M**, Homer S, Andrade J, Meinert E. The conceptualisation and measurement of engagement in digital health. *Internet Interventions*. 2024 Jun; 36:100735. doi:[10.1016/j.invent.2024.100735](https://doi.org/10.1016/j.invent.2024.100735)

Authors' contributions: MMI conceived the topic, conducted the analysis, and drafted and revised the paper. SH, JA, and EM supervised the research and contributed revisions.

8. Meinert E, **Milne-Ives M**, Lim E, Higham A, Boege S, de Pennington N, Bajre M, Mole, G, Normando E, Xue K. Accuracy and safety of an autonomous artificial intelligence clinical assistant conducting telemedicine follow-up assessment for cataract surgery. *eClinical Medicine*. 2024 Jul 4; 73:102692. doi:[10.1016/j.eclinm.2024.102692](https://doi.org/10.1016/j.eclinm.2024.102692)

Authors' contributions: NdP conceived of the study topic. The study design was developed by NdP, GM, EL, MMI, EN, KX, EM and AH. KX and EN led patient recruitment at the two study sites. AH and EL supervised Dora R1's calls and compiled the clinical data. MMI conducted the usability survey and MMI and EM completed participant interviews. Statistical analysis of the quantitative data was

conducted by the University of Plymouth. MMI, SB and EM conducted the thematic analysis of the qualitative data. MB was responsible for the cost analysis. EM set the structure for the manuscript, MMI drafted the paper based on this framework, and it was reviewed by all authors.

9. Meinert E, **Milne-Ives M**, Taylor E, Inkster B, Paik A, Ananthakrishnan A, Orr M, Costelloe CE, Shankar R. Real-world testing of an artificial intelligence conversational agent as an early intervention and support tool in the mental health referral care pathway: A randomised controlled trial. *International Journal of Social Psychiatry*. 2024; Under Review.

Authors' contributions: All authors contributed to the study design. Participants were recruited and screened and quantitative data was collected by AP. MMI conducted the participant interviews. CC and MO conducted statistical analysis of the quantitative data and MMI and AA conducted the thematic analysis of the qualitative data. EM set the structure for the manuscript, MMI and CC drafted the paper based on this framework, and it was reviewed by all authors with final revisions from EM.

10. **Milne-Ives M**, Homer S, Andrade J, Meinert E. Engagement with and impact of a mobile health app for childhood obesity prevention and management: a mixed methods study protocol. *JMIR Research Protocols*. 2025; Under Review. doi:[10.21203/rs.3.rs-4510115/v1](https://doi.org/10.21203/rs.3.rs-4510115/v1) (Preprint)

Authors' contributions: MMI and EM conceived of the study topic and designed and drafted the protocol. SH and JA contributed to the drafting and revision of the protocol. Final revision was conducted by MMI, AA, and EM.

11. **Milne-Ives M**, Homer S, Andrade J, Meinert E. Mapping the process of engagement with digital health interventions: A cross-case synthesis. *Mayo Clinic Proceedings: Innovations, Quality & Outcomes*. 2025; Under Review.

MMI conceived the topic, conducted the analysis, and drafted and revised the manuscript. SH, JA, and EM supervised the research and contributed revisions.

12. **Milne-Ives M**, Homer S, Andrade J, Meinert E. Exploring factors influencing engagement across different digital health interventions: thematic synthesis. *Nature Scientific Reports*. 2025; Under Review.

MMI conceived the topic, conducted the analysis, and drafted and revised the manuscript. SH, JA, and EM supervised the research and contributed revisions.

Chapter 1: Introduction

1.1 Chapter Overview

Chapter 1 introduces the research topic and provides an overview of the problem of achieving sufficient engagement with digital health interventions to demonstrate the rationale for the investigation. It provides an initial overview of the theoretical basis for the thesis' research question and aims and introduces the specific clinical areas that will be examined. The chapter concludes by laying out the structure of the thesis.

Publications: Sections of Chapter 1 have been drawn from previously published papers; references and author contributions for each of the papers that were used in the preparation of this chapter are listed below.

Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile Apps for Health Behavior Change in Physical Activity, Diet, Drug and Alcohol Use, and Mental Health: Systematic Review. *JMIR mHealth and uHealth*. 2020 Mar 18; 8(3), e17046. doi:[10.2196/17046](https://doi.org/10.2196/17046)

Authors' contributions: MMI and CL screened the studies, MMI conducted the data extraction and risk of bias assessment, which were validated by CC. The systematic review was written by MMI with revisions from MV and EM.

de Pennington N, Mole G, Lim E, **Milne-Ives M**, Normando E, Xue K, Meinert E. Safety and Acceptability of a Natural Language Artificial Intelligence Assistant to Deliver Clinical Follow-up to Cataract Surgery Patients: Proposal. *JMIR Research Protocols*. 2021 Jul 28; 10(7):e27227. doi:[10.2196/27227](https://doi.org/10.2196/27227)

Authors' contributions: NP conceived the study topic and designed and drafted the proposal. GM, EL, MMI, EN, and KX contributed to drafting and revising the proposal. Final revision was conducted by EM.

Milne-Ives M, Homer S, Andrade J, Meinert E. Associations Between Behavior Change Techniques and Engagement With Mobile Health Apps: Protocol for a Systematic Review. *JMIR Research Protocols*. 2022 Mar 29; 11(3):e35172. doi:[10.2196/35172](https://doi.org/10.2196/35172)

Authors' contributions: The protocol was conceived and written by MMI with revisions from EM, SH, and JA.

Milne-Ives M, Rahman E, Bradwell H, Baines R, Boey T, Potter A, Lawrence W, van Velthoven MH, Meinert E. Barriers and facilitators to parents' engagement with and perceived impact of a childhood obesity app: A mixed-methods study. *PLOS Digital Health*. 2024 Mar 27; 3(3):e0000481. doi:[10.1371/journal.pdig.0000481](https://doi.org/10.1371/journal.pdig.0000481)

Authors' contributions: MMI collected data, carried out the initial analyses, and drafted the initial manuscript, and critically reviewed and revised the manuscript. ER, AP, WL, and MvV conceptualised and designed the study and critically reviewed and revised the manuscript. HB, RB, and TB, analysed data and critically reviewed and revised the manuscript. EM conceptualised and designed the study, supervised data collection and analysis, and critically reviewed and revised the manuscript.

Milne-Ives M, Homer S, Andrade J, Meinert E. The conceptualisation and measurement of engagement in digital health. *Internet Interventions*. 2024 Jun; 36:100735. doi:[10.1016/j.invent.2024.100735](https://doi.org/10.1016/j.invent.2024.100735)

Authors' contributions: MMI conceived the topic, conducted the analysis, and drafted and revised the protocol. SH, JA, and EM supervised the research and contributed revisions.

Meinert E, **Milne-Ives M**, Lim E, Higham A, Boege S, de Pennington N, Bajre M, Mole, G, Normando E, Xue K. Accuracy and safety of an autonomous artificial intelligence clinical assistant conducting telemedicine follow-up assessment for cataract surgery. *eClinical Medicine*. 2024 Jul 4; 73:102692. doi:[10.1016/j.eclinm.2024.102692](https://doi.org/10.1016/j.eclinm.2024.102692)

Authors' contributions: NdP conceived of the study topic. The study design was developed by NdP, GM, EL, MMI, EN, KX, EM and AH. KX and EN led patient recruitment at the two study sites. AH and EL supervised Dora R1's calls and compiled the clinical data. MMI conducted the usability survey and MMI and EM completed participant interviews. Statistical analysis of the quantitative data was conducted by the University of Plymouth, MMI, SB and EM conducted the thematic analysis of the qualitative data. MB was responsible for the cost analysis. EM set the structure for the manuscript, MMI drafted the paper based on this framework, and it was reviewed by all authors.

Meinert E, **Milne-Ives M**, Taylor E, Inkster B, Paik A, Ananthakrishnan A, Orr M, Costelloe CE, Shankar R. Real-world testing of an artificial intelligence

conversational agent as an early intervention and support tool in the mental health referral care pathway: A randomised controlled trial. *International Journal of Social Psychiatry*. 2024; Under Review.

Authors' contributions: All authors contributed to the study design.

Participants were recruited and screened and quantitative data was collected by AP. MMI conducted the participant interviews. CC and MO conducted statistical analysis of the quantitative data and MMI and AA conducted the thematic analysis of the qualitative data. EM set the structure for the manuscript, MMI and CC drafted the paper based on this framework, and it was reviewed by all authors with final revisions from EM.

Milne-Ives M, Homer S, Andrade J, Meinert E. Engagement with and impact of a mobile health app for childhood obesity prevention and management: a mixed methods study protocol. *JMIR Research Protocols*. 2025; Under Review.

Authors' contributions: MMI and EM conceived of the study topic and designed and drafted the protocol. SH and JA contributed to the drafting and revision of the protocol. Final revision was conducted by MMI, AA, and EM.

1.2 Engagement with Digital Health - What's the Problem?

*Parts of this section have been previously published in [1–4].

The role of digital technologies in health, particularly those that centre the patient as a key user, has grown exponentially in the past two decades [5]. Digital health technologies have been applied to support a large range of needs, including collecting, storing, and managing data; analysing data to inform clinical decision-making, predict risk, and identify health issues; and delivering behavioural interventions [6,7]. Given the ubiquity of digital devices such as smartphones and wearables, digital health technologies have the potential to streamline, enhance, and transform the availability and delivery of healthcare services. This is a key need for healthcare systems that are facing increasing demands beyond their capacity to address in a timely manner, but the impact of these technologies is limited by a critical factor: lack of engagement. Low engagement with digital health interventions (DHIs) is commonly observed in clinical trials and real-world observational studies and is a significant barrier to achieving the potential of digital health [8–13]. Although it may vary by intervention and individual, a certain degree of engagement with a DHI is necessary to achieve its intended outcomes [14]. Despite a recent increase in focus on the conceptualisation and evaluation of engagement with DHIs, there is a need to better understand users' processes of engagement

with DHIs to determine how the different components of engagement can be best supported at different stages to improve the impact of the interventions.

As demand outstrips clinical resources, access to services is a growing issue in many areas of healthcare [15]. The lack of capacity to meet healthcare demands can result in long wait times, limited clinical contact, and clinician burnout [15,16]. As countries struggle to meet this demand, there has been an increasing focus on supporting healthcare service delivery through the development and adoption of digital health strategies [17], particularly since the Covid-19 pandemic. For example, the UK's 2019 NHS Long Term Plan had key aims to develop shared responsibility of health, achieve digitally-enabled care, and deliver P4 medicine (predictive, personalised, preventive and participatory care) [18] and the US CDC's Global Digital Health Strategy is working towards increasing the availability and uptake of digital tools and use of data across the health workforce [19]. Digital health has been variously defined in the literature [20], but generally refers to the use of digital technologies to gather, analyse, and share health data to empower and centre patients in care, improve the quality and efficiency of care delivery, and optimise health outcomes [20–22] (a specific definition of DHIs for this thesis is provided in Table 1-1). In addition to facilitating clinical tasks, digital technology can help empower patients to be actively engaged with their own health management. This is important because it has the potential to reduce the demand for healthcare services and improve health factors related to preventable mortality [23,24].

Since the first iPhone was released in 2007, personal digital technology has become increasingly prevalent and capable, offering a promising means of delivering healthcare interventions to the general population. The large number of mobile health apps currently available for download is a testament to their popularity; estimates have found that there are over 350,000 health-related apps across a wide variety of well-being and clinical health areas [25,26] and over 6 billion smartphone users worldwide [27]. Many smartphones now have the ability to passively collect a variety of health data - including physical activity, social interaction, sleep, and mobility patterns - and make inferences about mental and physical health [28,29], which can be further enhanced with wearable devices. Combining these capabilities with active user interaction allows digital tools to deliver many different behavioural interventions which can help users lead healthier lives and to collect a variety of biopsychosocial data that can inform their health management.

This perceived accessibility is one of the main selling points of digital health technologies. The ubiquity of digital devices makes digital an appealing avenue for healthcare, as a means of gathering and analysing data and delivering healthcare interventions, but their impact is

dependent on users' willingness to engage with them [30]. 'Effective engagement' has been defined as "sufficient engagement with the intervention to achieve intended outcomes" [14]. Achieving 'effective engagement' is a common challenge for DHIs, across clinical areas and study types. Studies of engagement with mobile health apps and wearable devices often observe poor long-term use [13,31,32] and high rates of attrition [12,33]. Systematic reviews and meta-analyses have found that, while engagement can vary significantly depending on the intervention and user, there are high overall rates of drop-out. A meta-analysis of app interventions for chronic disease estimated a pooled drop-out rate of 43%, with higher rates in real-world studies than supported trials [8]. Accounting for publication bias, a meta-analysis of clinical trials of mental health apps found a similar pooled drop-out rate of almost 50% [9]. A study investigating real-world use of mental health apps found that, although there was a large number of downloads, there was only a median of 4% of people opening the app on any given day and there was 30-day retention of only 3% of users [13]. Another review of real-world use of mental health apps found that the proportion of users who had sustained use or completed the intervention ranged from less than 1% to a high of 29% [34]. Health apps more generally also have been reported to have low retention, with one review reporting a median retention of 5.5 days [10]. This review also highlighted another issue with the perceived accessibility of DHIs: digital health equity. There were large variations in engagement based on user characteristics (e.g. age) and most of the studies were unable to recruit a sample that was representative of the US population in terms of ethnicity or geography [10]. This highlights the importance of considering demographic and contextual factors as well as psychological factors when examining engagement with DHIs and how it can be improved.

Understanding how these barriers influence engagement is a key step in developing DHIs that can have a significant impact on health behaviours and outcomes. Previous research has identified positive associations between engagement with DHIs and efficacy in a variety of contexts, including mental health and weight-related behaviours. One study specifically examined the association between subjective engagement and mental health outcomes and found significant correlations between each of the subscales of the Usefulness, Satisfaction, and Ease of Use (USE) Questionnaire and anxiety and depression [35]. Analyses of the associations between engagement and impact are hindered by limited reporting - DHI evaluations tend to focus primarily on objective usage metrics of engagement - and the heterogeneity of measures used when engagement is assessed in more depth [36–41]. To better understand how engagement is related to behavioural and health outcomes and how it can be supported to achieve an intervention's intended outcomes, there is a need to better understand how users are engaging with DHIs. Recently, there has been a growing awareness

and research focus on engagement as a critical element in the success of DHIs, with studies aiming to clarify how engagement is conceptualised [42–45] and measured [14,30,44,46] in the digital health field. [Section 1.3](#) provides an overview of recent conceptualisations of engagement in the literature and summarises the key definitions and theoretical propositions that can be derived from the research about digital health engagement that has been conducted so far.

1.3 Theoretical Background

1.3.1 Summary of engagement as a concept

*Parts of this section have been previously published in [1,3,47]

Although various models and definitions have been proposed [45] and no definitive conceptualisation has been established [48,49], there is a general consensus that engagement is a multifaceted concept [14,42,45,49–51]. Conceptualisations of engagement generally include a few common components that can be grouped into two main categories (Figure 1-1): what users are engaging with (the DHI or the target health behaviour [14,50]) and how they are engaging with it (e.g. behaviourally, affectively, cognitively [45]). Each of these categories includes various components that are grouped differently depending on the theory. For example, Figure 1a demonstrates how engagement can refer to micro engagement with the intervention, which can be further broken down into engagement with the intervention interface or its ‘active ingredients’ (contents and features), and macro engagement with the target health behaviour that the intervention aims to change [14,50]. Likewise, Figure 1b highlights key components that reflect users’ type of engagement with an intervention: affective and cognitive (sometimes considered as one ‘subjective experience’ component) and behavioural [42,45]. For the systematic review detailed in Chapter 2, key elements of each component were identified from the literature. Perski and colleagues define ‘subjective experience’ as ‘attention, interest, and affect’ [42]. Other literature considers affective engagement to capture motivation, positive and negative feelings, and sensory pleasure and cognitive engagement to include aspects relating to challenge, interest, thought, and attention [42,45,51]. Behavioural engagement can include various measures of usage (amount, duration, frequency) and interactivity [42,45,51]. All three components can relate to both micro and macro engagement. These concepts are critically examined in Chapter 3.

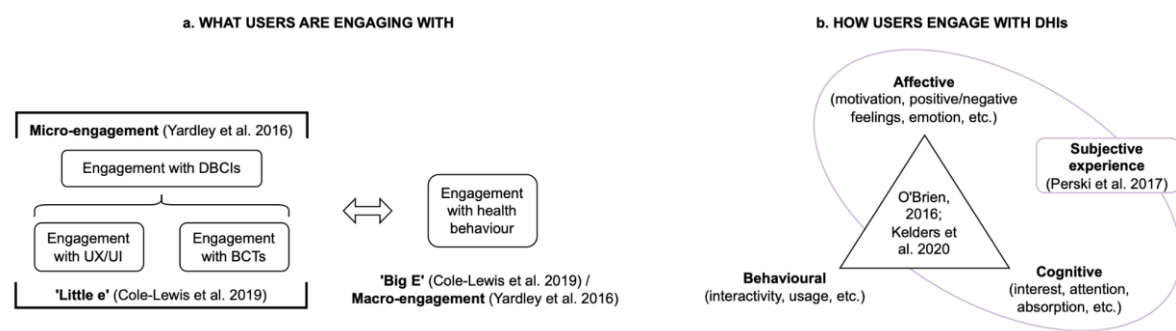


Figure 1-1. Summary of key theoretical concepts of engagement with digital health (based on [14,42,45,50,51], previously published in [1,47]); BCTs: Behaviour Change Techniques, DBCI: digital behaviour change intervention, DHI: digital health intervention, UI: user interface, UX: user experience

1.3.2 Key definitions

The definitions used for key terms and concepts throughout the thesis are provided in Table 1-1. Some definitions were adopted from existing literature; others were developed as part of work conducted during this thesis. For instance, the definitions of affective, cognitive, and behavioural engagement that will be used in this thesis were developed from a critical examination of the literature conducted in Chapters 2 and 3.

Table 1-1. Definitions of key terms and concepts

Term	Definition
Behaviour	How an individual acts “in response to external or internal stimuli” [52]
Behaviour change intervention	“Coordinated sets of activities designed to change specified behaviour patterns” [53]
Behaviour Change Techniques (BCTs)	“The smallest parts of the content of a behaviour change intervention that are observable, replicable and on their own have the potential to bring about behaviour change” [54–56]
Conversational agent (CA)	“Digital interfaces that use natural language to engage in a synchronous dialogue using [at least one] communication modality, such as text, voice, images, or video” [57] In this thesis, there are two conversational agents discussed: a text-based CA accessed via a mobile app (“Wysa”) and a voice-based CA delivered via telephone call (“Dora”).

Digital behaviour change intervention (DBCI)	A type of DHI with the intended purpose of modifying a particular behaviour to support positive health outcomes [58,59]
Digital health intervention (DHI)	An activity or set of activities intended to improve the health or well-being of an individual that is delivered via digital technology (including but not limited to mobile phones, websites, artificial intelligence, conversational agents, wearables, and other information technology systems) [20,60,61] In this thesis, the DHIs evaluated included mobile health apps and CAs (delivered via app and telephone call).
Engagement	A multifaceted concept that includes affective, cognitive, and behavioural components of a users' interpretation of, interaction with, and response to a digital health intervention [14,42,45,49–51]
Affective engagement	Defined for this thesis in Section 3.6.3 as “the positive and negative feelings, emotions, moods, interest, and attitudes (e.g. enjoyment, satisfaction, frustration) related to perceptions of and experiences engaging with the intervention and the behavioural goal [45,49,62–66]. It will also include elements of motivation relating to the alignment of the intervention and target behaviour to users' goals and values, in both the immediate moment and in the long-term [64,67–70].”
Behavioural engagement	Defined for this thesis in Section 3.6.3 as “individuals' interaction with a stimulus or their participation in or performance of a particular activity. This can include quantity and quality of use - frequency, amount, depth, duration, and accordance with intended use in line with intervention goal - of the intervention and performance of the target health behaviour [42,43,45,49,66]. Factors that can influence behavioural engagement (such as effort) will be considered in relation to it in case study evaluations, but are not included within the definition of behavioural engagement.”
Cognitive engagement	Defined for this thesis in Section 3.6.3 as “users' interest and attentional effort invested in the decision-making processes and learning associated with the intervention and the target behaviour [42,49,63,66,67]. It will also include elements of motivation relating to assessments of capability, risks, and benefits (perceived usefulness, control, and potential for success) [62,64,65,70].”
Effective engagement	“Sufficient engagement with the intervention to achieve intended outcomes” [14]
Micro	“Moment-to-moment engagement with the intervention” [14]; in this thesis this concept is defined as engagement with the digital

engagement	health device, interface, and content (active intervention ingredients) [50]
Macro engagement	In this thesis, the concept of macro engagement is defined as including Yardley <i>et al.</i> 's definition of macro engagement as "engagement and identification with the wider intervention goals" [14] and Cole-Lewis <i>et al.</i> 's definition of 'Big E' as "engagement with the targeted health behaviour" [50]
Health behaviour	Health behaviours can be broadly defined as "actions taken by individuals that affect health or mortality ... examples include smoking, substance use, diet, physical activity, sleep, risky sexual activities, health care seeking behaviours, and adherence to prescribed medical treatments" [71]. A broad interpretation of what can be considered a health behaviour was adopted for this thesis.

1.4 Rationale

The previous sections have demonstrated that engagement is an important factor influencing the success of DHIs. Low engagement with DHIs can limit their potential impact [12,30,31], although the variety in definitions and measures of engagement means that reliable quantitative estimates of the relationship between engagement and outcomes are still lacking [12]. As healthcare service delivery becomes increasingly digital and accessible through personal devices like smartphones and wearables [18,72], there is a need to ensure that these digital interventions are achieving their intended outcomes. The necessity and importance of this thesis can be argued from two perspectives: theoretical and clinical.

Despite the importance of engagement for achieving intervention aims, the examination of engagement in digital health studies is still limited. One of the issues that is limiting our understanding of engagement is that many digital health evaluations only capture behavioural (app use) measures of engagement [14,30,44,45]. This is an issue because, as demonstrated in [Section 1.3](#), these only measure one component of engagement and does not enable a more holistic understanding of how users engage with DHIs and how various factors influence the components of engagement. Although advances have been made in clarifying the theoretical conceptualisation of engagement in digital health, the lack of in-depth and multifaceted empirical investigation of engagement and a focus on micro behavioural engagement with DHIs means that our understanding of how the different components of engagement interact with each other over the course of users' engagement with a DHI and its aims is limited. Improving our understanding of how these different components are related will help identify which components may be most important to support effective engagement at various stages

in users' processes of engagement with a DHI, which would be expected to improve the intervention's behavioural and health impact. These patterns of engagement may vary depending on the type of intervention, health application, or individual characteristics; further investigation will enable the generation of recommendations for how to best support engagement with DHIs in general and strategies for tailoring interventions to meet the needs of different users and contexts.

So far, this chapter has focused on demonstrating the gap in the literature that this thesis will aim to address; however, the value in addressing this gap is grounded in how it could help meet patient health needs. This thesis examines three DHIs as case studies. The following section will introduce the three case studies by describing the patient and clinical needs that interventions are aiming to address and will demonstrate the practical benefits that better understanding of engagement with DHIs would bring to patients.

1.5 Overview of Case Studies' Clinical Needs

The three DHIs that will be examined in this thesis aim to address patient health needs in three distinct clinical areas: childhood obesity, mental health, and post-surgical follow-up. There is value for this thesis in using a diverse set of DHIs - in terms of technology, aim, and clinical area ([Section 4.4](#)) - because it will enable analysis of how engagement might vary in different contexts and transferability of findings. Although focused on distinct clinical areas, the included DHIs address two main interconnected needs: patient health outcomes and clinical burden. Childhood obesity and mental health concerns are among some of the most common health challenges of the 21st century [73,74]. They can have significant negative impacts on individuals' physical and mental health and quality of life and on healthcare resources and economies. Although both are complex conditions with many determinants, behavioural factors can help improve patient outcomes [75,76]. The prevalence of these conditions means that there is a lack of clinical capacity to provide behavioural support to all the patients who could benefit from it. Digital behaviour change interventions (DBCIs) have been developed in abundance to fill this gap; previous research has demonstrated the potential of such interventions, but many of them lack a strong theoretical basis or evidence for their proposed health benefits. The expected patient benefit in the third case study is slightly different: rather than aiming to change patients' behaviour, it aims to improve patient outcomes by improving the efficiency of clinical review post-cataract surgery. By automating routine tasks such as post-surgical follow-up, this application of digital health technology can reduce the time clinicians must spend on low-skill tasks and enable them to prioritise patients with the greatest need, so that they are seen more quickly.

The following subsections will examine the patient health need in these cases in more detail to illustrate key current challenges in healthcare provision and how digital technology could help address that need. As the demand for healthcare services exceeds systems' capacities to deliver timely care, digital health technologies become increasingly important as a means of bridging this gap. As highlighted in previous sections, achieving these potential patient and clinical benefits is dependent on the extent to which patients (and clinicians, in some cases) engage with the interventions. This is how the learnings that will be generated by this investigation could have a significant impact on patient outcomes and the healthcare system.

1.5.1 Childhood obesity

*This section was previously published in [77].

Non-communicable diseases are one of the major health challenges of the 21st century [73]. Childhood obesity is a growing global health concern affecting around 1 in 5 children worldwide [78–80]. The worldwide prevalence of overweight and obesity is growing rapidly in children [80], particularly in developed countries like the US and UK. In both countries, more than 2 out of 5 children are obese by age 11 [81,82]. Being overweight or obese can negatively impact a child's physical health [83–86], self-esteem, social interaction, quality of life, and lifespan [87–90] and strain healthcare resources [91]. In the UK in 2015, it was estimated that approximately £6.1 (USD\$9.3) billion was being spent on obesity-related healthcare annually; by 2050, this is expected to increase to £9.7 (USD\$14.8) billion [92,93]. In the US, obesity-related healthcare costs almost USD\$173 billion per year [94]. Determinants of childhood obesity are complex and interrelated - including structural socio-economic inequities and genetics [95–97] - but on an individual level, prevention and management primarily focuses on lifestyle behavioural contributors like diet and activity [85,98–100]. These behaviours are well-established factors in other non-communicable diseases such as cancer, heart disease, stroke, and type 2 diabetes [83–85].

There is a growing demand for digital interventions to ease the burden they cause healthcare systems [33,73]. Mobile apps are a promising tool to help support dietary and physical activity behaviour change [101–104], but thousands of health and fitness apps are available [105,106] and many do not have robust evidence of a long-term positive impact on weight [2,103,107–109]. The ubiquity of smartphones and the growing market for weight loss and health apps [105] has resulted in a large and growing number of digital interventions aiming to support healthy behaviour change and weight loss [25,110,111]. Digital technologies have a significant and well-recognized potential to improve the quality, accessibility, and

sustainability self-management support for weight and related behaviours [112–114]. A significant amount of research about the effectiveness and value of digital interventions for health behaviour change has been conducted [115]. The potential benefits of mobile apps for weight management has been supported by studies that have found evidence of effectiveness at improving health behaviours and weight [37,102,116,117], especially for highly engaged users [37,116]. However, the benefits of many publicly-available apps have yet to be proven or lack conclusive evidence of effectiveness at improving health behaviours and outcomes [2].

Evidence of the impact of digital interventions for childhood obesity specifically is still emerging, but some studies have observed positive impacts [107,117–120]. The mixed evidence of effectiveness in the literature as a whole points to issues with the design and implementation of DHIs. A key limiting factor of the effectiveness of DBCIs is user engagement and adherence [14,42]; to be successful, it is essential that mobile health apps achieve sufficient user engagement to deliver the intervention as designed [14,37,116]. Although there are many theory-based strategies to improve engagement with digital interventions and health behaviours [121–124], many publicly-available apps lack a theoretical basis [2,125]. Emerging research is starting to examine how to best engage various users (parents and children) with digital interventions for childhood obesity [126] but further investigation is needed to optimise intervention design to support engagement for different types of users [1,47,127,128]. Therefore, there is a need to understand how digital tools for weight management can best support users' engagement with the intervention and the health behaviours to achieve positive outcomes.

This first case study examined parents' engagement with a mobile app for childhood obesity ("NoObesity") to identify barriers and facilitators to their engagement with the intervention and to examine their perceptions of its impact on their families' motivation, self-efficacy, and engagement with relevant health behaviours (such as physical activity and healthy eating).

1.5.2 Mental health

*Parts of this section have been submitted for publication in [129].

Mental health is another key clinical area that can have significant negative impacts on health, well-being, and economies [130]. Mental health conditions can have wide ranging negative impacts on individuals' lives and well-being. Depression and anxiety disorders can lead to a range of adverse psychological, social, and employment outcomes [131]. Mental health conditions also affect physical health; anxiety, depression, and substance abuse disorders have

been associated with an increased risk of non-communicable diseases such as diabetes, cancer, and cardiovascular diseases [132–134]. By 2030, mental health concerns are projected to be the global leading cause of mortality and morbidity [135]. Lack of support at an early stage can lead to poorer physical health outcomes and the development of unhealthy coping strategies such as alcohol consumption or poorer diet choices [136]. Prolonged, severe mental illness is associated with reduced life expectancy; with people at risk of dying an average of 15 years sooner than those without severe mental illness. The cause of death in such cases is often avoidable, with two thirds of the deaths caused by heart disease and cancer, often as a result of excess smoking [137].

There is a growing recognition of the importance of mental health, which is reflected across the UK and international health sectors. However, common mental health concerns frequently go unsupported. It is estimated that 1-in-6 people will have experienced a common mental health concern in the past week below the threshold of a diagnosable mental health concern [138] and that many people (~75% in one study) who would benefit from support for common mental health concerns do not receive it, resulting in approximate costs of £105.2 billion a year to the UK economy [137,139]. Globally, depression and anxiety are estimated to cost USD\$1 trillion each year [140].

In 2008, a programme for Improving Access to Psychological Therapies (IAPT, now Talking Therapies) was launched by the NHS [141]. The large number of referrals demonstrates the demand for mental health services in the UK; over 1.5 million referrals to adult IAPT services were made in 2018/2019 [142]. For those patients with recognised mental health conditions who are able to access mental health services in the UK, negative experiences are commonly reported [143]. Over a quarter of these patients (n=441,775) were discharged without face-to-face contact following a telephone referral and another quarter (n=440,344) were discharged with a workbook [142]. Nearly half of surveyed patients (44%) who had received any NHS therapies in 2019 reported that they had to wait too long to access them [143] and in 2017, half of the respondents reported only being offered pharmaceutical treatment [144]. Long wait times to access mental health services can have significant health and economic impacts; a recent study by the Royal College of Psychiatrists found that 38% of people waiting for treatment reach out to crisis services and 11% go to A&E [145]. With each contact costing an average of approximately £100 [146], the total cost to the NHS was over £90 million. This figure does not include additional care needs or A&E attendances for those individuals who were unable to wait up to 12 weeks for assessment.

The economic cost of mental health conditions can occur from sick days, reduced quality of life, and reduction in productivity [130,135]. Depression, anxiety, or stress are some of the most common reasons for lost work days [147]. Symptoms of common mental health conditions like depression and anxiety have increased since before the Covid-19 pandemic, referrals to mental health services decreased, and the number of people presenting in crisis increased [148]. It is estimated that the economic cost for the UK due to poor mental health is greater than NHS spending on mental health services [149]. Reducing the impact of common mental illness will increase our national income and productivity.

The NHS Long Term Plan aims to promote digitally-enabled care, improve access to mental health support for 1.9 million people, and support NHS Trusts at meeting their referral waitlist key performance indicators [18]. Digital health tools have the potential to support these goals by providing a digital early intervention approach that can be implemented at early stages in the mental health service pathway, reducing the need for inappropriate assessments and escalating cases based on clinical outcome measures as needed. There is already a substantial number of DHIs that have been developed to support mental health concerns; a recent meta-analysis identified a wide variety of intervention types, including web- and mobile-based therapy, online peer support, therapist contact, virtual reality, and digitally-delivered cognitive training, aiming to address a range of mental health challenges (including substance abuse, mood and anxiety, and many specific conditions) [150]. The review found evidence supporting the potential efficacy of digital mental health tools, but identified a lack of consistency and comprehensiveness in the research and a need for further investigation, particularly around health equity considerations [150].

This second case study examined whether a mobile AI-enabled self-help tool (“Wysa”) could improve patients’ mental health and well-being while waiting for Talking Therapies compared to no intervention. Patients’ engagement with the app and their perceptions of its acceptability was also examined.

1.5.3 Post-surgical follow-up

*This section was previously published in [151].

In the UK, the ageing population is causing an increased demand for health care services that is exceeding clinical capacity [152]. With staff pressured to deliver high-volume workloads, the resulting burnout crisis is increasing costs for health care providers. Demand has been further exacerbated by the COVID-19 pandemic, as the widespread cancellation of elective care has created a large backlog of clinical work [153,154]. Digital technologies have the

potential to help address workforce constraints [155] by automating repetitive, lower-skill tasks. Routine ‘high volume, low complexity’ clinical pathways account for many waiting list backlogs [156] and offer an excellent opportunity for improved efficiency. One example of an area where increased efficiency could have a substantial benefit for healthcare providers and patients is ophthalmology. Ophthalmology has the highest outpatient clinic numbers and performs the most common surgery in the UK - cataract surgery [157–161]. Cataract surgery has low complication rates (1-2% [162–164]), and although postoperative review is recommended, it does not need to be hospital-based [165]. This makes cataract surgery follow-up assessment an ideal test case for automation of routine clinical interactions, as it could free up clinicians’ time for skill-demanding tasks and enable the prioritisation of patients who more urgently require clinical review and intervention [151].

This third case study examined the safety, accuracy, and patient perceptions of engagement with an autonomous conversational agent post-cataract surgery follow-up assessment (“Dora”, delivered via telephone).

1.5.4 Importance of engagement for patient benefit in these clinical cases

This section highlighted the potential clinical impact of this research by demonstrating how the mismatch between clinical demand and capacity can have negative impacts on patients’ ability to access timely and high-quality care in three different clinical areas. DHIs have the potential to help fill these gaps between demand and available resources by supporting preventative health behaviour change, empowering patients to self-manage their health in periods where they have no clinical contact and by improving efficiency by introducing automation and enabling prioritisation of patients through remote data collection and analysis. Without effective engagement - sufficient to achieve the aims of the interventions [14] - these potential benefits for patients and healthcare systems cannot be realised. This challenge is commonly observed in digital health research and can only be addressed through a better understanding of how and why users engage with DHIs and how intervention design can best support that engagement.

1.6 Research Question, Aim, and Objectives

1.6.1 Development of the research question

The previous sections have demonstrated the current state of the literature and argued why this research is important to conduct, from both a theoretical and applied perspective. They have demonstrated that engagement is a necessary component for DHIs to achieve their

behavioural and health aims. This section outlines how the research question was developed based on this investigation of previous theory around the concept of engagement with DHIs and future areas for investigation identified in recent literature [37,166,167]. To summarise what was already known about the conceptualisation and importance of engagement in digital health and to clarify remaining gaps in our understanding, I outlined key theoretical propositions based on previous research:

1. Various health-related behaviours (e.g. diet, physical activity, substance abuse, meditation) have the potential to influence individuals' health outcomes [168,169], but for many reasons (relating to capability, opportunity, and motivation [53,170]), people do not always engage in positive health behaviours.
2. DHIs have the potential to increase positive health behaviours, but the evidence for their effectiveness is mixed [2,116,171,172] and a key limiting factor of their potential impact is a lack of sufficient engagement to achieve the interventions' aims [12,14,31,33].
3. Engagement is a multifaceted concept comprising affective, cognitive, and behavioural engagement with a digital health intervention (micro) and target behaviour (macro) [14,42,45,49–51].
4. Engagement is positively associated with behavioural and health outcomes (as a whole [12,30,31] and as distinct components [35,37]).
5. During the initial stages of engagement, some research has suggested that affective engagement comes first, followed by cognitive engagement, and then behavioural engagement [49,66,173], suggesting that different engagement strategies may help support different components of engagement ([Section 3.7](#)).

These theoretical propositions, derived from the literature, highlight a two-stage problem that needs to be addressed: (1) how to support engagement with health behaviours and (2) how to support engagement with digital tools designed to support (1). Recent research has made strides in clarifying how engagement can be conceptualised within the field of digital health, but our understanding of how different components of engagement interact during a user's process of engagement with a DHI is still limited [49]; this is the gap that this thesis will aim to address.

To address this gap, I selected an exploratory, mixed-methods multiple-case study method. The justification for the selection of this study design is detailed in [Section 4.2](#); briefly, it

enabled an in-depth exploration of the process of users' engagement with DHIs. Examining multiple cases also allowed for comparison of the primary outcome - the associations between different components of engagement - across different types of digital intervention. This meant that the relationships could be examined in more detail and potential factors influencing them could be generated and tested against the other case studies to deepen our theoretical understanding and test generalisability. Specifically, the research question was defined as:

How are the different components of engagement (affective, cognitive, behavioural, micro, and macro) related in digital behaviour change app interventions for weight management and mental health and in an autonomous telemedicine post-surgical follow-up assessment?

1.6.2 Aims and objectives

The main aim of the thesis will be to advance the theoretical understanding of how people engage with different types of DHIs, to inform the development of digital tools that can achieve their intended impact. This thesis addresses a knowledge gap concerning how the different components of engagement are related to each other and how those relationships vary over time. Understanding how these components interact over the course of engagement with a DHI will enable the development of digital health solutions that use more tailored strategies to support the engagement needed to achieve the intervention aims.

To achieve this aim, the thesis research objectives are:

1. To map the different components of users' engagement with digital health interventions and how they influence each other;
2. To compare engagement with different types of digital health interventions to determine if identified processes of engagement are generalisable;
3. To generate evidence-based recommendations for strategies to optimise engagement at different stages.

1.7 Structure of Thesis

This section provides an overview of the structure of the rest of the thesis, which is organised into 7 main chapters. Chapter 1 detailed the rationale for the thesis, based on the theoretical background and clinical need, and described how this shaped the thesis' aims, objectives, and

research question. The remaining chapters examine this question using various methods and from different perspectives (Figure 1-2).

Chapter 2 provides a systematic review of the literature to summarise what is already known about the associations between different components of engagement and specific active ingredients of DHIs (Behaviour Change Techniques) and to further clarify the gaps in knowledge.

Chapter 3 examines the conceptualisation and measurement of the different components of engagement through a comparison of two self-report scales that aim to measure engagement as a multi-faceted concept and combines these insights with previous research to generate definitions for each of the components and to hypothesise a theoretically-based mapping of how the different components of engagement are related.

Chapter 4 provides an overview of the theoretical frameworks and methods used in the overall thesis and in the three case studies (see Figure 4-1 in [Section 4.2](#) for a more detailed diagram of how the case studies fit into the overall thesis methodology).

Chapter 5 presents the results and key insights from the within-case analyses of each case study and develops empirically-based mappings of the different components of engagement from the analysis of each case study to enable the cross-case analysis.

Chapter 6 addresses the thesis research question by comparing the mappings from these three cases in a cross-case analysis and conducts a thematic synthesis to generate factors influencing these components and their associations across different types of DHI.

Chapter 7 discusses the principal findings, the study's strengths and limitations, implications for future research, recommendations for digital intervention development, reflections on the process of conducting the thesis, and the conclusions and contributions of the thesis to knowledge.

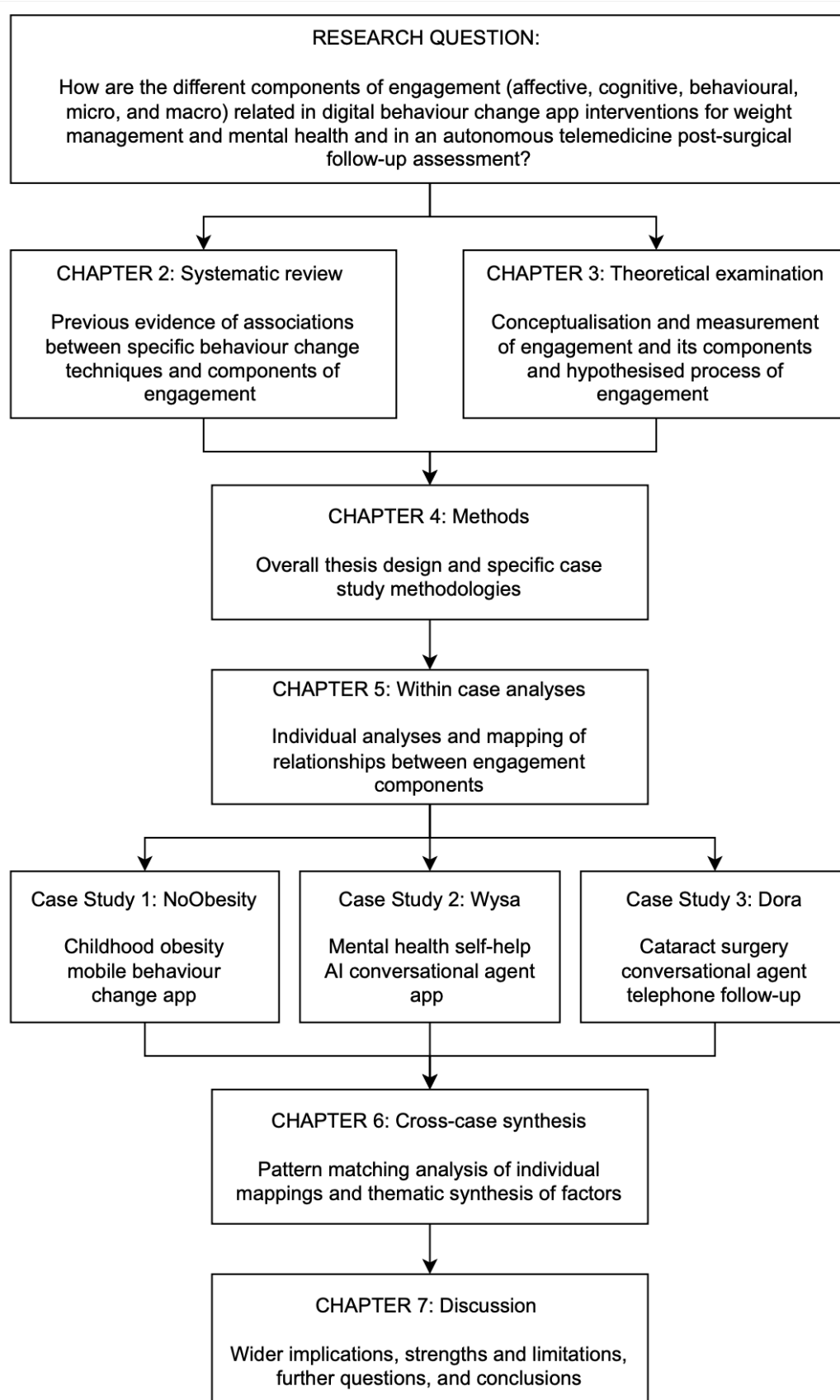


Figure 1-2. Thesis structure diagram

Chapter 2: Systematic Review

2.1 Chapter Overview

Chapter 1 of this thesis introduced the background and rationale for the study, specified the research questions and objectives, and provided an overview of the thesis structure. It demonstrated the gap in the literature relating to our understanding of how different components of engagement influence each other over the process of digital health engagement. The overall aim of the thesis is to improve our understanding of how users engage with digital health interventions (DHIs) and to explore how that can be optimised through theory-based design. The current limitations of evaluations of digital health engagement that were briefly highlighted in [Section 1.2](#) make it difficult to get a good understanding of the state of the literature on relationships between the different components of engagement, as they are rarely examined explicitly (as will be discussed in more detail in Chapter 3). Instead, the systematic review of the literature focused primarily on examining what is known around the second part of the aim - how theory-based intervention features can support digital health engagement - and used that analysis to explore the state of the literature around components of engagement and highlight both gaps and existing knowledge.

Chapter 2 presents an update to a published systematic review and provides an analysis of the literature to summarise and synthesise what is known about the potential associations between Behaviour Change Techniques (BCTs) - theory-based components of behavioural change interventions - and engagement with mobile health apps. A systematic review methodology [174,175] was used to conduct a comprehensive investigation of the available literature with the aim of identifying BCTs that might be likely to support engagement and how associations between specific BCTs and components of engagement are being measured in digital health evaluations. The following sections detail the background and rationale for this review and outline the methods, which follow the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines. The results and analyses are presented and interpreted in the context of the wider literature, the strengths and limitations of the review, and the gaps in the available data.

Publications: As of the thesis submission, Chapter 2 has been drawn from a peer-reviewed published protocol and a systematic review submitted for peer-reviewed publication.

Milne-Ives M, Homer S, Andrade J, Meinert E. Associations Between Behavior Change Techniques and Engagement With Mobile Health Apps: Protocol for a

Systematic Review. *JMIR Research Protocols*. 2022 Mar 29; 11(3):e35172.
doi:[10.2196/35172](https://doi.org/10.2196/35172)

Milne-Ives M, Homer S, Andrade J, Meinert E. Potential associations between behavior change techniques and engagement with mobile health apps: a systematic review. *Frontiers in Psychology*. 2023 Sep 18; 14:1227443.
doi:[10.3389/fpsyg.2023.1227443](https://doi.org/10.3389/fpsyg.2023.1227443)

2.2 Background

*This section has been previously published in [1,47].

Mobile health applications have the potential to empower people to improve their health behaviours and self-manage health conditions [113,114,176]. They are widely available, with a large global market [105], and provide a means of delivering far-reaching behavioural interventions. To have a significant impact on behavioural and health outcomes, mobile health apps need to be able to support sufficient engagement to achieve the aims of the intervention they are delivering [14,50]. Previous research has established an association between engagement with digital interventions and their impact on intended outcomes [31,37,42], demonstrating the importance of a certain degree of engagement for the efficacy of an intervention. Maintaining engagement is a common challenge for mobile health apps [8–13] and is a potential reason why evidence of their effectiveness remains mixed [2,116,171,172].

The conceptualisation of engagement has recently become a focus in digital health research, resulting in several theories and frameworks (Figure 1-1) [14,42,45,49–51]. As described in [Section 1.3](#), it is generally accepted to be multifaceted, with components relating to dimension (ie. affective, cognitive, and behavioural) and scale (micro and macro) [1]. Although evidence is still developing, previous research has demonstrated associations between both objective (behavioural) and subjective (affective and cognitive) engagement and clinical outcomes [35] and has emphasised the importance of incorporating behavioural theory into DHI design [32,115,176–178]. Incorporating theory enables evidence-based predictions to optimise context-specific intervention design and potential impact [121,176,179]. A variety of strategies have been proposed and incorporated into mobile health apps to help improve user engagement, including design features and BCTs [180–182], but their impact on engagement is often measured using only behavioural measures of app use [30,37]. As these measures cannot capture the full picture of users' multifaceted engagement, our understanding of how different components of engagement are associated with app features, engagement strategies, and outcomes remains incomplete.

Despite the growing presence of BCTs [179] in mobile health apps, there has been limited investigation of how they can help support engagement specifically. A preliminary review of the literature [1] identified several systematic reviews that have explored questions around this topic [122,183–185], but none that analysed how BCTs can best support the various components of engagement. One of these reviews [184] structured their analysis around the Capability, Opportunity, Motivation - Behaviour (COM-B) model, which is part of the Behaviour Change Wheel theoretical framework [53]. The authors identified 26 different factors relating to capability, opportunity, and motivation that have been associated with uptake of and engagement with mobile health apps in the literature [184]. This provides a valuable, theory-based contribution to the understanding of factors affecting engagement with mobile health apps. However, although the review included studies with either qualitative and quantitative (primarily system use data) measures of engagement and used a multi-faceted definition of engagement [42], there was no consideration of how the influence of these factors varied for the different components and measures of engagement. Additionally, PROSPERO was searched using various combinations of the keywords ‘engagement’, ‘digital health interventions’, ‘digital behaviour change interventions’, ‘behaviour change techniques’, ‘BCT’, ‘mobile health apps’, ‘mHealth’, ‘eHealth’, and ‘digital behaviour change’. None of the registered protocols aimed to examine the associations between BCTs and the different components of engagement; however, the search terms identified the PROSPERO pre-registration for one of the previous reviews cited in this rationale [184], indicating that the search terms were appropriate.

The purpose of this review was to address this gap in the literature by investigating how specific BCTs are associated with the different components of engagement (affective, cognitive, behavioural, micro, and macro) to provide insights for the development of mobile health apps. To examine this, the primary objective was to identify BCTs being incorporated in the design and development of mobile health apps and their hypothesised or evidenced associations with engagement and its various components. Learnings from this review will enable future studies to empirically test causal relationships between specific BCTs and specific components of engagement to further improve our understanding of their associations.

2.3 Methods

*Parts of this section have been previously published in [1,47].

2.3.1 Study design

The Preferred Reporting Items for Systematic Reviews and Meta-analysis Protocols (PRISMA-P) framework [186] was used to structure the protocol for this review ([Appendix 2-A](#)). The protocol was prospectively registered on PROSPERO (CRD42022312596) and published before searches were conducted on 12 July 2022 [1].

2.3.2 Eligibility criteria

The eligibility criteria and search strategy were based on the Population, Intervention, Comparator, Outcomes, Study (PICOS) framework (Table 2-1) [187,188]. The protocol scope was broadened to include development studies to capture research that hypothesised potential associations. Studies published before 2012 were excluded because I wanted to provide an overview of recent evidence (digital technology evolves rapidly [189]) and because the conceptualisation of engagement as a multifaceted construct emerged in digital health relatively recently [45].

Table 2-1. PICOS framework

Component	Inclusion criteria	Exclusion criteria
Population	Mobile health app users of any age (adults and children)	
Intervention	Mobile health apps developed using at least one behaviour change theory, framework, or technique to target at least one health behaviour (including but not limited to: physical activity, diet, sedentary behaviour, substance use, etc.)	Studies with no description of behavioural theory or BCT; authors needed to be able to meaningfully identify behaviourally-based features if they were not reported as BCTs
Comparator	No comparator is required	
Outcomes	App engagement (hypothesised or evidenced), defined broadly to include any components of engagement (affective, cognitive, behavioural;	

	relating to the intervention interface, components, or target health behaviour) and any type of measurement (quantitative or qualitative). Secondary outcomes will include the BCTs and theories incorporated in the apps, qualitative or quantitative engagement outcomes measured (including any components of engagement specified by a theoretical framework) and the behavioural and health outcomes reported.	
Study types	Studies that describe the design, development, or evaluation of digital health behaviour change interventions that aim to support user engagement (including randomised controlled trials, quantitative, qualitative, cohort, and case studies)	Reviews, protocols, and conference abstracts or posters without full texts

2.3.3 Search strategy

PubMed, Embase, Cumulative Index to Nursing and Allied Health Literature (CINAHL), APA PsycInfo, ScienceDirect, and Web of Science were searched using the search structure (Table 2-2, [Appendix 2-B](#)): engagement (MeSH OR Keywords) AND digital health interventions (MeSH OR Keywords) AND behaviour change (MeSH OR Keywords).

Table 2-2. Search terms

Category	MeSH ^a	Keywords (in full text)
Engagement	Treatment Adherence and Compliance OR Patient Participation OR Patient Compliance	Engag* OR “user engagement” OR immersion OR flow OR involvement OR presence OR adherence OR attrition compliance OR maintenance OR acceptability OR satisfaction

Digital health interventions	Telemedicine OR Mobile Applications OR Internet-Based Intervention	“mHealth” OR “eHealth” OR “mobile health” OR “eHealth” OR telehealth OR mobile OR phone OR smartphone OR cell OR digital OR “app” OR “apps” OR application* OR digital OR web OR internet
Behaviour change	Behavior Control OR Psychological Theory	“behaviour change techniques” or “behaviour change techniques” or “BCT” or “behaviour change technique” or “behaviour change technique” or “behavioural change strategies” or “behavioural change strategies” or “behaviour change wheel” or “behaviour change wheel” or “behavioural theory” or “behavioural theory” or “behaviour change theory” or “behaviour change theory” or “health behaviour change” or “behaviour change” or “behaviour change” or “digital behaviour change intervention” or “digital behaviour change intervention” or “DBCI” or “behaviour change intervention”

2.3.4 Screening and article selection

References were imported into EndNote X9 for duplicate removal and initial screening using the EndNote search function ([Appendix 2-C](#)). I screened titles and abstracts of remaining references in Rayyan and then conducted a full-text review.

2.3.5 Data extraction

Data was extracted into a predetermined form (see Textbox 2-1).

Textbox 2-1. Article information and data extraction

General study information

- Year of publication
- Country of study
- Sample demographics (eg. age, gender, target population)
- Initial sample size

- Analysed sample size
- Study duration

Mobile health app behavioural intervention

- App name
- Operating platform (iOS, Android, web)
- Target health behaviour
- Aim of the intervention
- Behavioural theory used
- Number of included Behaviour Change Techniques [179] (if any)
- List of included Behaviour Change Techniques [179] (if any)
- Other app engagement features

Evaluation

- Component(s) of engagement examined
- Engagement outcome measures
- Effect of intervention on engagement outcomes
- Effect of intervention on behavioural outcomes (if reported)
- Effect of intervention on health outcomes (if reported)

2.3.6 Data analysis

A descriptive analysis summarised study and intervention characteristics. A quantitative meta-analysis of effect was not possible due to outcome variety, so extracted evidence was analysed by one author to map potential associations between BCTs and engagement. Evidence for potential associations was divided into three categories: evidenced, hypothesised, or inferred. Associations were inferred if studies did not directly examine associations between features and engagement. Where studies did not explicitly report the included BCTs, features were coded using the BCT Taxonomy [179] as generically as possible (e.g. ‘non-specific’ reward) to ensure that they could be analysed in line with the rest of the included studies and the included BCTs were marked as ‘inferred’ rather than ‘explicit’. To examine components of engagement, a basic codebook was created using definitions from the literature ([Appendix 2-D](#)) to explore how they were assessed and associated with BCTs. The systematic review was conducted prior to Chapter 3, where the main definitions of the components of engagement were developed. While the codebook largely reflects the final definitions included in the thesis, there is a lack of consideration of

how concepts like interest and motivation can relate to both affective and cognitive engagement. Given the limitations both with this method of analysis and the available data, this analysis was exploratory.

2.3.7 Risk of bias and quality assessment

The quality appraisal was conducted using the Cochrane Collaboration Risk of Bias 2 (RoB 2) tool for randomised controlled trials (RCTs) [190,191] and the Mixed-Methods Appraisal Tool (MMAT) [192]. The MMAT tool was used instead of the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool [193] (which was initially planned in the published protocol) because it could be applied to the wide range of study types included in the review.

2.3.8 Systematic review update

An updated search was conducted on 15 January 2024 to identify studies published in the interim that met the eligibility criteria and could provide additional data for the analysis ([Appendix 2-E](#)). The update was initially planned to be a simple re-run of the search and screening process described above. Since I executed the initial systematic review, I have conducted and advised on several other reviews; in the process of re-running the EndNote screening, I identified terms and passes that I felt overly restricted the results at too early a stage in the screening process. To address this limitation, I amended the EndNote passes to be more inclusive and to ensure that references imported to EndNote without abstracts were not excluded by the pass structure ([Appendix 2-F](#)): additional keywords were included in some of the passes to capture more terms and title/abstract/keywords were searched rather than just title or abstract. After the final pass, as in the initial search, duplicates were removed, as were references that had previously been screened in the initial review. The titles and abstracts of the remaining references were screened manually in Rayyan, and full-texts of the relevant articles were screened and excluded with reasons. Data extraction followed the same form (Textbox 2-1).

2.4 Results

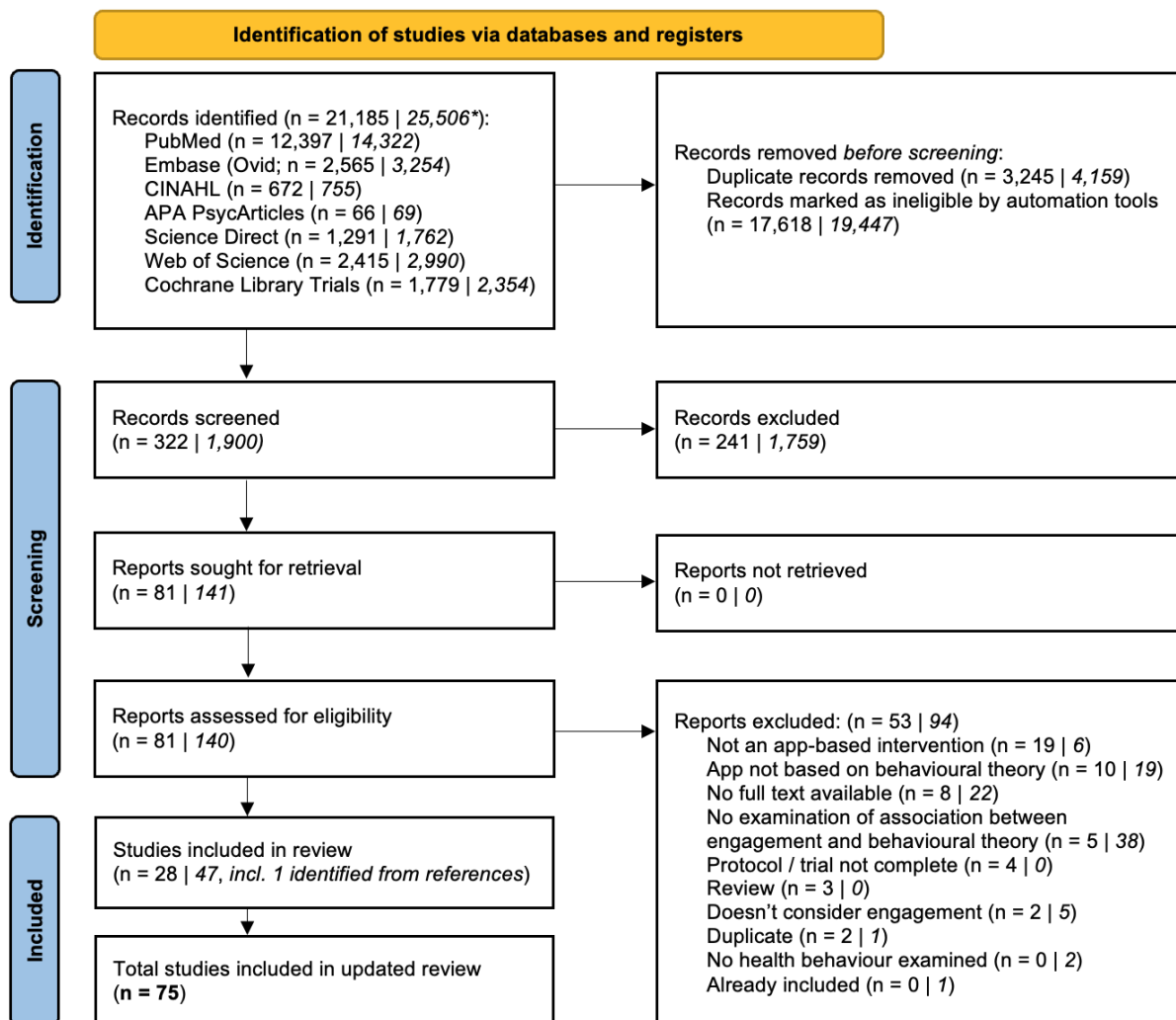
*This section has been updated from a previously published paper [47].

2.4.1 Included studies

In the first execution of the systematic review, 21,185 articles were retrieved and 28 were determined to be eligible for inclusion. The reasons for exclusion in the full-text review stage

are included in the PRISMA flow diagram (Figure 2-1). One of the included references was a Clinicaltrials.gov registration from which a full text article was identified. Rerunning the search 18 months later retrieved an additional 4,321 references, for a total of 25,506. The full texts of 140 articles were reviewed; 47 articles were identified as eligible for inclusion (including a subsequent analysis of an included development paper that was identified during full-text review and added), 19 of which were published in 2022 or later.

Figure 2-1. PRISMA flow diagram



*Reference counts are presented as (n = original | *update*) to provide a breakdown of the differences between the two screening processes, rather than simply providing the final totals

2.4.2 Study characteristics

The characteristics of the 75 included studies are summarised in [Appendix 2-G](#). Qualitative and randomised studies were most common, each accounting for approximately a quarter of included studies (19/75 and 18/75, respectively), although not all randomised studies were

traditional RCTs (e.g. pilot RCTs, micro-randomised trials). Reported study durations ranged from 1-2 weeks for pilot studies and feasibility testing [194,195] to 18-20 months for real world app use evaluations [196–198]. Sample sizes ranged from 6 [199] to almost 120,000 (real world app use evaluation) [200]. 70% of the studies had fewer than 100 participants (49/70, as not all studies included participants). Two thirds of the studies (52/75) focused on adults only.

2.4.3 App characteristics

67 individual apps were examined (Table 2-3) across the studies, with some studies including multiple apps [197,201–206]. The studies referenced a total of 30 unique behavioural theories (Table 2-4), with a fifth (16/75) referencing more than one. Just over half the studies (40/75) explicitly reported BCTs. Around a fifth of studies (17/75) referenced BCTs but no particular behavioural theory [182,200,202,203,205,207–218]. Among the 75 studies, 67 unique BCTs were identified as included in, or considered for inclusion in the development of, the apps.

Table 2-3. Summary of app characteristics

First author (year)	App name	Operating platform	Target health behaviour	Aim of intervention	Behav. theory ^a
<i>Originally included studies</i>					
Ahonkhai (2021) [219]	PEERNaija	Android	Medication adherence	Adherence	SCT
Baumel (2018) [197]	Many included	Android and web	Several (mental health, diet, physical activity, reduce substance use)	Self-management of health	Not specified
Buman (2016) [220]	BeWell24	Android	Sleep, sedentary, and physical activity behaviours	Improve cardiometabolic health outcomes	BCTs, SCT, sleep education, sleep hygiene, stimulus control therapy, self-regulation theory
Caon (2022) [221]	PEGASO Companion (eDiary)	Android	Dietary behaviours	Weight management	COM-B model

Cueto (2019) [211]	Kurbo	Not specified	Dietary and physical activity behaviours	Weight management	BCTs
Curtis (2019) [222]	MyMate&Me	Not specified	Medication adherence	Adherence	BCW
D'Addario (2020) [208]	Not specified	Not specified	Physical activity behaviour	Increase physical activity	BCTs
Davis (2019) [209]	1) Noom Walk Pedometer (NWP; low BCT) 2) MapMyFitness (MMF; high BCT)	iOS	Physical activity behaviour	Increase physical activity	BCTs
Delmas (2020) [196]	AirForU	iOS and Android	Air quality related behaviours	Educate about air pollution and protective behaviours	Theory of Issue Engagement, TPB
DeSmet (2019) [210]	Not specified	Not specified	Sleep, sedentary, and physical activity behaviours	Increase physical activity (and sleep, and decrease sedentary behaviour)	BCTs
Du (2016) [223]	Fittle	iOS and Android	Dietary and physical activity behaviours	Adherence (to healthy habits)	TPB, SCT
Duff (2018) [195]	MedFit App	Android	Physical activity behaviours	Self-management of health	SCT, BCW
Garnett (2015) [182]	Not specified	Not specified	Reduce substance use	Reduce alcohol consumption	BCTs
Hochsmann (2019) [224]	Mission: Schweinehund	iOS and Android	Physical activity behaviour	Increase physical activity	SDT
Hales (2016) [225]	Social Pounds Off Digitally (POD)	Android	Dietary and physical activity behaviours	Weight management	SCT
Lin (2021) [226]	Aim2Be	Not specified	Several (dietary and physical activity behaviours, screen time, sleep)	Adherence (to healthy habits)	SCT, SDT

Nuijten (2021) [227]	GameBus	iOS and Android	Several	Adherence (to healthy habits)	Social comparison model of competition
Nuijten (2022) [228]	GameBus	iOS and Android	Physical activity behaviours	Increase physical activity	COM-B, Fogg Behavior Model, TPB, SDT, Flow theory ^b
Nurmi (2020) [229]	Precious app	Android	Physical activity behaviours	Adherence (to healthy habits)	SDT
Smith (2022) [123]	Stay-Active	Not specified	Physical activity behaviours	Increase physical activity	BCW, COM-B model, TDF
Songtaweessin (2020) [230]	P3 (Prepared, Protected, emPowered) Thailand app	iOS and Android	Medication adherence	Adherence	SCT, Narrative Communication, Fogg behavioural model
Sporrel (2020) [231]	Playful data-driven Active Urban Living (PAUL)	Android	Physical activity behaviours	Increase physical activity	BCW, Behavioral Intervention Technology model
Szinay (2021) [201]	Many included	Not specified	Several	Not specified	COM-B model, TDF
Tang (2015) [202]	Many included	Not specified	Dietary and physical activity behaviours	Weight management	BCTs
Thornton (2021) [207]	Health4Life App	iOS and Android	Several (dietary and physical activity behaviours, reduce substance use)	Adherence (to healthy habits)	BCTs
Valle (2020) [212]	Nudge	iOS	Dietary and physical activity behaviours	Weight management	BCTs
Whiteley (2018) [232]	Battle Viro	iOS	Medication adherence	Adherence	IMB model
Whiteley (2019) [233]	Viral Combat	iOS	Medication adherence	Adherence	IMB model
<i>Additional studies included after the update</i>					

Abdul Khalil (2023) [234]	N/A	N/A	Dietary behaviours	Increase self-monitoring (diet)	TPB
Attwood (2017) [200]	Drinkaware	Android and iOS	Reduce substance use	Reduce alcohol consumption	BCTs
Beck (2021) [235]	SMART Track	Android and iOS	Reduce substance use	Self-management of health (addiction recovery)	SDT, Social Control Theory, BCTs
Bell (2020) [198]	Drink Less	Not specified	Reduce substance use	Reduce alcohol consumption	COM-B, TDF, BCTs
Bell (2023) [236]	Drink Less	Not specified	Reduce substance use	Reduce alcohol consumption	COM-B, TDF, BCTs
Brindal (2019) [237]	MotiMate	Android and iOS	Several (dietary and physical activity behaviours, mental health)	Weight management	Conservation of resources and self-regulation theory, Health Action Process Approach (HAPA)
Burgermaster (2023) [238]	Platano	Android and iOS	Dietary behaviours	Self-management of health (diabetes)	SDT
Chang (2023) [239]	NeuroFlow	Not specified	Mental health	Improve mental well-being	Behavioural economic principles
Curtis (2015) [240]	N/A	N/A	Dietary behaviours	Weight management	BCW, TDF
Daniels (2023) [241]	MIA (More in Action)	N/A	Physical activity behaviours	Increase physical activity	BCW
Direito (2015) [203]	1) Zombies, Run 2) Get Running	Android and iOS	Physical activity behaviours	Increase physical activity and cardiorespiratory fitness	BCTs
Domin (2022) [213]	N/A	N/A	Physical activity behaviours	Increase physical activity	BCTs

Edney (2020) [242]	Active Team	Android and iOS	Physical activity behaviours	Increase physical activity	SCT
Edwards (2018) [243]	Cigbreak	Android and iOS	Reduce substance use	Reduce / stop smoking	BCW, BCTs
Ehlers (2014) [244]	N/A	N/A	Physical activity behaviours	Increase physical activity	SCT
Ekezie (2021) [245]	Baby Steps	Web app	Physical activity behaviours	Self-management of health (diabetes)	Not specified
Fatima (2023) [246]	OnTheMove!	Android and iOS	Physical activity behaviours	Adherence (to healthy habits)	SDT, BCTs
Flaherty (2021) [205]	App A and App B (names not specified)		Dietary behaviours	Support healthier food behaviours	BCTs
Garnett (2019) [247]	Drink Less	Not specified	Reduce substance use	Reduce alcohol consumption	COM-B, TDF, BCTs, Social Norms Theory
Hendrie (2019) [248]	VegEze	iOS	Dietary behaviours	Support healthier food behaviours (increase vegetable consumption)	BCW, COM-B, BCTs
Hendrie (2020) [194]	VegEze	iOS	Dietary behaviours	Support healthier food behaviours (increase vegetable consumption)	BCW, COM-B, BCTs
Kekkonen (2023) [249]	Not specified	Android	Several (dietary and physical activity behaviours, mental health, sleep, etc.)	Adherence (to healthy habits)	SDT, transtheoretical model, persuasive system design
Lawrason (2023) [250]	SCI Step Together	Not specified	Physical activity behaviours	Increase physical activity	SDT, BCTs
Lim (2022) [251]	nBuddy Diabetes	Not specified	Dietary and physical activity behaviours	Weight management and glycemic control	Obesity-Related Behavioral Intervention Trials

					(ORBIT) Model; self-regulation theory
Litterbach (2017) [252]	Growing healthy program	Android and iOS	Dietary behaviours	Promote healthy infant feeding	BCW
Lubbans (2014) [253]	“Active Teen Leaders Avoiding Screen-time” (ATLAS)	Android and iOS	Dietary and physical activity behaviours	Adherence (to healthy habits)	SCT, SDT
McClure (2023) [254]	GEMS - standard vs. enhanced	Android and iOS	Reduce substance use	Reduce / stop smoking	SCT, PRIME theory of motivation, cognitive behavioral therapy, acceptance and commitment therapy
Montilva-Monsalve (2023) [255]	Gro Health	Android and iOS	Dietary behaviours	Self-management of health (diabetes)	BCW, COM-B, TDF, BCTs
Mummah (2016) [214]	Vegethon	Not specified	Dietary behaviours	Support healthier food behaviours (increase vegetable consumption)	BCTs
Mummah (2016) [215]	Vegethon	iOS	Dietary behaviours	Support healthier food behaviours (increase vegetable consumption)	BCTs
Nahum-Shani (2021) [256]	Substance Abuse Research Assistant [SARA]	Android or iOS	Reduce substance use	Increase self-monitoring (substance use)	Behavioral theory
Nicholas (2021) [257]	START app (Striding TowARds health and	iOS	Physical activity behaviours	Increase physical activity	SDT, BCTs

	well-being Trial)				
Patterson (2024) [258]	Vire app and ToDo-CR program	Not specified	Physical activity behaviours	Reduce sedentary behaviour	SCT, BCTs
Perski (2017) [204]	N/A (many apps included)	Android and iOS	Reduce substance use	Reduce alcohol consumption or stop smoking	N/A
Pfaeffli Dale (2019) [259]	Carrot Rewards	Not specified	Vaccine uptake	Encourage influenza vaccine uptake	Transtheoretical model
Resnick (2021) [260]	Healthier Together	Android and iOS	Several (dietary and physical activity behaviours, reduce substance use)	Self-management of health (promoting cancer prevention behaviors)	Goal-setting and social network theories
Roberts (2019) [206]	1) Human 2) The Walk 3) The Johnson & Johnson Official 7 Minute Workout 4) Gorilla Workout	Android and iOS	Physical activity behaviours	Increase physical activity	BCTs
Rohde (2019) [261]	Challenge to Go	Android	Dietary behaviours	Support healthier food behaviours	BCW
Sakane (2023) [216]	KENPO-app	Android and iOS	Dietary and physical activity behaviours	Weight management	BCTs
Sanders (2023) [262]	SnackApp	Android and iOS	Physical activity behaviours	Increase physical activity	SDT, BCTs, Control theory, self-regulation theory
Schulte (2022) [217]	Boozebuster	Android and iOS	Reduce substance use	Reduce alcohol consumption	BCTs
Simons (2018) [263]	Active Coach	Android	Physical activity behaviours	Increase physical activity	SDT, BCTs

Stancu (2022) [264]	N/A	N/A	Dietary behaviours	Self-management of health	SDT
Tombor (2019) [265]	SmokeFree Baby	Android and iOS	Reduce substance use	Reduce / stop smoking	COM-B, PRIME theory
Whale (2021) [218]	URApp	Android and iOS	Bladder control	Self-management of health (Daytime Urinary Incontinence)	BCTs
Willms (2023) [199]	Healthy Hearts	iOS	Physical activity behaviours	Self-management of health (hypertension prevention)	Multiprocess Action Control (M-PAC) framework
Zimmermann (2023) [266]	Vida	Not specified	Several (dietary and physical activity behaviours, sleep, etc.)	Self-management of health	SDT

^aBCTs = Behaviour Change Techniques; BCW = Behaviour Change Wheel; COM-B = Capability, Opportunity, Motivation - Behaviour model; HAPA = Health Action Process Approach; IMB model = Information, Motivation, and Behavioral Skills Model; M-PAC = Multiprocess Action Control framework; PRIME theory = Plans, Evaluations, Motives, Impulses, Responses; ORBIT = Obesity-Related Behavioral Intervention Trials model; SCT = Social Cognitive Theory; SDT = Self-Determination Theory; TPB = Theory of Planned Behaviour; TDF = Theoretical Domains Framework

^bDiscussed in paper, but unclear if incorporated in app design

Table 2-4. Behavioural theories used in apps targeting particular health behaviours

Target behaviour	BCTs	SDT	SCT	BCW ^a	COM-B model ^a	TDF	TPB	Fogg behav. model	IMB model	Trans-theoretical model	PRIME theory	Self-reg. theory	Other
Physical activity / sedentary behaviour	18	12	9	4	2	1	2	1	-	1	-	4	9
Dietary behaviour	11	6	4	6	4	2	2	-	-	1	-	2	4
Substance use	9	1	1	1	4	3	-	-	-	-	2	-	6
Sleep	2	3	2	-	-	-	-	-	-	1	-	1	2
Med. adherence	-	-	2	1	-	-	-	1	2	-	-	-	1
Mental	-	1	-	-	-	-	-	-	-	1	-	1	3

health													
Air quality behaviour	-	-	-	-	-	-	1	-	-	-	-	-	1
Screen time	-	1	1	-	-	-	-	-	-	-	-	-	-
Vaccine uptake	-	-	-	-	-	-	-	-	-	1	-	-	1
Bladder control	1	-	-	-	-	-	-	-	-	-	-	-	1
Several (not specified)	-	-	-	-	1	1	-	-	-	-	-	-	1
Total	41	24	19	12	11	7	5	2	2	5	2	8	27

^aThe core of the BCW is the COM-B model, but they are reported separately here in line with what the studies reported

Note: An occurrence was added each time a behavioural theory was reported in a study for each of the target behaviours supported by that app; therefore, columns and rows do not necessarily add up to the number of studies, as several studies included multiple target behaviours or behavioural theories.

2.4.4 Engagement measures

The most common methods of measuring engagement were through app usage data (e.g. frequency, duration) or qualitative methods (33/75 and 32/75 studies, respectively). Other methods included surveys, app-collected data (active or passive), data from synced devices (e.g. activity trackers and smart pill bottle caps), and discrete choice experiments [264]. Many of the studies that used qualitative methods were focused on user-centred design or early-stage usability or feasibility testing. There were a diverse range of questionnaires used to assess engagement; of the 25 studies that used surveys, 13 used surveys developed for that particular study. Ten different validated scales were used; the most common was the System Usability Scale (used 4 times), followed by the user version of the Mobile App Rating Scale (used 3 times; Table 2-5).

Table 2-5. Summary of engagement measures used in the studies

First author (year) ^a	Intervention data (freq., duration)	Behav. data self-reported on app	Behav. sensor data (app or synced device)	Survey	Qualitative
<i>Originally included studies</i>					
Baumel (2018) [197]	X	-	-	Enlight Quality Ratings (EQR)	-

Buman (2016) [220]	X	-	-	-	-
Caon (2022) [221]	X	X	-	-	-
Cueto (2019) [211]	X	X	-	-	-
D'Addario (2020) [208]	-	-	-	-	X
Davis (2019) [209]	-	-	-	User Mobile App Rating Scale (uMARS)	-
Delmas (2020) [196]	X	X	-	-	-
DeSmet (2019) [210]	-	-	-	Validated scales to measure intention- to-change	-
Du (2016) [223]	-	X	-	-	-
Duff (2018) [195]	-	-	-	-	X
Hales (2016) [225]	X	X	-	-	X
Hochsmann (2019) [224]	X	-	X	Intrinsic Motivation Inventory (IMI)	-
Lin (2021) [226]	X	-	-	-	-
Songtaweessin (2020) [230]	-	-	-	-	X
Nuijten (2021) [227]	X	X	-	-	-
Nuijten (2022) [228]	X	X	X	-	-
Nurmi (2020) [229]	-	-	-	-	X
Szinay (2021) [201]	-	-	-	-	X
Tang (2015) [202]	-	-	-	-	X
Thornton (2021) [207]	-	-	-	System Usability Scale (SUS)	X
Valle (2020) [212]	X	X	X	-	-
Whiteley (2019) [233]	-	-	-	Client Service Questionnaire (CSQ)	X
<i>Additional studies included after the update</i>					
Abdul Khalil (2023) [234]					X
Attwood (2017) [200]	X	X		Study-specific survey	X

Bell (2020) [198]	X	X			
Bell (2023) [236]	X				
Brindal (2019) [237]	X			Study-specific survey based on Technology Acceptance Model	
Burgermaster (2023) [238]					X
Chang (2023) [239]	X				
Curtis (2015) [240]					X
Daniels (2023) [241]				User Experience Questionnaire (UEQ)	X
Direito (2015) [203]				Study-specific survey	
Domin (2022) [213]					X
Edney (2020) [242]	X			System Usability Scale (SUS); study-specific survey	
Edwards (2018) [243]				Study-specific survey	X
Ehlers (2014) [244]				Study-specific survey based on Social Cognitive Theory and Unified Theory of Acceptance and Use of Technology (UTAUT2)	
Ekezie (2021) [245]					X
Fatima (2023) [246]	X			System Usability Scale (SUS); Usefulness, Satisfaction and Ease of Use (USE) Questionnaire; study-specific survey	
Flaherty (2021) [205]		X			X ^b
Hendrie (2019) [248]	X			Study-specific survey	
Hendrie (2020) [194]	X	X			

Kekkonen (2023) [249]					X
Lawrason (2023) [250]	X				X
Lim (2022) [251]	X				
Litterbach (2017) [252]					X
Lubbans (2014) [253]					X
McClure (2023) [254]	X	X		Study-specific survey	
Montilva-Monsalve (2023) [255]					X
Mummah (2016) [214]				Study-specific survey	X
Mummah (2016) [215]	X	X		Study-specific survey, adapted from previously developed user satisfaction surveys	
Nahum-Shani (2021) [256]	X	X		Study-specific survey	
Nicholas (2021) [257]	X	X		User Version of the Mobile Application Rating Scale (uMARS)	X
Patterson (2024) [258]	X				X
Perski (2017) [204]					X
Pfaffli Dale (2019) [259]	X				
Resnick (2021) [260]	X	X		System Usability Scale (SUS)	X
Roberts (2019) [206]					X
Sakane (2023) [216]	X				
Sanders (2023) [262]	X			User Version of the Mobile Application Rating Scale (uMARS)	

Schulte (2022) [217]				TWente Engagement with Ehealth Technologies Scale (TWEETS)	
Simons (2018) [263]				Study-specific survey	X
Stancu (2022) ^c [264]					
Tombor (2019) [265]	X				
Whale (2021) [218]					X
Willms (2023) [199]				mHealth App Usability Questionnaire (MAUQ)	X
Zimmermann (2023) [266]	X				
Total	33	17	3	25	32

^a9 papers are absent from the table, as they did not include engagement measures

[123,182,219,222,231,232,235,247,261]

^bThis evaluation included an observational study of app use

^cThis evaluation conducted a discrete choice experiment to determine users' feature preferences

2.4.5 Associations between Behaviour Change Techniques and engagement

[Appendix 2-H](#) details the analysis of the potential associations between BCTs and engagement, in the context of the available descriptions and quality of the evidence, highlighting key data from the papers that shaped the interpretation of evidence of associations. Methods in the 'evidenced' group of studies included comparisons between different mobile conditions (e.g. apps with and without social support) [203,216,217,223,227,228,236,237,239,242,253,265], rankings of BCT effectiveness for engagement by experts or users [182,201,202], modelling [196,226], qualitative user perceptions [206,213,245,252,255,260], and observational studies of app use or user preferences [198,205,244,251,259,260]. Overall, the 'evidenced' studies were of moderate quality: of the 11 randomised trials, four had low risk of bias and six had some risk of bias, with only one having high risk of bias [227] and most of the studies evaluated using the MMAT met most or all of the quality criteria (14/19; [Appendix 2-H](#)). The five most common BCTs associated with engagement (based on the summed number of studies from the three categories) were feedback on behaviour (48/75, 64% of studies), self-monitoring of behaviour

(46/75, 61%), goal setting (42/75, 56%), prompts/cues (40/75, 53%), and social support (unspecified; 32/75, 43%) (Table 2-6). Other BCTs that were associated with engagement in at least a quarter of studies included non-specific reward and information about health consequences (both 23/75, 31%), action planning (22/75, 29%), and instruction on how to perform the behaviour (20/75, 27%). The five most commonly supported BCTs were the same when only looking at the ‘evidenced’ studies. There was at least one study with evidenced, hypothesised, or inferred associations between engagement and 67 different BCTs; when only considering ‘evidenced’ studies, 34 unique BCTs were identified.

Table 2-6. Associations between BCTs and engagement identified in the studies

Evidence (# of studies in which BCT was associated with engagement) ^a	Original				Update				Evidenced total (n=30)	Overall total (n = 75)
	Evidenced (n = 8)	Hypoth. (n = 11)	Inferred (n = 9)	Total (n = 28)	Evidenced (n = 22)	Hypoth. (n = 9)	Inferred (n = 16)	Total (n = 47)		
1.1 Goal setting (behaviour)	5	6	6	17	10	7	8	25	15	42
1.2 Problem solving		3	2	5	1	1	6	8	1	13
1.3 Goal setting (outcome)		1	2	3	2	1	1	4	2	7
1.4 Action planning	3	2	3	8	5	4	5	14	8	22
1.5 Review behaviour goal(s)		1	3	4	2	3	5	10	2	14
1.6 Discrepancy between current behaviour and goal			1	1		4	3	7		8
1.7 Review outcome goal(s)						1		1		1
1.8 Behavioral contract							1	1		1
1.9 Commitment						2	2	4		4
2.1. Monitoring of behavior by others without feedback					3	1	1	5	3	5
2.2 Feedback on behaviour	3	7	7	17	11	7	13	31	14	48
2.3 Self-monitoring of behaviour	3	7	5	15	10	8	13	31	13	46
2.4 Self-monitoring of outcomes of behaviour			1	1	2	1	2	5	2	6
2.5 Monitoring of outcomes of behaviour			1	1	1			1	1	2
2.6 Biofeedback			1	1	4	2	1	7	4	8

3.1 Social support (unspecified)	5	5	3	13	7	2	10	19	12	32
3.2 Social support (practical)	2	4		6	1	4	1	6	3	12
3.3 Social support (emotional)		3		3	1	3	2	6	1	9
4.1 Instruction to perform the behaviour		3	3	6	5	3	6	14	5	20
4.2 Information about antecedents		1	1	2		1	3	4		6
5.1 Information about health consequences	1	5	2	8	2	7	6	15	3	23
5.2 Salience of consequences		2		2		2	1	3		5
5.3 Information about social and environmental consequences					1	4	1	6	1	6
5.4 Monitoring of emotional consequences						1		1		1
5.5 Anticipated regret	1			1		1		1	1	2
5.6 Information about emotional consequences	1	1		2	1	4		5	2	7
6.1 Demonstration of the behaviour	1	1		2	2		1	3	3	5
6.2 Social comparison		3	1	4	5	2	3	10	5	14
6.3 Information about others' approval						1	3	4		4
7.1 Prompts/cues	3	6	4	13	11	7	9	27	14	40
7.3 Reduce prompts/cues			1	1		1	1	2		3
7.8 Associative learning						1		1		1
8.1 Behavioural practice/rehearsal		1		1		4	1	5		6
8.2 Behaviour substitution			2	2		1		1		3
8.3 Habit formation	1	1	1	3	1	4	2	7	2	10
8.7 Graded tasks		1	2	3	1	2	2	5	1	8
9.1 Credible source		1		1	1	3	3	7	1	8
9.2 Pros and cons						2		2		2
9.3 Comparative imagining of future outcomes						1		1		1

10.1 Material incentive (behavior)					1	1		2	1	2
10.2 Material reward		2		2	2		3	5	2	7
10.3 Non-specific reward	2	7	3	12	2	5	4	11	4	23
10.4 Social reward			1	1	1	3	4	8	1	9
10.5 Social incentive						1	1	2		2
10.6 Non-specific incentive						2	1	3		3
10.7 Self-incentive						1		1		1
10.9 Self-reward	1			1		2		2	1	3
10.10 Reward (outcome)										0
11.1 Pharmacological support						1		1		1
11.2 Reduce negative emotions		1		1		1	1	2		3
11.3 Conserving mental resources					1			1	1	1
12.1 Restructuring the physical environment		1	1	2	1	3		4	1	6
12.2. Restructuring the social environment						1		1		1
12.3 Avoidance/reducing exposure to cues for the behavior					1	1	1	3	1	3
12.4. Distraction						2		2		2
12.5 Adding objects to the environment		1		1		3		3		4
13.1 Identification of self as role model						2	1	3		3
13.2 Framing/reframing						2	3	5		5
13.3 Incompatible beliefs						2		2		2
13.4 Valued self-identify						3		3		3
13.5 Identity associated with changed behaviour						3	1	4		4
14.2 Punishment		1		1						1
14.5 Rewarding completion						1		1		1

15.1 Written persuasion about capabilities		1		1	1	2	2	5	1	6
15.2 Mental rehearsal of successful performance			1	1						1
15.3 Focus on past success			1	1		1	2	3		4
15.4 Self-talk							2	2		2
Count of BCTs	14	29	26	40	32	58	43	64	34	67

^aSome studies did not provide sufficient evidence of a potential association between incorporated BCTs and engagement and some evidence suggested a lack of or negative association; these BCTs were not included in the counts in Table 2-6 and are identified in [Appendix 2-H](#) using italicised grey and red text, respectively.

2.4.6 Exploratory analysis of engagement components

Less than 10% (7/75) of the studies explicitly used multifaceted conceptualisations of engagement [201,204,205,217,229,243,256]. Of the studies which included measures of engagement with sufficient description to code various components, three quarters of papers captured some form of behavioural engagement (48/64), either with the intervention or with the target behaviour. Almost all of the studies that directly examined engagement with the intervention analysed app use data (33/38), sometimes supplemented by self-report via questionnaire (8/38) or interview (3/38).

More than two thirds of the studies captured data that could inform assessments of affective or cognitive (44/64) components of engagement, but affective and cognitive engagement were rarely captured explicitly or distinguished. The majority of studies examined outcomes such as acceptability, usability, and experiences using the app, which have elements relating to affective and cognitive engagement. Despite this, the analysis identified a couple of common themes in several studies, notably associations between particular BCTs and ‘motivation’ and ‘fun’ [195,201,202,207,208,220,225,241,244]. Participants suggested that features providing social support [201,208,213,241,258], feedback on behaviour [200,204,208,263], prompts and cues [195], rewards [201,204,225], goal-setting [204,213,263], and encouragement [201,225] were motivating and that gamification, such as challenges and competition, was both motivating and fun [201,202,225,242,244]. Not all participants felt the apps needed to be fun, as that was not the reason they were engaging with them [195,208]. Some studies also identified negative emotions (e.g. discouragement, guilt) associated with feedback that increased their awareness of their bad behaviour [202,205,220].

Constructs related to affective engagement were also captured in various questionnaires: irritation (EQR User Engagement subscale), sense of empathising from app (EQR Therapeutic Alliance subscale), inspiration/encouragement (EQR Therapeutic Persuasiveness subscale), fun and entertainment (uMARS engagement section), visual appeal (EQR Visual Design subscale and uMARS aesthetics section) [197,209], hedonic motivation (UTAUT2) [195], interest/enjoyment (IMI subscale) [224], and interface and satisfaction (MAUQ subscale) [267].

Data related to cognitive engagement also came primarily from qualitative feedback [195,201,202,207,220,225,229,233], although a few questionnaires (EQR User Engagement subscale [197], uMARS engagement section [209], IMI interest/enjoyment subscale [224], and the CSQ [233]) included at least one item relating to interest. Qualitative results identified a dichotomous association between cognitive effort and engagement; some feedback suggested that engagement would be negatively affected by a time-consuming app with a high cognitive load and that a basic, easy-to-use app was desirable [201], while other findings indicated that lots of content could help sustain engagement and disliked apps that were too basic [202,207].

2.4.7 Risk of bias and quality assessment

The risk of bias of the 18 randomised trials was evaluated using the Cochrane Collaboration RoB 2 tool [190,191] (Table 2-7). The domains with the lowest risk of bias were ‘measurement of the outcome’ (because most of the studies used objective measures of engagement), ‘missing outcome data’ (largely because engagement was the main outcome, so drop-out and attrition were data), and ‘deviations from intended interventions (assignment)’ (as most studies did not report any deviations). Bias in the selection of the reported result could not be determined for several studies, as pre-published protocols or registrations were not found or not sufficiently detailed.

Table 2-7. Risk of Bias 2 assessment of randomised studies

Study	Randomization process	Period and carryover effects (crossover trials only)	Timing of identification / recruitment (cluster trials only)	Deviations from intended interventions (assignment)	Deviations from intended interventions (adherence)	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall Bias
Buman et al. 2016 [220]	~	N/A	N/A	~	~	~	~	~	~
Davis and Ellis 2019 [209]	✓	N/A	N/A	✓	✓	✓	✓	~	✓
Du et al. 2016 [223]	✓	N/A	N/A	✓	~	~	✓	~	~

Hochsmann et al. 2019 [224]	✓	N/A	N/A	✓	✓	✓	✓	~	~
Nuijten et al. 2021 [227]	~	✗	N/A	~	~	✓	✓	~	✗
Nuijten et al. 2022 [228]	~	N/A	N/A	~	~	✓	✓	✓	~
<i>Additional studies included after the update</i>									
Bell (2023) [236]	✓	N/A	N/A	✓	✓	✓	✓	✓	✓
Brindal (2019) [237]	✓	N/A	N/A	✓	~	~	✓	~	~
Chang (2023) [239]	✓	N/A	N/A	✓	✓	✓	✓	✓	✓
Direito (2015) [203]	✓	N/A	N/A	✓	✓	✓	✓	✓	✓
Edney (2020) [242]	✓	N/A	✓	✓	✓	✓	✓	✓	✓
Lawrason (2023) [250]	~	N/A	N/A	✓	~	~	~	✓	~
McClure (2023) [254]	~	N/A	N/A	✓	✓	~	✓	✓	~
Mummah (2016) [215]	✓	N/A	N/A	✓	✓	✓	✓	~	~
Nahum-Shani (2021) [256]	✓	N/A	N/A	✓	✓	✓	✓	✓	✓
Sakane (2023) [216]	✓	N/A	N/A	✓	✓	✓	✓	~	~
Schulte (2022) [217]	✓	N/A	N/A	✓	✓	~	~	✓	~
Tombor (2019) [265]	✓	N/A	N/A	✓	✓	✓	✓	~	~

The Mixed-Methods Appraisal Tool (MMAT) [192] was used to assess the quality of the remaining studies ([Appendix 2-I](#)). Of the study types, the qualitative studies met the most criteria; however, as the individual quality criteria are different for each study type, a direct comparison is inappropriate. Common issues with the mixed-methods studies assessed were a lack of explicit justification for the mixed-methods design and a lack of detailed interpretation of the integration of results. For the quantitative studies, almost half (7/15) did not have participants representative of the target population. Across all of the study types, there was a common issue regarding the item about ‘clear research questions’; two-thirds of the studies (30/44) reported aims, but not research questions. Studies that were primarily focused on development rather than evaluation were not included in this assessment (13/75) [199,214,218,219,231,235,240,241,243,246–248,261], as was the one included content analysis [197].

2.5 Discussion

*This section has been updated from a previously published paper [47].

2.5.1 Principal findings

75 studies that developed or examined mobile apps for health behaviour change were examined to investigate how BCTs could influence different components of engagement. Although a variety of methods were used and various components of engagement were captured, most studies did not purposefully examine these different components. Different measures tended to capture different components: affective and cognitive engagement were primarily captured with qualitative methods, micro behavioural engagement with app use data, and macro behavioural data with self-report, questionnaires, and app tracking. This emphasises the importance of mixed methods to fully assess user engagement. By combining study-reported evidence with the coding of BCTs and engagement, I identified patterns of evidence that suggest the potential of 5 BCTs to support engagement with DHIs: feedback, self-monitoring, goal setting, prompts/cues, and social support. Additional BCTs with some evidence of association with engagement (with the intervention or behaviour) included rewards, information about health consequences, action planning, and instruction on how to perform the behaviour. A lack of explicit reporting of BCTs and preplanned analyses of engagement components limited my ability to identify potential associations between BCTs and specific components of engagement.

2.5.2 Comparisons with existing literature

The BCTs identified as having evidence of associations with engagement are in line with previous findings. The mixed influence of social support on engagement has also been observed by previous research [184,268] and could potentially be explained by a difference between hypothetical preferences and practical experiences [184]. The implementation of the social support BCT could also be a factor; the studies that found a positive influence of social support on engagement were team-based [223,227], provided practical social support, such as connection with healthcare professionals, or offered optional opportunities to interact with other users to get peer support, compete, or be held accountable to their goals [201,202]. There were mixed findings for social comparison and competition, with some studies indicating that these could support motivation and engagement for at least some participants [242,244–246] while others found that participants disliked these features [203,206,208,213].

2.5.3 Strengths and limitations

The broad scope of the review is both a strength and a limitation; the inclusive search strategy reduced the likelihood of missing relevant studies, but necessitated automated screening. To mitigate the risk of screening out relevant studies, I used a set of six articles that had been

discovered in preliminary searches and determined to be eligible to check the screening. After each pass in EndNote, if any of them were missing, the search terms were revised and the pass re-run to ensure I was not incorrectly eliminating relevant papers. Although the search terms were developed with a second author, the review was not independently executed by two reviewers, which would have reduced the risk of bias in the analysis and increased the credibility of the conclusions. It was not possible to have a second reviewer due to time and resource constraints.

To strengthen the study, I conducted a second round of searches and screening to incorporate studies published since the initial search and to amend the screening strategy to be more inclusive and avoid unintentional exclusion of relevant articles through automated screening. A significantly greater number of articles were screened manually in the second round, which resulted in almost twice as many articles being included in the second round compared to the first. The analysis of these findings produced similar results to the initial review, improving the confidence in and credibility of the initial and overall conclusions.

The diversity of outcomes studied prohibited a quantitative meta-analysis to examine effect. Additionally, the studies examined a wide variety of health apps, aimed at various populations and behaviours. In including this variety of studies, the review is limited in its ability to explore potential context-specific associations. Contextual factors such as behaviour, age, app type, and outcomes could influence engagement in ways that were not possible to capture with the current sample.

This study instead contributes to the literature by analysing the body of evidence to posit associations between BCTs and engagement for future testing. The analysis was limited by the available reported data and descriptions of the studies' apps and measures. The lack of explicit analyses of engagement components in the included studies meant the review relied on author coding and inference to identify potential associations. The BCTs that were identified as having the strongest evidence of association with engagement are also some of the most common BCTs included in mobile health apps [269,270]; it is likely that there are other BCTs that are associated with engagement, but are less frequently used, and therefore have less available evidence of an association to capture in a review. Likewise, although the review was intended to focus on BCTs to limit the scope, this also represents a limitation; engagement strategies beyond BCTs (such as personalisation, gamification, ease of use, etc.) should not be overlooked [184,271]. Examining a broader range of engagement strategies may have identified more evidence of associations between particular strategies and specific components of engagement. As such, this review represents a starting point for further

investigation of engagement and an identification of current limitations in digital health methodology and reporting.

2.5.4 Future research and recommendations

Through this review, I identified gaps in the literature that shaped the rest of the thesis. A challenge when analysing the data was that there were no clear and comprehensive definitions of affective, cognitive, and behavioural engagement in the literature. Constructs commonly associated with affective engagement included motivation, positive and negative feelings, and sometimes aesthetic pleasure, while cognitive engagement was described using terms such as interest, attention, thought, and challenge [42,45,51], but these were compiled from various sources and do not necessarily reflect complete and defined constructs. The findings also emphasised a gap in the literature that was previously identified; the focus on objective and behavioural measures of engagement, often restricted to usage metrics. When captured, qualitative and affective engagement data was primarily qualitative (with some self-report questionnaires), supporting recent discussions in the literature on the importance of using mixed-methods designs to enable engagement to be comprehensively assessed [14,42,44].

Based on the associations with motivation evidenced in the studies, BCTs such as goal setting, self-monitoring, feedback, social support, rewards, and prompts and cues could be associated with cognitive and affective engagement. For some BCTs (e.g. prompts and cues, social support, rewards), the evidence indicates that the way they are incorporated into interventions can affect how they are associated with engagement. Recent studies have identified a lack of justification and guidance for the operationalisation of BCTs [272,273]. It will be important to examine if there are certain operationalisations that better support engagement or whether BCTs' influences are related to individual characteristics.

Understanding when and how they have a positive or negative impact would enable interventions to be tailored to the specific context in which they will be implemented or dynamically to individuals, to emphasise specific BCTs and app features that those individuals are most likely to respond best to.

In view of BCTs and app design, there appears to be a balance needed to obtain the optimal level of cognitive load. Data from participants indicated a preference for a reduced cognitive load - limited complexity, less data input, limited number of features, not time-consuming [201,274] - while other data and motivation theory suggests that apps need to be interesting enough to keep users' attention [202,207] and provide the right level of challenge to achieve optimal motivation [274]. It is likely to be important for engagement to choose BCTs and

other app features carefully to fit the context and the aim of the intervention, rather than incorporating as many BCTs as possible.

2.5.5 Conclusions

The purpose of this systematic review was to incorporate theoretical conceptualisations of engagement and synthesise evidence of how mobile health apps and their BCTs are associated with different components of user engagement. Nine BCTs were identified as having the strongest evidence of a potential association with engagement: feedback, self-monitoring of behaviour, goal-setting and action planning, prompts/cues, social support, rewards, information about health consequences, and instruction on how to perform the behaviour. Recommendations for app development deriving from the findings are the inclusion and further investigation of the BCTs identified to support motivation and the importance of identifying the right level of cognitive involvement for the target population to keep the app interesting without being a burden. The importance of personalisation is also emphasised by these results, as some of the BCTs identified as having associations with engagement (e.g. social support/comparison and rewards) could have positive or negative impacts, depending on the study and individual user characteristics. This review contributes to the body of evidence for the influence of specific BCTs on engagement by identifying a gap between theoretical conceptualisations of engagement and its evaluation in mobile health apps and by positing potential associations for future empirical investigation.

An exploratory analysis of the engagement components identified key areas for future research, including further theoretical clarity of engagement components, how less researched BCTs might be theoretically linked to engagement components, and how BCTs can be operationalised in app design to best support affective, cognitive, and behavioural engagement. Greater clarity in the definitions of different components of engagement, and further development and incorporation of means of measuring specific types of engagement, could enable future research to examine how different components of engagement are related to specific BCTs and behavioural and health outcomes. This knowledge could inform the design and tailoring of mobile health apps to improve engagement with the intervention and target behaviours. By highlighting these gaps, the systematic review helped shape the rest of the thesis by identifying key questions that needed to be addressed before empirical case studies could be conducted and analysed. These gaps around the conceptualisation and measurement of engagement components will be examined further in Chapter 3.

Beyond highlighting these gaps for further investigation, the systematic review also contributes additional data that will be incorporated with empirical data from the case studies in Chapter 7 to provide a list of evidence-based recommendations for how intervention designers can help support effective engagement. Basing these recommendations on a synthesis of the in-depth case study analyses as well as a broad analysis of existing literature helps to improve confidence in the conclusions drawn and in the potential value of incorporating the recommendations into intervention design.

Chapter 3: Conceptualisation and Measurement of Engagement

3.1 Chapter Overview

Chapter 2 highlighted the gap in the literature on how different components of engagement may influence how users engage with interventions and target behaviours over time and how Behaviour Change Techniques (BCTs) and other app features could help support these different components. Previous research has also identified other strategies that can support engagement, such as personalisation [271], user-friendly and non-burdensome design [275], and reciprocity [256]. Despite this, high attrition and low engagement in digital health interventions (DHIs) remain common and the effectiveness of particular strategies varies over time and by individual [256,275,276]. Compounding this gap in knowledge is the fact that evaluations of engagement are commonly limited to behavioural measures of app use [30,37] rather than exploring engagement as a multifaceted concept involving affective, cognitive, and wider behavioural components.

Chapter 3 aims to further explore the gaps identified in the literature in the systematic review. To address the lack of clarity around the conceptualisation of components of engagement, it examines how theory-based, multifaceted measures of engagement with DHIs capture different components of engagement (affective, cognitive, behavioural, micro, and macro) and what insights about the relationships between these components can be generated from an analysis of such scales. The chapter begins by summarising the current issues with the measurement of digital health engagement and the methods used in this analysis. The first part of the chapter discusses the implications of this analysis for how affective, cognitive, and behavioural components of engagement can be defined and whether there is value in distinguishing between them. Expanding on a published study, it summarises the theoretical bases of the included measures and conducts an item-level comparison to assess how they capture cognitive, affective, behavioural, micro, and macro components of engagement. The second part of the chapter aims to build on these insights. Based on previous evidence and theory about affective, cognitive, and behavioural processes, a model of the relationships between these components over the process of digital health engagement will be generated. This model will serve as a hypothesised mapping for the empirical case study analyses to be compared against. The chapter concludes with the implications of this analysis for research, recommendations for the measurement of engagement, and a discussion of the chapter's strengths and limitations.

Publications: Sections of Chapter 3 have been drawn from a previously published paper; its reference and author contributions are listed below.

Milne-Ives M, Homer S, Andrade J, Meinert E. The conceptualisation and measurement of engagement in digital health. *Internet Interventions*. 2024 Jun; 36:100735. doi:[10.1016/j.invent.2024.100735](https://doi.org/10.1016/j.invent.2024.100735)

Authors' contributions: MMI conceived the topic, conducted the analysis, and drafted and revised the protocol. SH, JA, and EM supervised the research and contributed revisions.

3.2 Origin of this Study

The concept for this chapter emerged from two different methodological challenges. The first was a need to further unpack issues underlying the multifaceted view of engagement in the literature when conducting the systematic review (Chapter 2). The lack of established definitions for the different components of engagement posed a practical issue for the review, as it made it difficult to code what aspects of engagement the measures were capturing. The second challenge was the need to select a self-report measure of engagement to use in a planned future evaluation of one of the case studies to triangulate with app use data and qualitative feedback. More broadly, these challenges make it difficult for digital health evaluations to conduct in-depth, replicable analyses of engagement.

On a preliminary search, two self-report measures that captured engagement as a multifaceted concept were identified. The shape of this study solidified through the process of trying to determine which measure to include in the planned future evaluation. Although the measures use slightly different theoretical conceptualisations of engagement, both aim to capture various affective, cognitive, and behavioural components and have relatively similar psychometrics [277,278]. When examining them in more depth, I became intrigued by their differences - although they appear to capture similar components of engagement, the items used to capture these latent factors seemed quite distinct. I was curious as to how two measures aiming to measure the same concept, with similar theoretical backgrounds and relatively well-correlated scores, had such different items and what an examination of these two frameworks could reveal about the conceptualisation of engagement and its components.

The purpose of this chapter was to use an examination and comparison of these scales as a starting point for exploring the conceptualisations of different components of engagement (affective, cognitive, and behavioural). The aims were to 1) unpack how the facets of engagement have been conceptualised in psychometric scale development, 2) consider what is

still missing from current measures, 3) explore what can be learnt to improve the measurement of engagement in digital health, 4) develop the definitions of the engagement components that will be used in this thesis, and 5) generate hypotheses about how those components interact over the process of engagement. As I conducted this theoretical examination, I realised how pivotal it would be for the within- and cross-case analyses. Without a clear understanding of how I was defining engagement components and hypothesising their relationships over the process of engagement, it would have been impossible to clearly articulate how those components influenced each other in the context of the three case studies.

The initial version of the chapter examined two self-report measures that were specifically developed to evaluate digital health engagement as a multifaceted concept (the Digital Behaviour Change Intervention Engagement Scale [277,279] and the TWente Engagement with Ehealth Technologies Scale [278,280]). These were identified from my general reading of the recent engagement theory and literature. When part of the chapter was submitted for publication, peer review highlighted the lack of transparency and replicability in the process of scale selection for the analysis. Based on this feedback, I identified five recent reviews that summarised the literature on measures of engagement with DHIs and assessed all of the scales identified against a set of eligibility criteria ([Section 3.4](#)). No additional scales met the criteria and were included, but the revisions of this paper based on this feedback improved its quality and rigour and have been incorporated into this chapter.

3.3 Issues with the Current Measurement of Engagement

The evaluation of engagement with DHIs is important because of the role engagement plays in moderating the impact of the intervention on health behaviours and outcomes. If an intervention has no observed effect on behavioural or health outcomes, there is a need to disambiguate what might be contributing to that lack of impact so that it can be addressed and improved. Recent efforts have advanced the theoretical conceptualisation of engagement with digital health ([Section 1.3](#)) but the evaluation of DHIs has not yet caught up with these theoretical advancements. As demonstrated in Chapter 2, when digital health evaluations assess engagement, most focus primarily on system usage data as their measure [14,41,44,45,281]. The key issue with this reliance on system usage metrics to assess engagement is that system use metrics only capture behavioural engagement with the digital intervention ('micro' engagement) [14,30,44]. This limits researchers' understanding of user engagement as a dynamic and multifaceted concept, because system usage metrics cannot

capture affective or cognitive elements of engagement [30,281] or offline ('macro') engagement with the behaviour change process [14,44].

The use of system usage data as an outcome measure is more aligned to the related concept of 'adherence,' which refers to the degree to which users are using an intervention as intended [43]. However, 'intended use' of the intervention may not align with the concept of 'effective engagement': "the degree of engagement needed to achieve intended outcomes" [14]. This highlights the crucial distinction between engagement with an intervention and engagement with a behaviour, as indefinite engagement with an intervention is not necessarily required to support sustained engagement with the behaviour [14,281]. High system use does not necessarily correlate with positive behavioural or health outcomes and existing evidence of the extent to which usage mediates these outcomes is limited [14,30]. For example, discontinued use of a DHI could indicate that the user has completely disengaged from the process or that they have been sufficiently successful in modifying their behaviour that they no longer need the intervention's support [14,282]. Conversely, long durations of engagement with a DHI could equally be due to meaningful engagement with content and intervention components or frustration and difficulties with the system [30]. Without further contextual data, system use metrics cannot provide a clear understanding of engagement.

The difficulty with evaluating 'effective engagement' is how it can be captured so that it can be compared between studies, enabling synthesis and the generation of new insights. This is a common challenge in digital health literature; definitions and measures used to assess engagement are inconsistent [41,45,50,281] and the lack of a clear threshold for what is considered 'good' or 'sufficient' engagement makes it difficult to compare engagement across studies [37,171]. One review of the measurement of engagement with mental health apps found that all 40 studies concluded that their interventions had good engagement, but all used different measures and criteria thresholds [171]. This is a complex issue, as the criteria for what is considered 'effective engagement' (in terms of users "getting what they need" out of the intervention [14,30]) can vary significantly by intervention or individual. For one patient, just logging into a mental health app might reflect a significant step in their willingness to get support and have its own benefit for their mental health [30]; another patient might only derive benefits if they are regularly engaging with app resources to build skills and habits [30].

There are a couple of ways to improve the credibility of assessment of engagement; for example, pre-specifying the criteria thresholds in published protocols or pre-registrations, based on theory or previous evidence, would increase transparency. The use of multiple

measures of engagement would enable triangulation, improving credibility and providing a more valid assessment of engagement by capturing a more holistic perspective [35,44,46,171]. In addition to system usage data, there are a variety of methods that can be used to measure engagement, including “qualitative measures, self-report questionnaires, ecological momentary assessments, sensor data, social media data, and psychophysiological measures” [44]. In terms of cost and feasibility, self-report questionnaires tend to be the easiest to implement in studies; however, there is a large variety of questionnaires that have been used in the literature to assess engagement or related constructs (Table 2-5, Table 3-2) and most do not (at least explicitly) capture engagement as a multifaceted concept.

3.4 Methods

3.4.1 Scale selection

A preliminary search for a theoretically-based, self-report measure to evaluate engagement with a DHI identified two promising scales: the Digital Behaviour Change Intervention Engagement Scale (DBCI-ES) [277,279] and the TWente Engagement with Ehealth Technologies Scale (TWEETS) [278,280]. To identify other potentially relevant measures for inclusion in the analysis and to formalise the method for scale selection, I established a set of eligibility criteria (Table 3-1) and examined recent reviews of engagement measures. Five recent reviews (published between 2018-2023) were identified from searches of Google Scholar and PubMed using keywords relating to ‘engagement’, ‘measure’ or ‘method’, ‘review’, and ‘digital health’ [44,46,171,185,279]. All the scales identified from these reviews (which included the DBCI-ES and the TWEETS [46]) were assessed against the eligibility criteria for scale selection to determine the scales included in the analysis.

Table 3-1. Criteria for scale selection

1	Self-report questionnaire measure
2	Theoretically-based on a multifaceted conceptualisation of engagement
3	Aims to capture affective, cognitive, and behavioural components
4	Developed to measure engagement with digital health

3.4.2 Analysis of included scales

A comparison of the included measures was performed by laying out the scale items, grouped by the theoretical components specified by the scale. The included scales were compared on an item level within the framework of a three-component conceptualisation of engagement

(affective, cognitive, and behavioural). I chose to use a three-component conceptualisation rather than a two-component conceptualisation for how users engage with DHIs (see Figure 1-1) to structure the paper because it enabled a more in-depth discussion of whether affective and cognitive components should or should not be considered separately. The theoretical implications of individual items and their associations with the different components of engagement in relation to theory and literature more broadly were examined and recommendations for the use and improvement of measures of digital health engagement were generated.

3.5 Results of Scale Comparison

3.5.1 Selection of measures for comparison

After assessing the scales used to measure engagement identified in previous reviews (Table 3-2), only the DBCI-ES and TWEETS met the four inclusion criteria. One other scale - the Website Evaluation Questionnaire [283] - met most of the criteria; however, this questionnaire was developed for a particular study and a particular type of DHI by its authors and there was insufficient discussion of its theoretical basis and conceptualisation in that paper to justify its inclusion.

Table 3-2. Assessment of scales used to measure engagement with DHIs

Scale	Identified from	Inclusion criteria			
		1. Self-report measure	2. Theoretically-based on a multi-faceted conceptualisation of engagement	3. Captures A, C, and B components*	4. Developed for digital health
Client Evaluation of Services Questionnaire [284]	[171]	✓	×	×	×
Client Satisfaction Questionnaire [285]	[46,171]	✓	×	×	×
Cognitive Absorption Scale [286]	[44]	✓	×	×	×
Computer System Usability Questionnaire [287]	[171]	✓	×	×	×
Credibility and Expectancy Scale [288]	[171]	✓	×	×	×

Digital Behaviour Change Intervention Engagement Scale [277,279]	[44,46]	✓	✓	✓	✓
eHealth Engagement Scale [289]	[44,279]	✓	×	×	✓
Enjoyment of website experiences scale [290]	[44]	✓	×	×	×
Flow State Scale [291]	[279]	✓	×	×	×
Gaming Engagement Questionnaire [292]	[44]	✓	×	×	×
Immersive Experience Questionnaire [293]	[44,279]	✓	×	×	×
Mobile Application Rating Scale [294]	[279]	✓	×	×	✓
Personal Involvement Inventory [295]	[279]	✓	×	×	×
Post-Study System Usability Questionnaire [296]	[171]	✓	×	×	×
Service Engagement Scale [297]	[298]	✓	×	×	×
System Usability Scale [299]	[46,171,298]	✓	×	×	×
Technology Acceptance Model [300,301]	[171]	✓	×	✓	×
Therapeutic Alliance Scales for Children [302]	[171]	✓	×	×	×
TWente Engagement with Ehealth Technologies Scale [278,280]	[46]	✓	✓	✓	✓
Usefulness, Satisfaction, and Ease questionnaire [303]	[171,298]	✓	×	×	×
User Engagement Scale (and short-form) [304,305]	[44,279]	✓	✓	×	×
User Experience Questionnaire [306]	[44]	✓	×	×	×

Website Evaluation Questionnaire [283]	[46]	✓	×	✓	✓

**Affective, cognitive, and behavioural components of engagement*

3.5.2 Summary of included multifaceted measures of engagement

The DBCI-ES and the TWEETS were developed to provide validated self-report scales that could capture the different facets of engagement in a quick and easy-to-use measure [279,280]. This section will provide a brief overview of how each of the scales were developed to inform the comparison and examination of their theoretical constructs.

3.5.2.1 Digital Behaviour Change Intervention Engagement Scale (DBCI-ES)

The DBCI-ES was developed based on a systematic review and theoretical conceptualisation conducted by its authors [42]. The systematic review examined studies that measured or discussed engagement with DHIs to identify and code concepts relating to engagement that were included in the measures or discussions [42]. Based on this coding process, the authors generated two constructs that comprised their definition of engagement: ‘engagement as behaviour’ and ‘engagement as subjective experience’. Engagement as behaviour refers to the usage of the digital intervention and reflects dimensions relating to amount, duration, frequency, and depth of use. Their conceptualisation of engagement as subjective experience includes both cognitive and affective elements that are defined as attention, interest, and affect. The authors derived these three elements from theoretically similar concepts from other fields identified in their review, such as ‘flow’ and ‘immersion’ [42].

Based on this theoretical conceptualisation, the authors identified 5 elements as necessary and sufficient conditions to capture engagement: attention, interest, enjoyment, amount of use, and depth of use [279]. Attention involves the process of paying attention to the DBCI, which interest and enjoyment can help to support [279]. Amount of use refers to any time spent on the app, including passive or active engagement (eg. reading content vs. setting goals), while depth of use reflects the intensity of a user’s engagement by capturing how many elements of the intervention (eg. specific features or BCTs) they accessed. Without the subjective experience elements, however, the behavioural indicators were not considered to be sufficient to capture engagement, as users could spend time on an intervention without paying attention to it or involving themselves with it for the purposes of behaviour change [279].

An initial set of scale items was generated using a deductive approach [307] based on the theoretical definitions of these 5 elements. The initial items were then assessed by 20 experts in behavioural science and human computer interaction and 50 non-experts [204]. The

participants classified each of the items as one of the five elements (or ‘unclassified’) in a content adequacy task; items were included in the final scale if at least 70% of the experts and non-experts classified the item as belonging to the element that it was intended to represent. The authors did report one exception; the ‘depth of use’ item did not quite meet the cut-off criteria (~65%) but was included anyway as the authors considered it a necessary theoretical component to capture [204].

The final questionnaire ([Appendix 3-A](#)) has 10 items - 8 relating to users’ subjective experience of the intervention and 2 relating to the extent of their use of the intervention [277]. The experiential items are measured based on users’ reports of how strongly they experienced a particular term on a 7-point Likert scale (1 = not at all, 4 = moderately, 7 = extremely). The experiential items include negatively and positively valenced terms; the three negative items (inattention, distraction, and annoyance) are reverse-scored [277]. The experiential subscale has a minimum score of 7 and a maximum score of 56, with a higher score indicating greater engagement. The two behavioural items are measured separately in seconds and in the proportion of app components used (0-100%).

3.5.2.2 TWente Engagement with Ehealth Technologies Scale (TWEETS)

The TWEETS was developed based on a largely similar theoretical conceptualisation as the DBCI-ES but using different methods. It aimed to provide a holistic measure of engagement that addressed the psychometric limitations of the DBCI-ES (issues around the independence of the two factors and criterion and divergent validity) [277–279]. As with the DBCI-ES, the development of the TWEETS started with a systematic review conducted by the authors. Unlike the other review [42], the authors limited their review to papers that provided a definition, conceptualisation, or theory of engagement but expanded the scope to include conceptualisations of engagement from any field, not just digital health [45]. They found that, although definitions varied across fields, there was general agreement that engagement was multifaceted and broadly included affective, cognitive, and behavioural components [45]. To develop their measure of engagement, the authors interviewed 20 digital health app users and 10 experts in eHealth development and research [280].

Responses that reflected participants’ perspectives of engagement with DHIs were deductively categorised as affective, cognitive, or behavioural by the authors and then inductively coded to generate key themes [280,307]. From their analysis, themes around behavioural engagement related more to the incorporation of digital interventions into participants’ daily lives and routines than the frequency of their use [278]. With respect to cognitive

engagement, themes focused on the personal relevance of the intervention for users and their confidence in its usefulness and its ability to help them achieve their goals. Affective engagement themes included users' feelings towards the intervention and how it related to their goals as well as their ability to identify with the intervention and its aims [278]. These key themes were used to create the scale items ([Appendix 3-B](#)) [280].

The TWEETS includes 9 items (3 each for affective, cognitive, and behavioural engagement) [278]. The items are measured on a 5-point Likert scale (from strongly disagree = 0 to strongly agree = 4) with a minimum score of 0 indicating low engagement and a maximum score of 36 indicating high engagement. None of the items are reverse-scored. The scale items include sections that can be tailored for the particular intervention being assessed [278].

3.5.2.3 Psychometric evaluations

Psychometric evaluations of both scales have been conducted by their authors; Kelders and colleagues also conducted a comparison between their TWEETS and the DBCI-ES [277,278]. Exploratory factor analysis (EFA) determined that the DBCI-ES comprised two main factors: behavioural and experiential. The experiential subscale had high internal reliability ($\alpha = .78$) and the behavioural moderate ($\alpha = .45$), but the subscales were not significantly correlated with each other [277]. In terms of predicting future engagement, the total scale scores were significantly associated with subsequent logins, although this effect was mainly driven by the experiential element. The study did not test the scale's ability to predict subsequent engagement with the health behaviour. Overall, the authors concluded that further refinement and validation in larger and different samples was needed [277].

Unlike the DBCI-ES, EFA found that the TWEETS best fitted a one factor structure [278]. Both the TWEETS and DBCI-ES experiential subscale were assessed over the four time points (baseline, 1 day, 1 week, and 2 weeks); both had generally good internal consistency (defined as $\alpha > .80$; $\alpha = .86, .86, .87$ and $\alpha = .72, .80, .84$ at T1, T2, and T3, respectively). Pearson's correlations of both measures between each time point were all significant and demonstrated relatively stable moderate test-retest reliability [278]. Both measures demonstrated predictive validity for a general self-report of perceived behaviour change ("Do you feel that you have changed your behaviour because of using the step counter app?"), although TWEETS explained slightly more of the variance between the first and last time point [278]. The authors also found moderate to strong correlations between the TWEETS and DBCI-ES experiential subscale ($r = 0.57, 0.73, \text{ and } 0.80$ at T1, T2, and T3) but neither had strong correlations with reported use frequency. Overall, the authors concluded the

TWEETS performed well - similarly to, and occasionally slightly better than, the DBCI-ES [278].

3.5.3 Comparison of scales

Considering the scales are relatively strongly correlated and propose to measure similar components of engagement, their items look quite different (Table 3-3) [277,278]. The following subsections compare the items addressing the various engagement components.

Table 3-3. Comparison of DBCI-ES and TWEETS

DBCI engagement components	DBCI sub-components	DBCI items	TWEETS items	TWEETS engagement components
Subjective experience	How strongly did you experience the following? (7-point scale)		Thinking about using [the technology] the last week, I feel that: (5-point scale)	
	Attention	1. Focus 2. Inattention (R)* 3. Distraction (R)	1. [This technology] makes it easier for me to work on [my goal] 2. [This technology] motivates me to [reach my goal] 3. [This technology] helps me to get more insight into [my behaviour relating to the goal]	Cognition
	Interest	4. Interest 5. Intrigue	4. I enjoy using [this technology] 5. I enjoy seeing the progress I make in [this technology]	Affect
	Enjoyment	6. Enjoyment 7. Annoyance (R) 8. Pleasure	6. [This technology] fits me as a person	
Behaviour	Amount of use	9. How much time (in minutes) do you roughly think that you spent on the app?	7. [This technology] is part of my daily routine 8. [This technology] takes me little effort to use 9. I'm able to use [this technology] as often as needed (to achieve my goals)	Behaviour
	Depth of use	10. Which of the app's components do you remember visiting? (0-100% of components)		

*(R) denotes items that are reverse-scored.

3.5.3.1 Cognitive engagement

The DBCI-ES includes two sub-components that relate to cognitive engagement - the interest and attention users experience when using the intervention (micro engagement). Attention includes an investment or focusing of cognitive resources towards specific perceptual

information at the expense of other information [44,308,309], which is captured well by its items (1-3). The concept of interest includes interactions between cognitive and affective factors [44,310,311], which highlights the overlap between cognitive and affective engagement. This is not an issue for the DBCI-ES, as it does not seek to separate cognitive and affective components but groups them together under ‘subjective experience’. Compared to the DBCI-ES, the TWEETS cognitive items (1-3) are less obviously aligned with interest and attention to the app (micro engagement) and more with users’ motivation to achieve their behavioural goal (macro engagement) [278]. However, the TWEETS items are framed such that the focus is on the space between micro and macro engagement - they capture how the intervention supports the behavioural goal. Participants are asked how much the DHI supports their motivation, making it more explicit that DHI use is a means to an end.

3.5.3.2 Affective engagement

The affective component is the most similar across the scales, with both explicitly referencing ‘enjoyment’. Unlike the TWEETS, the DBCI-ES includes both positively- and negatively-valenced items (6-8). Whether negative emotions should be considered part of affective engagement is still an open question, but evidence suggests that negative emotions can play a role in increasing or decreasing engagement [45,49,312,313]. As above, the concepts within affective engagement are broader in the TWEETS than the DBCI-ES; although TWEETS item 4 also captures enjoyment associated with using the technology, item 5 relates this affective experience to the users’ behaviour change progress and item 6 captures how users feel the technology fits their personal values.

3.5.3.3 Behavioural engagement

The trend of differing scopes continues with behavioural engagement, where the main difference between the scales is the degree of focus on micro engagement. The DBCI-ES focuses only on intervention use (items 9-10) [279] - data which could be captured objectively by system use metrics - whereas the TWEETS captures behavioural engagement with the intervention in the context of daily routines (item 7), effort (item 8), and goals (item 9). The latter aligns more with Yardley’s concept of ‘effective engagement’ as “sufficient engagement with the intervention to achieve intended outcomes” [14].

3.6 Theoretical Analysis of the DBCI-ES and TWEETS

3.6.1 Overall comparison of the engagement scales

Considering that the two scales were relatively strongly correlated, and that both propose to measure similar components of engagement - behavioural and subjective (affective and cognitive) - the scale items look quite different. In their exploratory psychometric comparison of their TWEETS and the DBCI-ES, Kelders *et al.* [278] allude to the conceptual similarities between the scales but do not discuss the differences between the scale items and what this means in terms of the conceptualisation of engagement. One potential explanation for why they appear different but are correlated is that each scale is capturing different sub-aspects of a holistic construct of engagement. This could have emerged because, although the development of both measures incorporated expert and non-expert perspectives, it also relied on the authors' judgement. In the DBCI-ES, the authors generated the initial list of items and provided categories for experts and non-experts to classify them into, to confirm if participants classified them in the same way as the authors [279]. The potential limitation with this strategy is that, even though the items were generated based on theoretical definitions and successfully matched by participants with the appropriate sub-component (attention, interest, enjoyment, amount of use, or depth of use), it is unclear whether they actually capture users' real experience of engagement. The method of developing the TWEETS had the opposite issue - items were generated based on what experts and non-experts think it is to be engaged with a DHI, but were grouped into affective, cognitive, and behavioural components based on the authors' interpretations [280]. This section will consider, in reference to broader theory, how well the items capture the different components of engagement and how the examination of the scales can generate insights about the conceptualisation of the components of engagement.

3.6.2 How many components of engagement are there?

3.6.2.1 Overview of analysis

The DBCI-ES and TWEETS used 2- and 3-component conceptualisations of engagement, respectively. This refers to components relating to how users engage with DHIs (subjective and behavioural vs. affective, cognitive, and behavioural; Figure 1-1b). The item level comparison of the DBCI-ES and TWEETS found that, although both scales have a similar theoretical basis and methods of development, a key difference between the scales is their scope in terms of what users engage with (Figure 1-1a). The DBCI-ES instructs users to answer the questions "with regards to [their] most recent use" of the intervention, which

focuses the scale clearly on micro engagement [279]. In contrast, the TWEETS blurs the distinction between micro and macro engagement by relating intervention use to goal achievement [278]. The assumption that the TWEETS captures micro engagement to some extent is supported by the correlation between the two scales [278].

3.6.2.2 Cognitive engagement

First, let us look at how the DBCI-ES and TWEETS items capture cognitive engagement. The DBCI-ES groups affective and cognitive engagement as subjective experience [42], so to focus on cognitive engagement we must first determine which of its three subjective sub-components to examine. On first glance, attention and interest might be picked out as the ‘cognitive’ components of the DBCI-ES [45,49,63], but Short *et al.* define situational interest as “an emotional state brought about by situational stimuli” that can play a role in how people direct their attention [44]. In the broader literature, definitions of elements related to cognitive engagement vary; in the field of game engagement, the concept of interest has been defined as including elements relating to preference, perceived relevance, importance, and motivation [314]. This suggests that interest relates to both cognitive and affective engagement. As the DBCI-ES does not aim to distinguish between these components, this is not an issue for the scale, but highlights the close integration between the components.

The TWEETS does aim to distinguish between affective and cognitive engagement. Items 1 and 3 seem aligned with cognitive effort and consideration, but item 2 focuses on motivation, which, like interest, is generally considered to be driven by both cognitive and affective elements [53,70,315]. Across theories [70,315–318], motivation tends to include two main components: self-efficacy or beliefs about competence (“can I do it?” [70]) and value (“do I want to do it?” [70]). For example, Expectancy-Value Theory identifies two factors influencing motivation: the degree to which an individual believes they will be successful and the degree to which the task has value (shaped by associated emotions from similar experiences) [70,317]. Likewise, Social Cognitive Theory identifies goals, perceived progress and expected outcomes, values, and self-efficacy as key motivational factors [315]. Self-Determination Theory suggests that the type or quality, as well as the degree, of motivation can vary [70,316]: extrinsic motivation is driven to varying degrees by social values or requirements, while intrinsic motivation drives behaviour simply because the task is interesting, challenging, or enjoyable [70,316]. Motivation has also been defined as “cognitive-affective events” that reflect “desires to reach a goal state” [319]. Although these definitions capture slightly different factors that drive motivation, they all highlight both affective and cognitive elements.

For these reasons, it is interesting that Kelders and colleagues classified their item relating to motivation as cognitive, when it can more broadly include both cognitive and affective elements. This could be due to the way that the participants expressed their thoughts relating to motivation in the interviews (as responses were initially categorised as affective, cognitive, or behavioural before items were generated), although this is not clear from the paper. The example quote provided for motivation does reflect a more cognitive element “*it helps me to ... get more conscious about [my] own behaviour and this motivates me*” [280]. The phrasing of the item in the scale, however, does not specify that it reflects a cognitive interpretation of motivation.

3.6.2.3 Affective engagement

Turning to affective engagement, the scales seem more aligned: both capture the positive emotional experience associated with the intervention, although TWEETS item 5 extends this beyond the intervention and DBCI-ES item 7 captures negative emotional experience. TWEETS item 6, however, focuses on the technology’s fit with the user “as a person”, which has conceptual similarities to the ‘ethicality’ construct of the Theoretical Framework of Acceptability (TFA) [320] and the ‘value’ component of motivation (particularly the task value aspect of Expectancy-Value Theory [70,317]. ‘Ethicality’ in the TFA refers to “the extent to which the intervention has good fit with an individual’s value system” [320] while ‘task value’ refers to the “degree to which individuals perceive personal importance, value or intrinsic interest in doing the task” [70,317]. This demonstrates again how the concept of motivation integrates the cognitive and affective components. As with the cognitive engagement component, the perspective of what should be included in the scope of affective engagement is broader in the TWEETS than in the DBCI-ES.

3.6.2.4 Behavioural engagement

Both scales consider behavioural engagement separately from cognitive and affective engagement, although they again differ in scope (micro vs. macro engagement). DBCI-ES items 9 and 10 and TWEETS items 7 and 9 focus on use of the intervention, but TWEETS item 8 relates to the effort needed to use the technology. It is interesting that the authors chose to categorise item 8 as behavioural engagement; for able-bodied individuals, the physical effort required to use a DHI is low, so the effort spent behaviourally engaging with the DHI is likely to be primarily cognitive [321]. The authors state that “when cognitively engaged, users are willing to spend mental effort to work with the technology” [280], which implies that cognitive effort is an intermediary between cognitive and behavioural engagement rather than

belonging strictly to one or the other. If this item is interpreted as relating to usability, rather than cognitive effort, it still blurs the boundaries of the categorisation, as the concept of usability includes affective, cognitive, and behavioural elements (such as satisfaction, learnability, and accessibility, among others) [322].

Another lens for examining cognitive effort is through the concept of boredom - people are generally averse to cognitive effort, but are equally (if not more) averse to its opposite [323]. The Meaning and Attentional Components (MAC) model defines boredom as having two main components - attention and meaning [69]. This is conceptually aligned with the cognitive and affective components of engagement and the competence and value components common in theories of motivation. The 'attention' component relates to cognitive effort - it represents a balance between the cognitive demands of the task and the available cognitive resources. 'Meaning' refers to the degree to which the task is "congruent with currently activated goals that are both valued and task salient" [69] - i.e. the degree to which it relates to something that the individual wants and is emotionally relevant in that moment. The MAC model provides a breakdown of how these two dimensions combine to provide profiles of boredom that predict how the individual in each profile will act to mitigate the boredom. A novel element of this model is how it distinguishes between pursuit of interest or enjoyment to alleviate boredom by considering the different ways in which cognitive resources can match or differ from demands [69]. For example, given that a task has high meaning, a person will experience enjoyment if they have low cognitive resources and the task has low cognitive demands but interest if they have high cognitive resources and the task has high cognitive demands. If the task is meaningless, the model predicts that low cognitive resources (regardless of cognitive demand) would prompt the individual to seek an enjoyable activity while high cognitive resources would prompt them to seek an interesting activity [69]. The MAC model highlights how motivation to engage could be driven by interest or enjoyment and highlights how availability of cognitive resources influences which is a stronger predictor of behaviour.

Through the examination of this 'effort to use' scale item, we can start to see how different components of engagement can influence each other. For example, micro behavioural engagement with an intervention at a particular time is likely to depend on how well the intervention aligns with an important health goal, how relevant the goal is to the user at that moment, the degree of cognitive resources required to engage with an intervention, and the degree of cognitive resources currently available to the individual. If an intervention cannot

trigger the salience of the health goal and match both low and high cognitive resource states, it is more likely that the individual will disengage.

3.6.2.5 How many components should be included in conceptualisations of engagement?

Empirically, EFAs found that the DBCI-ES fit its expected two-factor solution (behavioural and cognitive-affective) [277,279] but that the TWEETS fit a one-factor solution [278]. The TWEETS authors acknowledged the possibility that engagement only includes one component but felt this was at odds with the theory and could be due to conceptual overlap or the small number of items per component [278]. They also argued that the strong loading of the items onto the one-factor structure does emphasise that - whether the components should be considered distinct or not - engagement does comprise affective, cognitive, and behavioural aspects [278]. Interestingly, a study of engagement in education found a two-factor solution with a different combination of components (affective and behavioural-cognitive) [63], which aligns with the overlap between cognitive and behavioural engagement identified in the previous section and highlights the difficulty of using factor analyses to define the components of engagement. Given the conceptual overlap observed in this analysis of the DBCI-ES and TWEETS, is there value in distinguishing between components of engagement?

There is a history of discussion within psychology on the relationship between affect and cognition in various concepts. For example, there are long-standing unresolved questions around whether emotion is affectively or cognitively-driven [324,325], innate or constructed [324,326–328]. Recently, it has been argued that decoupling affective and cognitive processes does not reflect their close integration at different levels: neural, cognitive, behavioural [326,329]. This integration is explicit in some theories of engagement, like Perski and colleagues' [42]. It is also a core concept in Cognitive Behavioural Therapy (CBT) [330–333]. The close reciprocal links between affect, cognition, and behaviour are represented in the 'cognitive triangle' (Figure 3-1), which highlights how each component can dynamically influence the others [330,333].

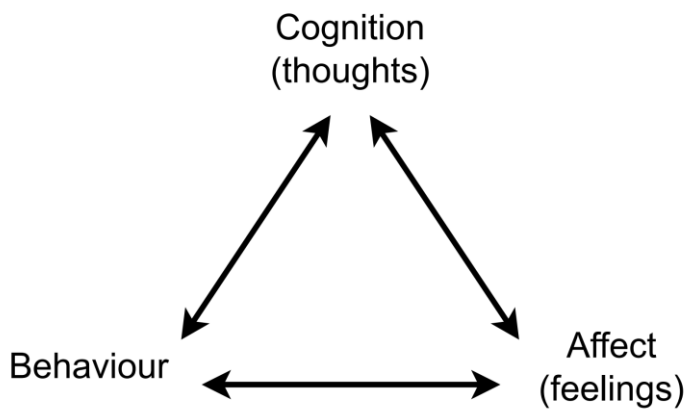


Figure 3-1. Cognitive triangle tool in Cognitive Behavioural Therapy

The idea that an alteration in one component can affect the other two is a key tool in CBT, because it means that each component should be considered to understand a particular issue and that there can be different pathways to addressing these issues [333]. Depending on the issue, context, and individual, one component may be a better starting point to influence the others [330]. However, despite these close interactions, CBT does not conflate the components into one; it is important for patients to recognise and have strategies to address each of the different components depending on context [330]. Using Early and Grady’s [333] example of a fear of elevators, Table 3-4 demonstrates how each component could be targeted in CBT to address the initial state and influence the other components. Similarly, despite the empirical evidence from the factor analyses and their close interactions, there is value in considering affective, cognitive, and behavioural components of engagement separately. Much like how the CBT cognitive triangle emphasises the importance of all three elements in treatment [330,333], making an effort to disentangle the components of engagement in intervention design and evaluation can generate insights that might not be gleaned if engagement was considered as one factor.

Table 3-4. Example application of CBT ‘cognitive triangle’ (adapted from [333])

Triangle component	Initial state	Challenge	Outcome
Cognition	Belief that elevators are dangerous	Provide evidence about the low risk of elevator accidents	Weakened belief of danger leads to more use of elevator and reduced feelings of anxiety
Affect	Feeling anxious	Use breathing exercises to regulate feelings and physiological reactions	Reduced feelings of fear could enable more rational analysis of risks and confidence in ability to use elevator without panic

Behaviour	Avoidance behaviours (taking the stairs)	Ride on an elevator, experiencing no issues	Safe experience using elevator challenges belief that elevators are dangerous and reduces feelings of anxiety

3.6.3 Definitions for this thesis

Based on the literature and the theoretical examination conducted in this chapter, the different components of engagement will be conceptualised for the purpose of this thesis as follows. For each component, the definitions include how that component can relate to micro and macro engagement.

Affective engagement will refer to the positive and negative feelings, emotions, moods, interest, and attitudes (e.g. enjoyment, satisfaction, frustration) related to perceptions of and experiences engaging with the intervention and the behavioural goal [45,46,49,62–66]. It will also include elements of motivation relating to emotions (e.g. pursuit of enjoyment or avoidance of negative emotions) and to the alignment of the intervention and target behaviour to users’ goals and values, in both the immediate moment and in the long-term [64,67–70].

Cognitive engagement will refer to users’ interest and attentional effort invested in the decision-making processes and learning associated with the intervention and the target behaviour [42,49,63,66,67]. It will also include elements of motivation relating to assessments of capability, risks, and benefits (perceived usefulness, suitability, control, and potential for success) [46,62,64,65,70].

Behavioural engagement will refer to the individuals’ interaction with a stimulus or their participation in or performance of a particular activity. This can include quantity and quality of use - frequency, amount, depth, duration, and accordance with intended use in line with intervention goal - of the intervention and performance of the target health behaviour [42,43,45,46,49,66]. Factors that can influence behavioural engagement (such as effort) will be considered in relation to it in case study evaluations, but are not included within the definition of behavioural engagement.

It is clear from even a brief review of the theory that affect, behaviour, and cognition are closely integrated. This is reflected in the definitions included here; I have included concepts of interest and motivation in the definitions of both affective and cognitive engagement. Despite a large body of research into motivation and engagement in the field of education, Skinner and Raine [334] argue that they are insufficiently integrated, leading to confusion

about how they are both distinct yet closely related concepts. For this study, which focuses on engagement rather than motivation, I have included elements of motivation within my definitions of the components of engagement to ensure that this intricate relationship is reflected.

Why have I kept a three-factor model of engagement distinguishing between affective, cognitive, and behavioural engagement despite evidence of their close integration? It is because I think value can be generated by distinguishing between affective, cognitive, and behavioural components. While there is overlap between them - the cognitive and affective elements of interest and motivation provide a good example - they also have conceptual differences that should not be discounted. The TWEETS one-factor solution is likely due to the high degree of conceptual overlap between different components of engagement that occurred in several items; the authors themselves were sceptical that their one-factor solution meant that engagement included only 1 component rather than the hypothesised 3 from the literature. As for the DBCI-ES's combination of affective and cognitive components into one experiential factor, I think that, despite the degree to which they might be correlated, there is value in considering them separately. Much like how the CBT cognitive triangle emphasises the importance of all three elements in treatment, making an effort to disentangle these components can generate insights that might not be gleaned if they were lumped together.

While this disentangling might be difficult to execute in a self-report questionnaire, it has promise for qualitative methods of examining engagement. For example, in an interview, researchers have the space and flexibility to investigate participants' thoughts, feelings, and actions in relation to a DHI and - as in CBT - how each of those influenced the others, enabling a more in-depth understanding of users' processes of engagement. This is not to say that quantitative self-report measures of engagement are useless - research would benefit from continued development of scales like the TWEETS and DBCI-ES and further investigation into how to design items to capture more refined components of engagement. This will be discussed further in [Section 3.9.1](#).

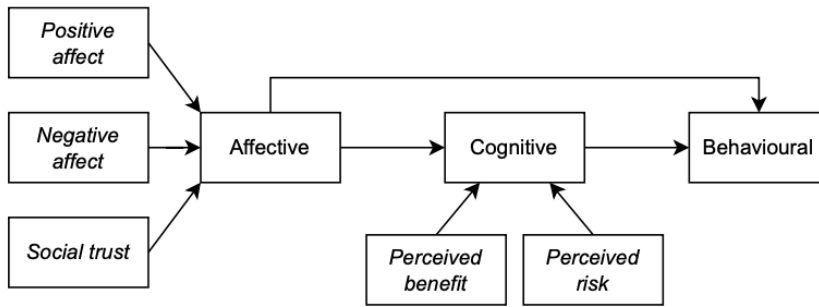
3.7 Evidence for Potential Patterns of Engagement Processes

Up to this point, Chapter 3 has focused on describing and examining two theoretically-based, multifaceted measures of engagement and how that previous research can inform our conceptualisation of the components of engagement and how they are related. The purpose of including definitions is not to provide a definitive answer to the evolving questions about the conceptualisation of engagement and how it aligns with other constructs, but to summarise

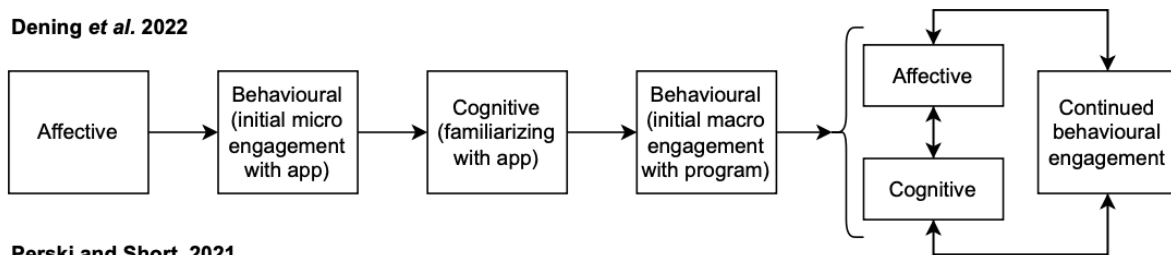
previous conceptualisations in the context of the theoretical examination conducted in this chapter and to clarify how these terms will be used in the rest of the thesis. Based on these insights and previous theory and research, the following two sections will generate a series of hypotheses about how the different components are expected to influence each other over the process of engagement with DHIs and their target behaviours.

Previous research has proposed models of how affective, cognitive, and behavioural responses proceed. Figure 3-2 visualises a few of these models, but as these were not identified via a systematic search, it should not be assumed to be a comprehensive summary. Though not all of these models are specific to digital health engagement, a brief overview of the different processes will be examined to generate hypotheses for how those insights might be generalised to engagement with DHIs and their target behaviours. The first model was derived from a study of user acceptance of self-driving vehicles, which proposed and tested a pattern of how affective and cognitive factors influenced participants' acceptance of road tests and behavioural intentions to use self-driving vehicles [62]. Liu *et al.*'s model suggests that affective factors (positive and negative affect and social trust) can directly and indirectly (via cognitive evaluations of perceived risks and benefits) affect participants' behavioural response. Their study provided evidence to support this hypothesised model, although the direct affective pathway was observed to have a larger influence on acceptance and behavioural intentions than the indirect pathway mediated by cognitive processes [62].

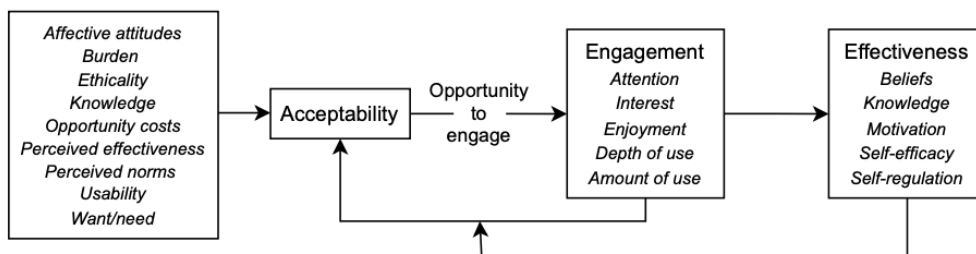
Liu et al. 2019



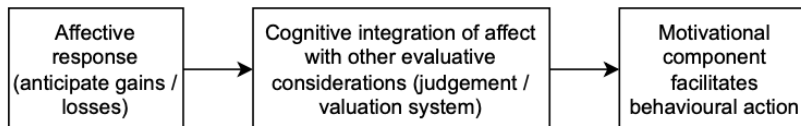
Dening et al. 2022



Perski and Short, 2021



Samanez-Larkin and Knutson, 2015 - AIM framework



Nahum-Shani et al. 2022 - AIM-ACT framework

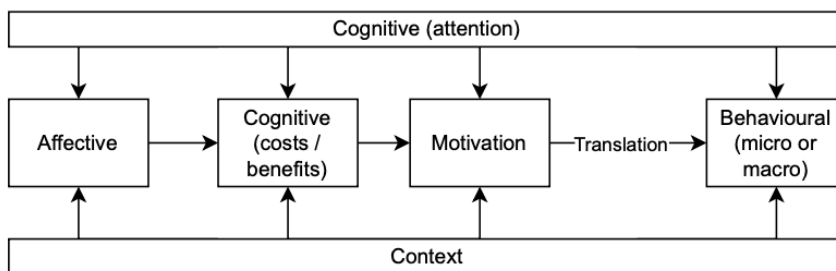


Figure 3-2. Simplified diagrams demonstrating proposed processes of engagement adapted from studies and theories [48,49,62,66,173]

Another study also found evidence for a similar pattern of factors in the context of digital health engagement. Dening *et al.* aimed to examine factors influencing effective engagement and how different components of engagement influenced each other in a web-based dietary intervention for people with Type 2 diabetes [66]. Qualitative data from semi-structured interviews (SSIs) indicated that participants' initial motivation to engage with the intervention

was affective - concern about health risks and needing medication and gratitude for the opportunity to participate in the program. After this initial affective engagement, participants cognitively engaged with the program structure and content, which led to behavioural engagement with the intervention guidelines and target behaviours [66]. As participants moved through the program, the authors observed that affective, cognitive, and behavioural engagement combined and influenced each other - for example, positive emotions associated with achievement of goals and visible progress supported further behavioural engagement with the intervention and behaviour. They also observed that while negative affect could have a negative influence on behavioural engagement, sufficient cognitive engagement with the educational elements of the program could still influence behavioural engagement [66].

Aligned with these studies, Perski and Short proposed a complex system perspective of how acceptability, engagement, and effectiveness interact [48]. They draw on existing theories of acceptability (including the Theoretical Framework of Acceptability [320] and the Technology Acceptance Lifecycle, which posits a process of acceptance that includes pre-use acceptability, initial use acceptability, and sustained use acceptance [335]) to highlight key factors that contribute to acceptability as an emergent property. Viewed from a strictly behavioural perspective, engagement is distinct from acceptability. However, when considering the definitions of affective and cognitive engagement outlined in [Section 3.5](#), there is conceptual overlap between engagement and acceptability [48]. Many of the acceptability factors that Perski and Short identify could fall under the definitions of affective and cognitive engagement. For example, affective attitudes (how the user feels about the intervention) and ethicality (the alignment of the intervention with the user's values) both align with the definition of affective engagement [48,320], while factors such as usability, burden, opportunity costs, and perceived effectiveness relate to the cognitive demand of engaging with the intervention and the assessment of risks and benefits of using it [48,320,335]. Thus, the model suggests that both cognitive and affective engagement influence initial decisions around behavioural engagement, which then influences the updating of cognitive and affective perceptions of the intervention as an ongoing system of influences [48].

Other studies have examined the question of how engagement proceeds from different perspectives. A model of how affective and cognitive processes support decision making was developed based on neuroscientific evidence and supported by a range of empirical evidence [173]. It was recently extended to help explain digital health engagement [49]. The original Affective-Integration-Motivation (AIM) Framework proposed a series of hierarchical and

sequential processes that take place to generate decisions: affective processing of the stimulus anticipates potential gains and losses, which is followed by cognitive integration of this affective response with other factors to produce a value judgement, which in turn motivates the individual towards a particular action [173]. In the application of this framework to a digital health context, Nahum-Shani *et al.* added an additional three influences to the framework to make it the AIM-ACT framework: attention, context, and translation of motivation to behaviour [49]. They posit that attention is a prerequisite for engagement at all stages of the process - affect, integration, and motivation. Without attention to the intervention or target behaviour, it is unlikely to elicit meaningful affective responses, generate cognitive processes of learning or decision-making, or trigger engagement with the behaviour. All of these elements of the process of engagement can also be shaped by the context in which they are occurring (the environment, other demands on attention or motivation, etc.), which also contributes to whether an individual's motivation to engage in a behaviour is successfully translated into action [49].

With all these proposed processes, it is important to consider what individuals are engaging with. Different types of DHIs may blur the line between engagement with the intervention (micro) and engagement with the target behaviour (macro). For example, a mobile health app targeting physical behaviour might send reminders or prompts via the app that encourage the user to go for a run [49]; in this case, engagement with the reminder (noticing and paying attention to it, the emotion it triggers, and whether the user clicks on it or swipes it away) is relatively distinct from engagement with the behaviour (e.g. weighing costs and benefits of going for a run when a notification is received or making a plan to go later). In other cases, engagement with the intervention and behaviour might be more closely related [49]; for instance, if a mental health app sends a notification that contains a short breathing exercise or meditation where the app guides the user through the behaviour, then the user can be engaged with the notification, the mental health content delivered via the app, and the health behaviour [49]. This highlights the gradients of engagement within the broader distinction of micro and macro engagement and is captured to some extent in Cole-Lewis *et al.*'s distinction of engagement with the UX/UI, BCTs, and target behaviours (Figure 1-1a) [50]. It is clear that engagement with the intervention and the target behaviour are related and influence each other over time. In many cases, the aim of a DHI is to build sufficient engagement with the behaviour that it becomes habitual, such that the intervention is no longer regularly used and re-engagement only occurs when or if needed [14]. Denning *et al.*'s study found that after the initial stages of engagement, affective, cognitive, and behavioural engagement combined and influenced each other [66], in line with the CBT cognitive triangle discussed in [Section 3.5](#).

The next section will build upon these theories and existing evidence to generate hypotheses for how the processes of engagement with DHIs progress.

3.8 Hypotheses Regarding the Process of Engagement

The review of the theories and evidence in [Section 3.7](#) provides the basis for a series of hypotheses about how engagement with DHIs (and their target behaviours) proceeds. It is a dynamic process that includes affective, cognitive, and behavioural engagement with various levels of the intervention and target behaviour and can fluctuate over time. Aligned with the concept of acceptability, we would expect that if there is sufficient affective or cognitive engagement with the *idea* of the DHI (e.g. app, wearable, conversational agent), this would drive initial behavioural engagement with it, given the opportunity. Behavioural engagement with the digital tool could be driven by either positively or negatively-valenced emotions or affective attitudes - for example, potential users could be motivated to engage with it because they expect the intervention to be enjoyable, because it supports a goal that is important to them, or because they worry about negative consequences of their current health behaviours. From a cognitive perspective, potential users would need to be aware of the intervention (having sufficient attention to recognise it); beyond that, interest in the intervention and a judgement that it will be more beneficial than costly would be expected to increase likelihood of initial micro behavioural engagement (with the user interface). These predictions are summarised in hypotheses 1a to 1c:

H1a: Given that an individual is aware of and able to access a digital health intervention, initial micro behavioural engagement will be driven by affective engagement with the idea of the intervention either directly or indirectly by influencing cognitive engagement (the individual's appraisal of the expected benefits, risks, and demands of intervention).

H1b: Both positively- and negatively-valenced emotions can support affective engagement and initial micro behavioural engagement.

H1c: Only positive cognitive engagement will support initial micro behavioural engagement.

From the point of initial behavioural engagement with the intervention, users' experiences of behaviourally engaging with it will influence how they think and feel about the digital tool. For instance, a user interface that is too cognitively demanding and difficult to use could reduce interest or cause frustration, increasing the likelihood of behavioural disengagement from the tool. Likewise, a dislike of the intervention's aesthetic design or a perception that it

does not align with the users' values and goals could also result in behavioural disengagement. Conversely, an intervention that is aesthetically appealing and easy to use while providing new information or activities that are interesting and aligned with the users' goals is more likely to support positive affective and cognitive engagement with the intervention, and thus help sustain behavioural engagement with the interface.

H2a: An individual's initial behavioural engagement with a digital health intervention will influence their affective and cognitive engagement with the intervention by providing further experiences which generate subsequent affective responses and cognitive assessments.

H2b: The intervention needs to be sufficiently affectively OR cognitively engaging to maintain behavioural engagement with the intervention interface and content.

If there is sufficient affective or cognitive engagement during a user's initial behavioural engagement with the intervention, behavioural engagement is likely to extend as the user explores the various features, information, or BCTs comprising the intervention. If users invest their attention in the intervention and it is an enjoyable or valuable experience that aligns with their goals and desires and the suggested actions are considered to be within the users' ability, we would expect micro behavioural engagement with the intervention to continue and motivation to engage with the health behaviour that the intervention intends to support to increase (macro behavioural engagement).

H3a: Affective or cognitive engagement with the intervention content will increase an individual's motivation to engage in the target health behaviour.

H3b: The translation of motivation into behaviour is dependent on sustained cognitive engagement (attention) and influenced by the individual's context.

Often, the purpose of DHIs is to help support users to engage with health behaviours that they are otherwise not sufficiently engaging in. The features and content of these interventions should therefore be designed to support affective, cognitive, and behavioural engagement with that target behaviour. The way that various interventions aim to increase engagement with the behaviour is likely to vary depending on the target behaviour, population, and context as well as the main barriers that people experience to engaging with that behaviour. The components of engagement that need support may also vary; for example, if a person who wants to engage in more physical activity knows that this will be beneficial for their health but really does not enjoy exercise, the intervention may need to help improve the users' affective engagement with exercise (rather than increasing cognitive engagement by

emphasising health benefits or risks). Therefore, while positive affective and cognitive engagement with the behaviour are expected to increase behavioural engagement, the influence of these components is likely to differ depending on individuals and context.

As users become more engaged with the health behaviour, those experiences are expected to form a feedback network between micro and macro affective, cognitive, and behavioural engagement. For many interventions, increased engagement with the health behaviour may result in decreased engagement with the intervention. For example, an intervention may become less cognitively engaging once a user has learned all of the information available or mastered the skills that it aimed to build. If engagement with the health behaviour falters, there may be re-engagement with the intervention (micro behavioural) to provide support for the users’ goals, if there is still sufficient affective or cognitive engagement.

H4a: As the individual engages with the health behaviour, affective, cognitive, and behavioural components of engagement with the health behaviour will all influence each other to determine whether macro behavioural engagement is sustained.

H4b: High macro engagement with the health behaviour may result in decreased micro engagement with the intervention as the health behaviour becomes more habitual.

H4c: Decreasing macro engagement with the health behaviour may prompt an increase in micro engagement with the intervention, if there is still macro affective or cognitive engagement with the health behaviour goal.

The series of hypotheses is summarised in the proposed model of engagement with DHIs depicted in Figure 3-3.

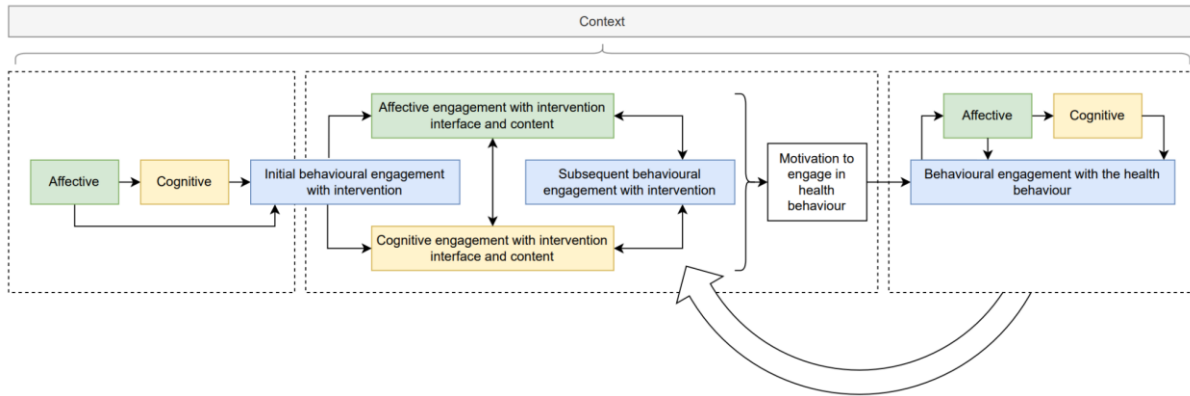


Figure 3-3. Proposed model of the process of engagement with a DHI and its target behaviour

3.9 Implications for Broader Research

Based on the analyses conducted in this chapter, recommendations were generated to help researchers determine which measure of engagement is the best fit for their study and to highlight key areas for future research, including how further scale development could support evaluations of engagement with DHIs.

3.9.1 Recommendations for scale selection and future measure development

The comparison of the scales highlights a few key considerations for researchers who are evaluating engagement with a DHI, primarily what the study is looking to answer and what other data can be captured. The DBCI-ES is recommended for researchers who are specifically looking to capture engagement with the digital tool because of its narrower scope focusing on micro engagement. It may also be beneficial for studies that are unable to capture system usage data, as the scale includes a self-reported approximation of system use. The caveat is that evidence for the correlation between objective and self-reported usage data was inconsistent, indicating that the behavioural subscale may lack validity [277]. On the other hand, the TWEETS is recommended for researchers who are looking for a holistic measure that captures various elements of engagement but are not aiming to investigate or compare specific components. This is because its items capture micro and macro engagement, as well as cognitive, affective, and behavioural components, but both our analysis and the authors' EFA indicated that the affective, cognitive, and behavioural elements should be examined as a whole rather than as separate components. This can add an additional layer of information for studies that are already capturing system usage data and extend the assessment beyond micro engagement. For all studies, we recommend using mixed-methods (if feasible) to enable a more in-depth investigation of the different components of engagement, how they interact, and what factors influence them [14,47].

Despite providing useful new tools for assessing engagement as a multifaceted construct, the gaps in the TWEETS and DBCI-ES identified in this analysis demonstrate continued room for improvement. While both scales are useful, theoretically-based tools for examining engagement as a multifaceted concept, both lack the ability to examine engagement components as separate constructs. The differing interpretations of the components of engagement across the two scales emphasises the need to establish clarity and agreement around their definitions. Future scales developed should include clear definitions of the components of engagement that it aims to capture, ideally derived both deductively from

theory and inductively from the target population [307]. This would facilitate judgments of whether a particular scale is the appropriate measure for a particular study aim.

This analysis also highlighted the relative lack of consideration of macro engagement in digital health engagement measures. The DBCI-ES specifically excluded it, while the TWEETS did not clearly distinguish between micro and macro engagement. When it comes to evaluations of engagement in digital health, measures tend to focus on micro engagement with a DHI (such as system use). This could be because macro engagement is already captured in other ways as behavioural outcomes (e.g. for physical activity, as step count or minutes spent exercising per day) or because it is assumed to be a separate concept. Although macro engagement with a health behaviour may be captured with other measures, these tend to focus on the behavioural component, either via objective data collection measures such as sensor data or via self-report measures that ask “how much” or “how often”. The development and potential benefits of a scale that captured affective, cognitive, and behavioural components of micro and macro engagement separately should be explored, as this could be a potentially valuable tool for understanding how engagement with the intervention and the behaviour are related. Although this would be beneficial for enabling comparison and meta-analysis of engagement across studies, there are also advantages to having a variety of measures of engagement available. Engagement is theoretically complex and multi-faceted; a variety of measures might be needed to address different research questions, intervention characteristics, and implementation contexts [46].

Future scale development should also explore at what periods it is most appropriate to measure engagement. The DBCI-ES behavioural subscale focused on most recent use - this could potentially reduce recall bias, but might result in inaccurate reporting if users felt their most recent use was not reflective of their typical use. If using a similar measure of micro behavioural engagement, the testing of future scales should assess whether validity could be improved by asking about engagement over a period of time rather than the most recent case. Timing is also a consideration for the TWEETS; initial encounters with a DHI may be important in determining whether users continue to engage with it to try and change their health behaviour, but the framing of the TWEETS items assumes some user experience engaging with the DHI in pursuit of a goal. This could miss critical information about first impressions of the app, highlighting again the importance of having various measures to address particular stages and components of engagement.

The challenge for developing additional scales will be the difficulty in differentiating any categories defined. Concepts like motivation, interest, cognitive effort, and flow (a state of

intrinsically-rewarding absorption in a task characterised by involvement, enjoyment, and attention [44]) include elements from various components [65,67]. Some conceptualisations have attempted to address this by removing these elements from their definitions - for example, defining cognitive engagement as “thinking and paying attention” without any motivational aspects [63], but this may not capture the whole of the concept. The definitions generated for this thesis purposefully included these concepts as elements of both affective and cognitive engagement to highlight the integration between all of these different constructs. It will be necessary to thoroughly test developed scales to determine if elements are separable. Although any distinctions drawn will be to some degree artificial, a measure with clear components would enable new questions to be examined more broadly - for example, which features of an intervention influence which components of engagement or whether micro engagement affects macro engagement with behaviour change. This would enable strategies to be incorporated into the intervention to support more effective engagement. For example, an intervention might be cognitively engaging (easy to use with interesting content) but affectively disengaging (aesthetically unappealing interface or negative tone). Some users may be sufficiently cognitively engaged to overcome the affective disengagement, but improving the aesthetic or tone would mitigate barriers to continued engagement with the intervention and its target behaviour.

3.9.2 Strengths and limitations

A strength of this chapter is that it provides a novel in-depth exploration of the definitions, relationships between, and measures of the components of engagement in digital health. The examination of several recent reviews of measures used to assess digital health engagement provided a wide pool of self-report measures that were considered for inclusion in the chapter against a set of criteria. This helped to ensure that the study did not exclude relevant measures in the analysis and that the methods were transparent and replicable. The recommendations generated based on the analysis are a valuable contribution for researchers selecting a measure of engagement and for future considerations about how engagement can be better evaluated to enable currently unanswered questions to be addressed.

A limitation of the chapter is that similar methods were not applied in the second part of the analysis that assessed proposed models of the relationships between affective, cognitive, and behavioural components ([Sections 3.7](#) and [3.8](#)). Upon reflection, a more systematic process of identifying studies that proposed or examined these patterns would have likely identified additional sources of evidence, strengthening the quality of the analysis and informing the theoretically-based hypotheses that were generated. However, unlike the analysis of the

measures of engagement, where several recent reviews were available and could be used to identify relevant measures for inclusion, it was difficult to identify previous literature examining the interactions between components of engagement. Although this chapter drew on a variety of theoretical and empirical sources to generate insights on the conceptualisation of different components of engagement in digital health evaluation and measurement, it did not systematically review the literature. A systematic review of evidence from various fields regarding the patterns of association between affective, cognitive, and behavioural components (of engagement, or decision-making more generally) would be a valuable addition. The analysis of the identified papers discussing patterns of engagement was conducted to help inform the thesis hypotheses and to generate a theoretical comparator for the empirical mappings of the process of engagement. It was not intended to represent the sum of the state of knowledge; additional sources of evidence and theory may shed additional light or challenge some of the conclusions.

3.9.3 Conclusions

This chapter set out to examine the measurement of engagement as a multifaceted concept from a theoretical perspective, provide recommendations for when these measures might be most appropriate, and explore how this theoretical background can inform conceptualisations of the components of engagement and hypotheses of the process of engagement with DHIs. This theoretical investigation highlighted the strengths and limitations of currently available measures of engagement with DHIs, which have implications for their use in future studies. Few of the scales being used to evaluate engagement with digital health were developed specifically for digital health or theoretically-based on a multifaceted conceptualisation of engagement. In explicitly considering the multifaceted nature of engagement, both the DBCI-ES and the TWEETS represent an advance in measures of engagement; however, there are still gaps in their ability to capture the different components of engagement. Although the TWEETS divides its items into affective, cognitive, and behavioural components, this analysis indicates that many of them blur these boundaries; likewise, the framing of the items so that they capture intervention engagement in relation to behavioural goals blurs the boundary between micro and macro engagement. In contrast, the DBCI-ES focuses on micro engagement with the intervention and has a two-factor solution that splits the measure into behavioural and subjective sub-scales. Although progress has been made in addressing the need for theoretical- and evidence-based measures of engagement as a multifaceted concept, the DBCI-ES and TWEETS fit different research needs and there is still a need for further

scale development to address gaps in our ability to comprehensively capture and differentiate components of engagement.

Despite evidence from other fields demonstrating substantial integration and overlap between the components of engagement, there are practical benefits that can be obtained through their investigation as distinct elements. Through the process of trying to tease these components apart, the ways in which they are connected can become clearer. This could enable a better understanding of processes of engagement and how they could be supported. There are several unanswered questions in our understanding of digital health engagement: whether engagement has distinct components, how those components should be defined and - if distinct - how each can be measured, and how the components influence each other over the process of engagement. Previous research has made headway in addressing these questions, but there is still relatively little examination of the relationships between different components over the process of engagement with DHIs. The definitions and hypotheses generated in this chapter will form the basis for the rest of the analyses conducted in this thesis, which aim to explore this question and provide a hypothesised mapping of the relationships between engagement components and a framework for how future research can continue to investigate these questions.

Chapter 4: Methods

4.1 Chapter Overview

Chapter 3 outlined the current issues with the conceptualisation and measurement of engagement in digital health. Two self-report scales aiming to measure engagement as a multifaceted concept were compared to generate insights about engagement's components and how they can best be evaluated. Based on this analysis, and an exploration of previous literature proposing models of interactions between affective, cognitive, and behavioural processes, I generated definitions of the components of engagement that will be used in this thesis, hypotheses for how the components of engagement are related over the process of digital health engagement, and recommendations for the use of current, and development of future, measures of engagement.

Chapter 4 provides an overview of the case study research design, the rationale for why it was chosen to investigate the research question, theoretical frameworks used, and a rationale for and description of the cases selected. It then details the methods of the individual case studies examining engagement with a family-focused app for childhood obesity, a mental health conversational agent (CA) app, and an autonomous CA for post-cataract surgery follow-up. The results of these individual case studies will be presented in Chapter 5 and include the original analyses addressing the specific case study objectives and the re-analysis of the individual case study data to enable the thesis cross-case analysis. The final section outlines these cross-case analysis methods, which will address the main thesis research question and will examine the generalisability of the findings. These results will be presented in Chapter 6.

Publications: Sections of Chapter 4 have been drawn from previously published papers; references and author contributions for each of the papers that were used in the preparation of this chapter are listed below.

Milne-Ives M, Rahman E, Bradwell H, Baines R, Boey T, Potter A, Lawrence W, van Velthoven MH, Meinert E. Barriers and facilitators to parents' engagement with and perceived impact of a childhood obesity app: A mixed-methods study. *PLOS Digital Health*. 2024 Mar 27; 3(3):e0000481. doi:[10.1371/journal.pdig.0000481](https://doi.org/10.1371/journal.pdig.0000481)

Authors' contributions: MMI collected data, carried out the initial analyses, and drafted the initial manuscript, and critically reviewed and revised the manuscript. ER, AP, WL, and MvV conceptualised and designed the study and critically reviewed and revised the manuscript. HB, RB, and TB, analysed data

and critically reviewed and revised the manuscript. EM conceptualised and designed the study, supervised data collection and analysis, and critically reviewed and revised the manuscript.

de Pennington N, Mole G, Lim E, **Milne-Ives M**, Normando E, Xue K, Meinert E. Safety and Acceptability of a Natural Language Artificial Intelligence Assistant to Deliver Clinical Follow-up to Cataract Surgery Patients: Proposal. *JMIR Research Protocols*. 2021 Jul 28; 10(7):e27227. doi:[10.2196%2F27227](https://doi.org/10.2196%2F27227)

Authors' contributions: NP conceived the study topic and designed and drafted the proposal. GM, EL, MMI, EN, and KX contributed to drafting and revising the proposal. Final revision was conducted by EM.

Meinert E, **Milne-Ives M**, Lim E, Higham A, Boege S, de Pennington N, Bajre M, Mole, G, Normando E, Xue K. Accuracy and safety of an autonomous artificial intelligence clinical assistant conducting telemedicine follow-up assessment for cataract surgery. *eClinical Medicine*. 2024 Jul 4; 73:102692. doi:[10.1016/j.eclinm.2024.102692](https://doi.org/10.1016/j.eclinm.2024.102692)

Authors' contributions: NdP conceived of the study topic. The study design was developed by NdP, GM, EL, MMI, EN, KX, EM and AH. KX and EN led patient recruitment at the two study sites. AH and EL supervised Dora R1's calls and compiled the clinical data. MMI conducted the usability survey and MMI and EM completed participant interviews. Statistical analysis of the quantitative data was conducted by the University of Plymouth, MMI, SB and EM conducted the thematic analysis of the qualitative data. MB was responsible for the cost analysis. EM set the structure for the manuscript, MMI drafted the paper based on this framework, and it was reviewed by all authors.

Meinert E, **Milne-Ives M**, Taylor E, Inkster B, Paik A, Ananthakrishnan A, Orr M, Costelloe CE, Shankar R. Real-world testing of an artificial intelligence conversational agent as an early intervention and support tool in the mental health referral care pathway: A randomised controlled trial. *International Journal of Social Psychiatry*. 2024; Under Review.

Authors' contributions: All authors contributed to the study design. Participants were recruited and screened and quantitative data was collected by AP. MMI conducted the participant interviews. CC and MO conducted statistical analysis of the quantitative data and MMI and AA conducted the

thematic analysis of the qualitative data. EM set the structure for the manuscript, MMI and CC drafted the paper based on this framework, and it was reviewed by all authors with final revisions from EM.

4.2 Research Design

4.2.1 Choosing a research design

The process of determining the research design for the thesis was intricately linked with the development and refinement of the research question. After conducting the systematic review of the literature (Chapter 2), I identified a gap in knowledge regarding how different BCTs are associated with different components of engagement and generated initial research questions around how specific BCTs could influence the different components of engagement. This could have been evaluated quantitatively - for example, by conducting a factorial randomised controlled trial to compare the influence of particular BCTs on engagement [336] - or by conducting in-depth qualitative analyses of users' preferences for and perceptions of particular BCTs in relation to engagement. However, the process of laying out the theoretical propositions ([Section 1.6.1](#)) revealed the importance of taking a step back; before examining how BCTs (or other features) can influence the different components of engagement, it is first necessary to have a clear understanding of how those components interact over the process of digital health engagement. This was confirmed by the analysis conducted in Chapter 3, which highlighted how, despite the recent advancements in the conceptualisation of engagement, there is still a need for a framework mapping these relationships. Such a framework would facilitate the future evaluation of particular associations between intervention features and components of engagement, strengthening evidence for how to best support digital health engagement.

The final thesis research question - focused on understanding how the different components of engagement with DHIs influence each other - could still have been addressed by various methods. From early in the thesis design process and literature review, it was clear that a mixed-methods design would provide the best approach for investigating engagement. The dominance of system usage data as a measure of engagement is an important limitation of many digital health evaluations. Mixed-methods designs are particularly important when examining engagement because of its multifaceted nature; using a combination of various methods - including qualitative methods, questionnaires, system use data, among others - can help capture different components and provide a more holistic understanding of engagement and enable triangulation to improve the credibility of findings [35,44,46,171].

Mixed-methods can be included in most study designs and “how” focused research questions can also be addressed by different methods, such as experiments and case studies [337]. Given the lack of previous evidence about factors specifically affecting the different components of engagement, it would have been difficult to design an experiment that manipulated the levels of affective, cognitive, and behavioural engagement so that their influences on each other could be untangled. Additionally, this would likely fail to represent real world processes of engagement with a DHI. For these reasons, I considered instead how a case study approach could be used to examine my research question.

A case study design was an appropriate methodology for this investigation because it enables an in-depth exploration of the case in its real-world context [337]. This is important for the study of engagement with DHIs because contextual factors can have an important influence on how an intervention is adopted and used in the real world, which can have a significant effect on its impact [338–340]. A controlled experiment of engagement and use of the interventions would be unlikely to reflect real-world engagement, which is the context in which the intervention needs to be effective. Beyond the introduction of the intervention, an experimental design would not have been appropriate for an examination of how different components of engagement were related, as users’ experience of these components of engagement cannot be experimentally controlled. Traditional experimental designs are also limited in their ability to investigate how and why an intervention engaged or did not engage users; case studies using mixed-methods are more able to examine these questions and enhance our theoretical understanding.

Initially, the idea for the thesis design was to focus on the mixed-methods evaluation, refinement, and re-evaluation of a particular digital health evaluation (NoObesity). The issue with this design was that it was focused on only one case, which would have limited my ability to assess whether my findings in that case would be transferable to engagement with other types of DHI. Instead, an exploratory multiple case study design was chosen to investigate the process of engagement and to examine the transferability of patterns of engagement across different DHI types and clinical areas [337]. A multiple case study approach strengthened the quality of the analysis by enabling replication and cross-case analyses to generate a generalisable framework of the patterns of association between different components of engagement [337].

4.2.2 Exploratory multiple case study design

The exploratory multiple case study design included three cases and followed Yin's procedures for a multiple case study approach (Figure 4-1) [337]. The rationale for the specific case selection is described in [Section 4.4.1](#). Each case was examined individually using implementation science-centred approaches and mixed-methods ([Section 4.5](#)). Although the overall studies included a variety of outcomes, the case studies here used qualitative SSI data and quantitative questionnaire and system use data to investigate factors associated with engagement, acceptability, and usability.

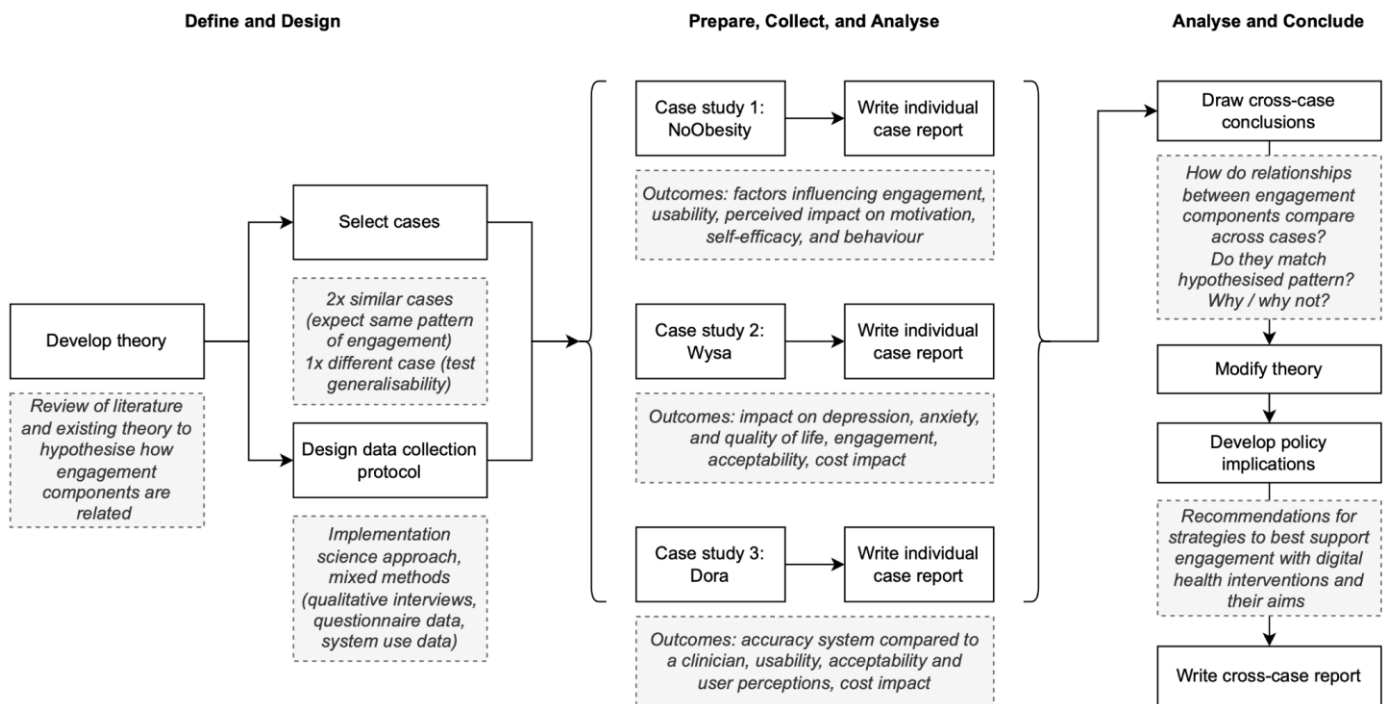


Figure 4-1. Procedures of a multiple case study approach (adapted from [337])

4.3 Theoretical Frameworks

*Parts of this section have been previously published in [77].

Five main theoretical frameworks were used to guide the individual case study designs and analyses (Table 4-1).

Table 4-1. Theoretical frameworks included in the case studies

Framework	Case Study 1	Case Study 2	Case Study 3
A multi-faceted conceptualisation of engagement [14,42,45,49–51]	✓	✓	✓

Non-adoption, Abandonment, Scale-up, Spread and Sustainability (NASSS) framework [341]	✓	✓	✓
Reach Effectiveness Adoption Implementation Maintenance (RE-AIM) framework [342]	✓	×	×
Theoretical Framework of Acceptability [320]	×	✓	✓
Capability, Opportunity, Motivation - Behaviour (COM-B) model [53,170]	✓	×	×

The Non-Adoption, Abandonment and Challenges to the Scale-up, Spread and Suitability (NASSS) framework [341] provides a structure for examining factors influencing long-term success. The Reach Effectiveness Adoption Implementation Maintenance (RE-AIM) framework [342] is a method which seeks to understand the impact of a technology within a given implementation context. Both frameworks emphasise the consideration of the multiple levels that influence adoption and sustainability and were used to inform initial study designs, to ensure that a comprehensive set of individual- and system-level factors were explored [341–344]. The adoption, implementation, and sustainability outlook of these frameworks is well aligned with investigations of engagement because interventions can only add value if they are successfully adopted and used. Although engagement with DHIs is often examined on an individual level, these frameworks emphasise the consideration of the multiple levels - individuals and systems - that influence adoption and non-adoption and help provide a broader perspective of what can influence individual-level engagement. The RE-AIM and NASSS frameworks were synthesised in the first case study, as its study aims had a particular emphasis on understanding real-world implementation, adoption, and use. The main aims of the second and third case studies were focused primarily on effectiveness, accuracy, and safety; as such, the NASSS framework was included in study design to capture the implementation perspective, but the additional considerations added through the synthesis of RE-AIM were not needed.

The Theoretical Framework of Acceptability (TFA) was used to help structure the SSI analyses in the second and third case studies [320]. This was a learning from the first case study, which did not use a theoretical framework to structure the interview topic guide, to help ensure comprehensive data was captured. This framework was chosen because it was developed specifically for health interventions, captures a multifaceted view of acceptability (with seven components), emphasises the importance of considering different time points when evaluating acceptability, and is highly cited in the field. As discussed in Chapter 3, acceptability can influence engagement with DHIs and many of the factors influencing

acceptability relate to cognitive or affective engagement with the idea of a DHI; for example, affective attitude and ethicality (related to affective engagement) and burden, perceived effectiveness, and opportunity costs (related to cognitive engagement).

In contrast, the Behaviour Change Wheel, and primarily its component Capability, Opportunity, Motivation - Behaviour (COM-B) model, was used to provide a theoretical structure for the thematic analysis for the first study. This framework was used instead of the TFA because of the different questions being examined by the case studies. The first case study was specifically examining barriers and facilitators to engagement with and perceived impact to inform a redesign of the DHI [170]. The COM-B model posits that capability, opportunity, and motivation interact to drive behaviour and was chosen because it was developed specifically to help support the design and evaluation of behaviour change interventions [53] and is theoretically linked with the BCT taxonomy [170,179].

4.4 Case Selection

4.4.1 Rationale for case selection

All three cases selected represent different fields of healthcare where digital technologies are being introduced to help support a gap in the availability and accessibility of healthcare services: childhood obesity, mental health, and post-surgical follow-up. The clinical need for each of these cases is described in [Section 1.5](#). Engagement is an important factor influencing the DHIs' potential benefits for patients, clinicians, and the healthcare system. Investigating engagement with these interventions can provide insight into how different components of engagement are related and how barriers to those components can be mitigated to optimise effective engagement with interventions and target behaviours.

The value in selecting cases from three different clinical areas was that it provided a better opportunity to examine the transferability of the individual case findings to a more general framework (Table 4-2). The challenge with including three disparate cases (in terms of clinical area, intervention type, and purpose) is that the comparability of the cases in the cross-case synthesis had to be carefully examined ([Section 6.3](#)). The three cases were selected to enable both direct (literal) and theoretical replication [337]. Literal replication expects similar cases to produce similar results, while theoretical replication expects different results for predictable reasons [337]. The first two cases are similar in terms of intervention type and aim, despite focusing on different clinical needs, so we would expect similar patterns of interaction between the different components of engagement if these patterns are dependent on intervention type and aim and not specific clinical context [337].

Table 4-2. Characteristics of the DHIs

Characteristics	Case 1	Case 2	Case 3
<i>Name</i>	NoObesity	Wysa	Dora (R1)
<i>Clinical area</i>	Childhood obesity	Mental health	Ophthalmology (cataract surgery)
<i>Intervention type</i>	Mobile app	Mobile app	Telephone call
<i>Conversational agent</i>	No	Yes	Yes
<i>Purpose</i>	Behaviour change	Behaviour change	Clinical assessment
<i>Intervention aim</i>	Increase positive weight- related health behaviours (physical activity, diet)	Build mental resilience skills	Conduct post-surgical follow-up screening to improve clinical efficiency
<i>Duration</i>	Extended	Extended	One-time

The third case is different to the first two: it is an administrative intervention delivered over a one-time telephone call, rather than a behaviour change intervention delivered by a mobile app that users' are intended to access regularly. Engagement with this type of intervention is expected to be different in some ways from engagement with behaviour change-focused DHIs. The third case includes both literal and theoretical replication: based on the theoretical investigation conducted in Chapter 3, it is hypothesised that the process of engagement that drives the initial behavioural engagement with the intervention is likely to be the same. The expected 'contrasting result' is that as this case only requires a one-time (or very occasional) micro engagement and no macro engagement, it will lack the progression to engagement with a health behaviour. This may alter the importance of various components in supporting micro behavioural engagement, as the main aim of the intervention is to keep users engaged for the duration of the call and have a sufficiently positive experience that they would be willing to use it again. However, although the relative importance of individual components may vary, up to the point of macro behavioural engagement, the pattern of how the components of engagement interact is expected to be the same ([Section 3.8](#)).

The inclusion of these three cases in the multiple case study design will enable an examination of whether the hypothesised patterns of interaction between the different engagement components are observed in a real-world context and whether they are generalisable to different types of DHIs.

4.4.2 Detailed description of interventions

*Parts of this section have been previously published [77,151,345] or submitted for publication [129].

4.4.2.1 Case 1: NoObesity

The “NoObesity” app was developed by Health Education England (HEE) to support the prevention and management of childhood obesity. The term HEE is used as that was the organisation at the time of the study, but the organisation has since been incorporated into the Workforce Training and Education Directorate, NHS England. The impetus for the app came from a gap in workforce training identified by healthcare professionals (HCPs): they have an important role in helping manage childhood obesity but can find it a difficult and sensitive topic to discuss with families and lack sufficient education and training in this skill [346–350]. Two linked apps were developed in collaboration with the Universities of Bournemouth and Southampton in 2017-2018 - one for families and one for healthcare professionals [351]. The HCP version of the app was developed to provide a tool to increase HCPs’ skills and confidence in communicating about weight and healthy lifestyles and enable them to connect and work more collaboratively with families [352,353] (Figure 4-2). For families, NoObesity supported goal setting - identifying potential barriers and strategies for overcoming them - and enabled families to track their progress and learn more about healthy behaviours through advice and games [352]. The apps could link HCP and family accounts to enable HCPs to work collaboratively with the family to identify barriers and develop personalised strategies for behaviour change. No behavioural theory was used in its initial development, but it was created in collaboration with HCPs [351] and included several BCTs within the app features ([Appendix 4-A](#)).

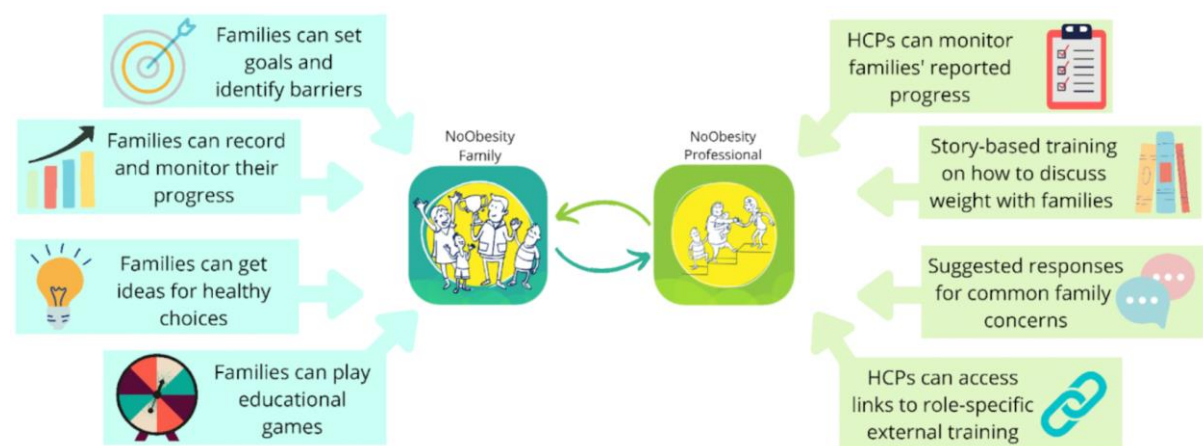


Figure 4-2. NoObesity version 1 app features

4.4.2.2 Case 2: Wysa

Wysa is a guided self-help and triaging tool delivered to patients via an Android or iOS app or web widget to provide additional support to individuals' usual care when waiting for mental health treatment. The app asks users to provide information and goals to personalise recommendations and collects routine outcome measures aligned to existing service referral measures [136] - the Patient Health Questionnaire (PHQ-9) [354] and General Anxiety Disorder (GAD-7) tool [355]). Self-help support is provided using AI-enabled tools - with natural language processing, conversation engines, a predetermined content library, and evidence-based therapeutic exercises and videos (Figure 4-3). Data is pushed directly into the patients' iaptus Electronic Patient Record (EPR) by an Application Programming Interface (Mayden Prism system) and can generate automated safety alerts. Compared to other mental health apps, it has been rated highly on privacy and security [356,357]. Previous evaluations of Wysa have found evidence for higher improvement in more engaged users [358] and positive feedback from users and clinical staff [359,360].

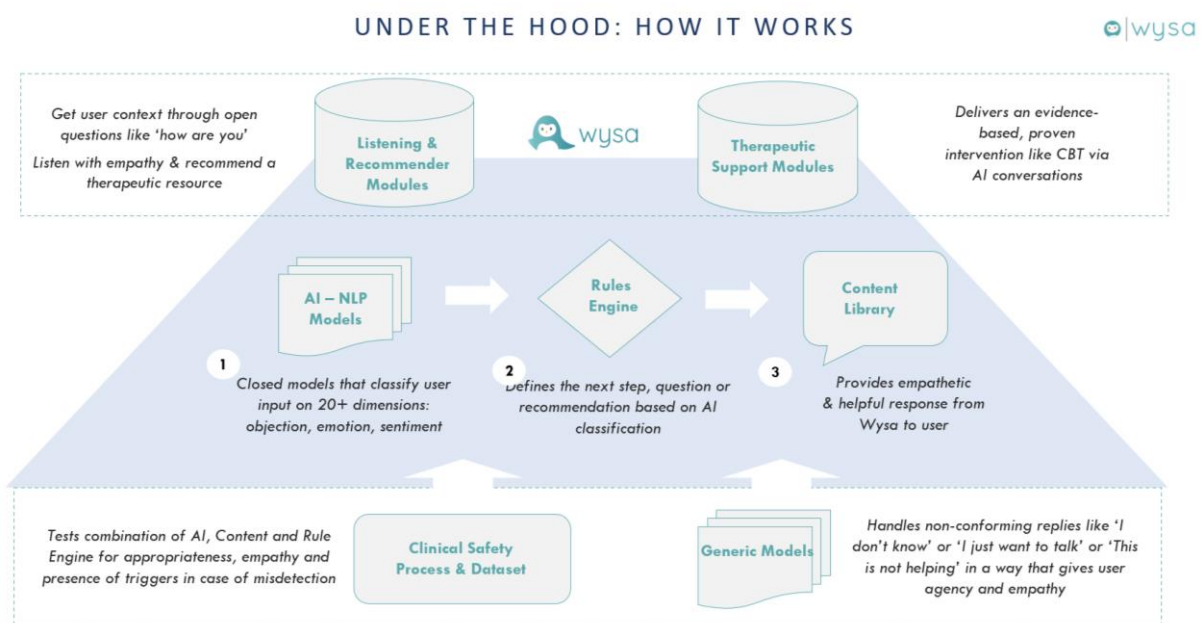


Figure 4-3. Diagram of Wysa's functionality [provided by Wysa Ltd.]

4.4.2.3 Case 3: Dora

Ufonia Limited developed a voice-based natural-language telephone conversational assistant (Dora R1) to deliver post-cataract surgery follow-up clinical consultations. Dora uses a variety of AI technologies to deliver the patient follow-up call, including speech transcription, natural language understanding, a machine learning conversation model, and speech generation to ask patients a series of (pre-determined) standard clinical symptom-based

questions [361]. Dora is delivered via a telephone call (no additional training or technology required for patients) as a real-time, stand-alone system: the operator inputs individual patient details to initiate the call and completes a summary in the electronic health record (EHR) afterwards. For this case study, the entire conversation was supervised live by a clinician who could interrupt the call at any point if the system failed, the patient struggled to interact with it, or Dora did not collect sufficient information from the patient. The detailed steps of the call are outlined in Figure 4-4.

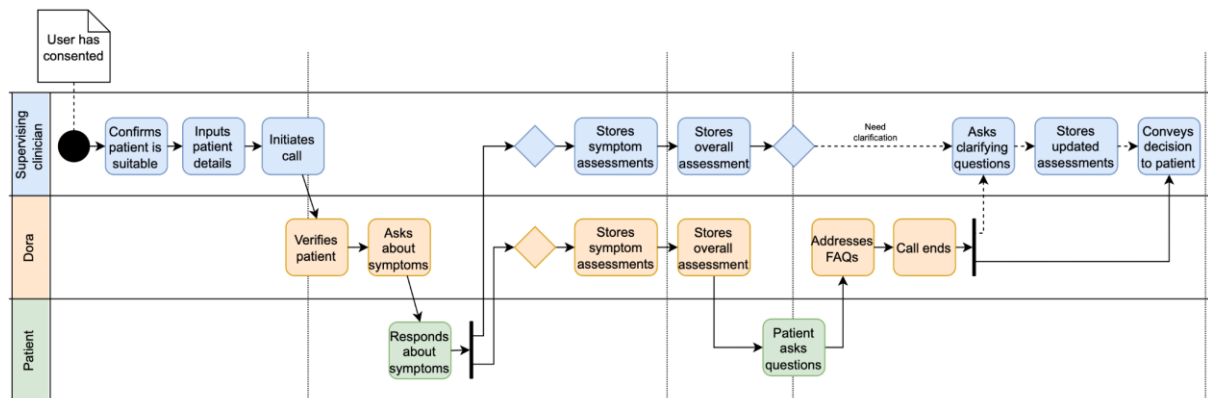


Figure 4-4. Flow diagram of Dora's call with patient

4.5 Individual Case Study Methods

The following sections present the full methods of the studies included in this thesis. Not all of the data described in this section are relevant for addressing the thesis research question. The full methods are provided here for completeness and to demonstrate the different aims and analyses of the case studies, as this will be relevant for their comparison in the cross-case analysis (Chapter 6). For clarity and conciseness, only the case study results pertaining to the research question are presented in the within-case analyses (Chapter 5). Table 4-3 provides an overview of the outcomes captured in each case study and which were included in this thesis; outcomes and measurement are described in more detail in the following sections.

Table 4-3. Case study outcome results reported in this thesis

Outcome	Case Study 1	Case Study 2	Case Study 3
<i>Included in Chapter 5 (within-case analyses)</i>			
Engagement	✓	✓	✓
Usability	✓	-	✓

Barriers and facilitators to engagement and perceived impact	✓	-	-
Motivation	✓	-	-
Self-efficacy	✓	-	-
Perceived impact on behaviour	✓	-	-
Opinions about app features	✓	-	-
Acceptability	-	✓	✓
Satisfaction	-	-	✓
<i>Excluded from Chapter 5 but included in published papers [151,345]</i>			
Depression severity	-	✓	-
Anxiety severity	-	✓	-
Crisis identification	-	✓	-
Health-related quality of life	-	✓	-
Accuracy and agreement	-	-	✓
Safety	-	-	✓
Cost impact	-	-	✓

4.5.1 Case Study 1: NoObesity

*This section was previously published in [77].

4.5.1.1 Aim

The purpose of this study was 1) to identify barriers and facilitators that may influence users' engagement with a mobile app for childhood obesity and 2) to evaluate its impact on perceived motivation, self-efficacy, weight-related health behaviours, and communication between HCPs and families [353].

4.5.1.2 Study design

A mixed-methods interventional study was conducted based on a Phase 1 implementation science design (Figure 4-5) [338,339]. Phase 1 studies focus on generating evidence around the implementation strategies and other factors that could influence the acceptability, effectiveness, and successful adoption [338,339]. This is a critical step in the evaluative process, as it informs intervention refinement to improve impact in later efficacy- and effectiveness-oriented evaluations. The primary analysis examined qualitative data about barriers and facilitators to users' engagement with the app and its impact on motivation, self-efficacy, and behaviours. To enhance the credibility of the qualitative results, they were triangulated with quantitative survey and app use data from a larger sample [44,353]. The RE-AIM and NASSS frameworks informed the data collection plan, to ensure a comprehensive set of individual- and system-level factors was captured ([Appendix 4-B](#)) [341–344]). The Standards for Reporting Qualitative Research (SRQR) [362] and Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) [363] checklists were used to ensure transparency and completeness ([Appendix 4-C](#) and [4-D](#)).

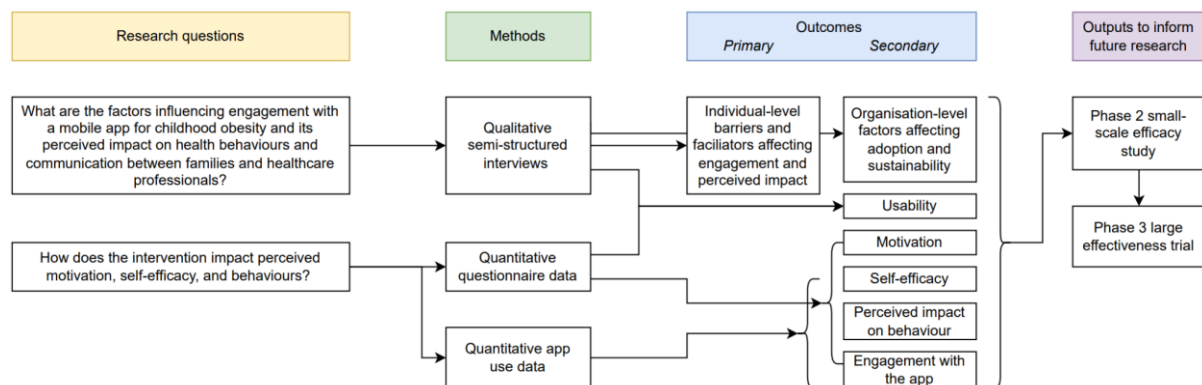


Figure 4-5. Case Study 1 logic diagram

4.5.1.3 Participants and recruitment

4.5.1.3.1 Recruitment and eligibility

From March to August 2020, a convenience sample of UK parents and HCPs was recruited using social media advertising (Google Adwords, Instagram Ads) [364]. Informed consent was collected using a Qualtrics survey. Participants could use the app as, when, and where they wanted. Those who completed all questionnaires received a £100 (USD\$139) Amazon voucher.

Participants self-selected to participate based on the following inclusion criteria: (1) a parent/guardian (with at least one child under 18) OR an HCP working with families, (2)

willing and able to use the app, (3) fluent in English, and (4) no previous use of the app. Children and vulnerable adults, individuals with hearing impairments, and individuals known to the research team were excluded to facilitate ethical approval and because of a lack of capability to manage these participants. Eligibility was not limited by weight to reflect real-world app use. Two employees from HEE who helped develop NoObesity were also recruited to capture organisational-level factors influencing adoption and sustainability ([Appendix 4-B](#)).

4.5.1.3.2 Sampling

Stratified random sampling (based on gender, ethnicity, and income) was used to select a sample for interviews from all participants who completed the demographic questionnaire and provided a valid email. To ensure that a representative sample of participants were interviewed, the 61 participants who had consented and completed the demographic survey were grouped based on gender, ethnicity, and household income. One participant was randomly selected using an online random number generator from each group that had at least one member; if there was only one participant in a particular demographic group, they were invited to interview. Only one respondent of the demographic survey was a healthcare professional working with children and was also invited for an interview; the rest were members of the public with children. In December 2020, these 20 participants were invited to take part in an interview to discuss their experiences using the NoObesity app. After 1 week and a reminder email, only 8 participants had signed up for an interview. Where possible, another participant was randomly selected from the unrepresented demographic groups and invited to participate. By January 2021, only 11 participants had agreed to an interview, so all of the remaining eligible participants who provided valid emails were invited to an interview. 15 interviews were conducted with members of the public with children between 26 January and 15 February 2021. Two HEE employees involved in the development of the NoObesity app were also interviewed to capture organisational-level factors influencing adoption and sustainability.

4.5.1.4 Data collection and outcomes

4.5.1.4.1 Data collection procedures

Ethical approval was obtained from the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (R62092/RE001) and the University of Plymouth's Faculty Research Ethics and Integrity Committee (19/20-1316). Thirty-minute SSIs were conducted on Microsoft Teams based on a topic guide ([Appendix 4-E](#)) and

manually transcribed by a third party company (Rev). To mitigate response bias, participants were informed that the interviewer was independent from HEE and had not been involved in intervention development. Three questionnaires were administered using Qualtrics: consent, demographics based on Oxford Internet Surveys (OxIS) 2019 [365] (September 2020), and a final questionnaire (April 2021), which examined usability, use, and impact on motivation, self-efficacy, and behaviour and included attention check questions ([Appendix 4-E](#)).

4.5.1.4.2 Outcomes

The primary outcome - barriers and facilitators to engagement and health behaviours - was explored through the qualitative SSIs. Secondary outcomes (Table 4-4) were examined through qualitative interviews, questionnaire, and app use data (including self-reported data and objective login data captured via the app admin portal).

Table 4-4. Case Study 1 quantitative secondary outcome measures

Outcome	Outcome measure
Engagement with the app	App-captured dates of first and last logins
	Quantitative self-reported app use duration
Engagement with the study	Uptake (number of participants recruited)
	Dropout rates (number of participants completing each stage of the data collection)
Usability	System Usability Scale (SUS, scale: 0-100) [299]; completed by parents twice - once for themselves and once from their children's perspective
Motivation	Quantitative ratings of the impact of specific features on motivation (5-point Likert scale, see Table 5-2 for details)
Self-efficacy	Bandura's self-efficacy scale [366] (confidence in ability to do goal activities; scale: 0-100)
	In-app self-reported confidence in ability to achieve goals (6 star rating scale)
Perceived impact on behaviour	Quantitative questionnaire items about physical activity and healthy eating behaviours (5-point Likert scale, see Table 5-2 for details)
	In-app self-reported goal progress data
Opinions about app features	Quantitative questionnaire items about app features (5-point Likert scale, see Table 5-2 for details)

4.5.1.5 Data analysis

Qualitative data was analysed using a codebook thematic analysis approach [367,368] in Dedoose (version 9.0.17) [369]. To improve credibility - the alignment of participants' feedback with our representation of their feedback - several trustworthiness criteria were followed, including prolonged engagement with data, peer debriefing and researcher triangulation, testing for accuracy, and triangulation with quantitative data [370]. Four authors engaged with the data (first independently, then collaboratively), which ensured that individual preconceptions were challenged and alternative interpretations considered [370]. An initial list of codes was generated inductively by one author based on the transcripts and provided to three independent authors to deductively code the data (generating additional codes if needed). Participants' statements were coded at the semantic level (interpreting only the explicit meaning of the participants' words) [371]. Codings then evolved through intensive collaboration between the authors. Themes and sub-themes were generated by one author as topic summaries [368] and then discussed and refined by all coders.

Once the thematic framework was agreed by all coders, the themes and subthemes were examined to identify factors influencing the key outcomes the study aimed to explore: engagement and perceived impact on motivation, self-efficacy, and behaviour. To do this, I determined, based on the codes, whether each subtheme related to engagement or perceived impact on one of the outcomes. The factors were based on the subthemes, but could include multiple subthemes that related to one factor and were then deductively mapped to the COM-B model to facilitate interpretation [53,170,372]. A codebook for the mapping was created using the COM-B components [170] and the Theoretical Domains Framework (TDF) [373], which captures more detailed subthemes of the COM-B components, and examples from the literature to guide coding in a digital health context ([Appendix 4-F](#)) [184,201]. These factors are the main outcomes discussed. Quantitative survey and app use data were analysed using descriptive statistics. Where possible, these outcomes were triangulated with qualitative data.

An exploratory examination of the influence of the different components of engagement on each other was conducted based on the factors generated from the thematic analysis; this data analysis is described in [Section 4.6.2](#) as it is common across the three case studies.

4.5.2 Case Study 2: Wysa

*This section has been submitted for publication in [129].

4.5.2.1 Aim

The primary aim of this study was to examine the efficacy of real-world use of an AI-enabled self-help CA app (Wysa) as a supplementary tool for patients waiting for mental health support through the Talking Therapies program (formerly Improving Access to Psychological Therapies (IAPT)). The primary hypothesis was that access to Wysa would improve symptoms of depression compared to a waitlist control; secondary hypotheses also predicted improvement for symptoms of anxiety and health-related quality of life. To provide a more in-depth understanding of real-world use, the study also qualitatively investigated engagement with and acceptability of Wysa (Figure 4-6).

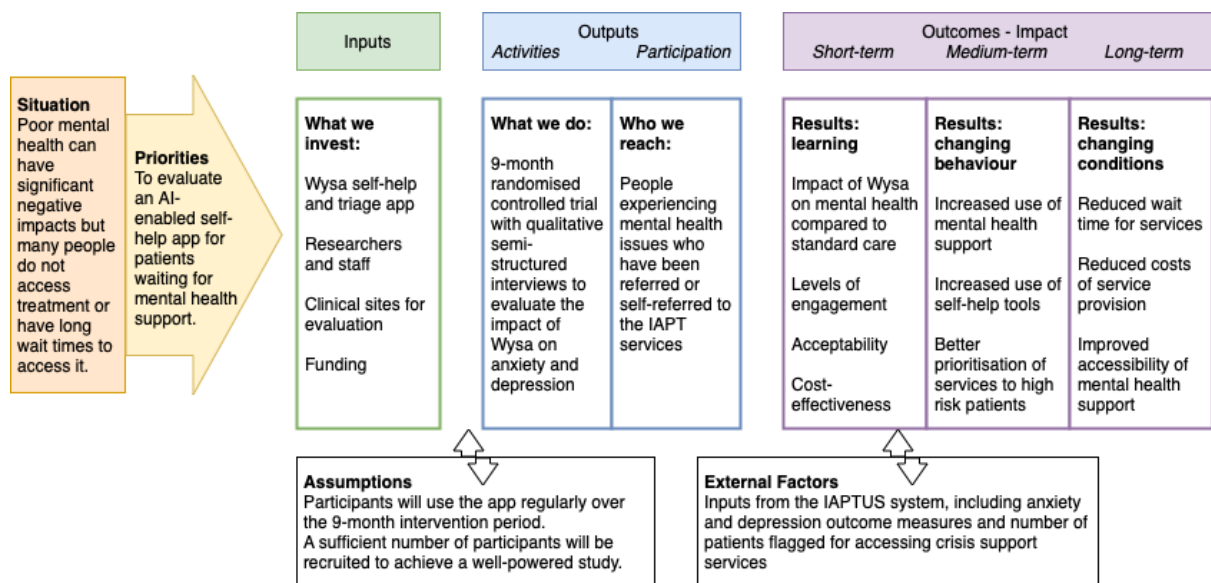


Figure 4-6. Case Study 2 Wisconsin logic diagram [374]

4.5.2.2 Study design

The overall evaluation of Wysa was a single-centre, parallel-group, randomised control trial (RCT) and used mixed methods to compare the intervention group against a waitlist control group. This comparator was selected to compare Wysa to the current standard of care for patients waiting for standard IAPT assessment and treatment. Participants were allocated 2:1 (Wysa:control) and followed up over 12 weeks. Participants in both groups received standard care, but the intervention group also received access to the Wysa AI CA app.

Quantitative methods were used to evaluate Wysa's efficacy at improving mental health outcomes. Quantitative and qualitative methods, specifically SSIs, were used to gather a more in-depth understanding of users' engagement with Wysa and their feedback regarding its acceptability (Figure 4-7). Two theoretical frameworks were used to shape the topic guide for the SSI, to ensure that data was being captured to comprehensively examine the different

components of engagement and factors influencing them before, during, and after engagement with Wysa: the TFA [320] and a multi-faceted conceptualisation of engagement [14,42,45,49–51].

Following the Consolidated Standards of Reporting Trials AI extension (CONSORT-AI) checklist ([Appendix 4-G](#)) [375], the study was prospectively approved on 25 May 2022 by the London - Stanmore Research Ethics Committee (22/PR/0467).

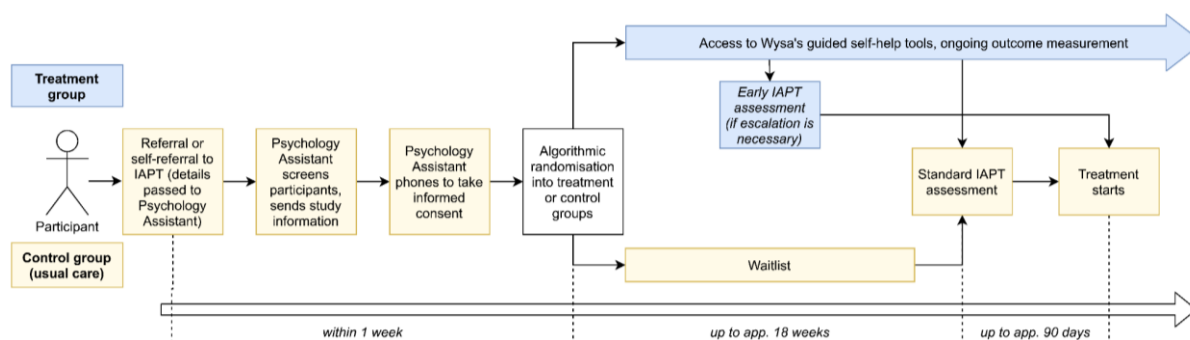


Figure 4-7. Case Study 2 study flow diagram

4.5.2.3 Participants and recruitment

4.5.2.3.1 Recruitment and eligibility

Participants referred to three Talking Therapy services within the Central North West London (CNWL) NHS Foundation Trust and were screened for eligibility by a psychology assistant. Participants on the standard care pathway who were 18 or older, proficient in English, willing and able to provide informed consent, referred or self-referred to IAPT with mild to moderate mental health concerns, and owned a device capable of supporting Wysa were eligible. Exclusion criteria included: a) previous and current major mental illness; b) current psychosis or a history of psychotic symptoms within 6 months; c) suicidal ideation; d) significant cognitive disorders; e) noted neurodevelopmental conditions; f) personality disorder diagnosis; g) PHQ-9 or GAD-7 score >15 (indicating moderately severe or severe symptoms); h) under the care of specialised mental health services in the last 2 years; i) previously failed Talking Therapies; j) referrals for a specialised assessment beyond the standard clinical pathway; k) incapable of self-consent; or l) in a dependent/unequal relationship with any project team member. A psychology assistant contacted eligible participants by telephone and email via an electronic data capture system (Castor EDC, Castor Research Inc.).

There were delays in Health Research Authority (HRA) approval, MHRA approval, and site set up. This meant that recruitment took place between 13 December 2022 and 27 April 2023,

lasting 4.5 months rather than the planned 7 months, as the funder end date was 30 August 2023. The protocol was registered on ClinicalTrials.gov (NCT05533190) on 05 September 2022 and ISRCTN (14644939) on 7 December 2022.

4.5.2.3.2 Sampling and randomisation

Preliminary work suggested that the use of Wysa was associated with a mean clinical improvement in PHQ-9 of 5.84 (SD6.66) [358]. Based on a standard deviation of 6.6 and a 2:1 recruitment ratio (usual care: intervention) a sample size of 393 would allow a clinically meaningful effect size (a difference of at least 2 points) on the PHQ-9 scale to be detected with 80% power [358]. This sample size was increased to 480 to allow for 20% attrition. This sample size was to be achieved over a 7-month recruitment period from an estimated pool of 1100 patients referred to CNWL monthly (service target).

All participants who consented were randomised into the intervention group (Wysa + standard care) or control group (standard care) in a 2:1 ratio after they completed the baseline questionnaire. A masked computer randomisation algorithm (Sealed Envelope [376]) ensured that the groups were balanced on key variables (limited to three - gender, age, and iaptus triage PHQ-9 score - due to the strata equation). Randomisation was done by the psychologist assistant and recorded encrypted in Castor to blind the evaluation teams; the qualitative team was un-blinded at the point of scheduling qualitative interviews with patients in the intervention group. Participants were sent an automatic email to inform them of their group assignment and participants in the intervention arm received an SMS link to download the premium (full) version of Wysa on their Android or iOS devices.

20 participants were selected using stratified random sampling to be invited to participate in a SSI to share their experiences and thoughts about Wysa. To ensure a diverse sample was collected, the primary sampling was stratified by gender, ethnicity, and age. Deprivation levels (above or below median)) and starting PHQ-9 levels (mild or moderate) were also reviewed to ensure there was sufficient diversity; this did not require any adjustments from the randomly generated lists. A backup list was generated by matching primary participants with remaining participants who most closely matched their demographic data.

4.5.2.4 Data collection and outcomes

4.5.2.4.1 Data collection procedures

The timeline for the data collection is detailed in Figure 4-8.

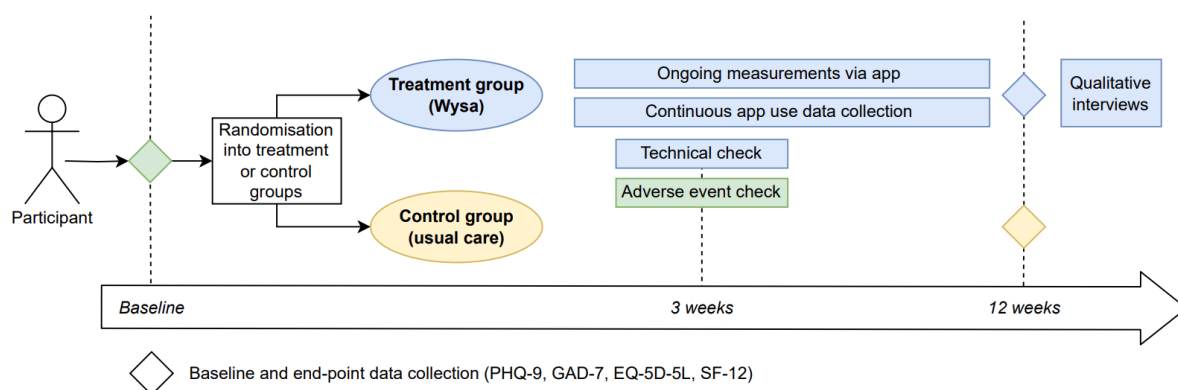


Figure 4-8. Case Study 2 data collection sequence diagram

The Castor EDC system was used to store all patient research data collected. Pre-randomisation, patients completed a baseline questionnaire and demographic data was collected from the patients' EHR on iaptus. Three weeks after a patient was randomised, the psychology assistant called them to identify any technical issues or adverse events. Twelve weeks post-randomisation, patients were sent a final questionnaire, which captured the same 4 outcomes collected at baseline. Participants were randomly selected for SSIs in June 2023 and invited for interviews as they completed their 12-week intervention periods. Interviews were conducted based on a topic guide ([Appendix 4-H](#)) via Microsoft Teams with one researcher between 22 June and 25 July, 2023. Participants were asked if they consented to audio and video recordings; if not, notes would be taken by hand and shared with the participant following the interview for verification. At the end of the study period, app use data from all intervention group participants were downloaded from Wysa servers. Demographic data and PHQ-9 and GAD-7 scores were collected from patients' EPRs via iaptus.

4.5.2.4.2 Outcomes

The primary outcome was change in depression severity from baseline to 12 weeks between groups, measured using the PHQ-9 [354] and compared between users of the app and patients in the standard clinical pathway (Table 4-5). Depression severity was chosen as the primary outcome because the PHQ-9, a depression symptom measure, is recommended for use in all patients presenting to IAPT services [136].

Secondary outcomes, including anxiety severity (GAD-7 [355], in line with the national recommendations laid out by the IAPT Manual [136]) and health related quality of life (EQ-5D-5L [377]) were also measured at baseline and 12-weeks post randomisation. The EQ-5D-5L [377] provides a descriptive self-assessment of health in 5 dimensions - mobility, self-care, usual activities, pain/discomfort and anxiety/depression [377] - and is commonly used as

a quality of life measure to conduct cost effectiveness analyses [378]. Wysa’s ability to identify users who are at high risk of a crisis and who need additional support was measured by capturing the number of patients that Wysa flags for escalation and the number of patients that flagged themselves as needing further support.

User engagement with Wysa was assessed using a combination of complementary measures. App use data collected included a tally of exercises completed, sessions completed, and messages to Wysa. Interviews were structured based on the Theoretical Framework of Acceptability (TFA) [320] and a multifaceted conceptualisation of engagement [14,42,45,49–51] ([Appendix 4-H](#)).

Table 4-5. Case Study 2 outcome measurement

Outcome	Outcome Measure
Depression severity	Score on the PHQ-9 [354]
Anxiety severity	Score on the GAD-7 [355]
Health-related quality of life	EQ-5D-5L [377]
Crisis identification	Number of users identified by the app for escalation of care
Engagement	- Uptake rates - Dropout / ‘did not attend’ rates - App usage data - Qualitative feedback from SSIs
Acceptability	Qualitative feedback from SSIs

4.5.2.5 Data analysis

Descriptive statistics assessed baseline comparability and the distributions of the primary and secondary outcomes. Linear regression was used to carry out comparative analysis of PHQ-9 and GAD-7. As the PHQ-9 and GAD-7 scores were captured from multiple sources (study questionnaires and iaptus triage scores), a within-item average was taken for patients whose scores on these two data points differed. Because of the small sample size, the study statistician determined that the planned imputation of missing data was not advisable; missing values were excluded in descriptive analyses. Additional post hoc regression was carried out for individuals who had a PHQ-2 score ≤ 5 and ≤ 4 at baseline. Health-related quality of life was obtained via the ‘eq5d’ package in R: this package transforms ordinal survey responses into a single index. Risk events (user-triggered or flagged by Wysa) were counted.

Qualitative data was analysed using a codebook thematic analysis approach [367,368] in ATLAS.ti (version 23.2.1). The seven constructs of the TFA [320] and the five main components of a multifaceted theoretical conceptualisation of engagement (affective, cognitive, behavioural, micro, and macro) [14,42,45,49–51] comprised the initial coding framework ([Appendix 4-I](#)). Using this framework, one author inductively generated sub-codes. The TFA constructs were theoretically associated with the components of engagement to facilitate interpretation [48,320,335]. Categorical theming of the data was conducted to organise the patterns constructed through the coding process [379]. Themes could relate to multiple constructs or components. A second author independently coded the transcripts, adding additional codes if needed. Independently-developed thematic frameworks were compared and synthesised collaboratively to generate the final interpretation and thematic map. Qualitative and quantitative measures of engagement were triangulated.

4.5.3 Case Study 3: Dora

*This section was previously published in [151].

4.5.3.1 Aim

This study aimed to build on preliminary evidence of acceptability [380,381] and evaluate the accuracy, safety, and potential for adoption into routine clinical care of the first version (R1) of Dora to gain CE mark approval. The primary objective was to assess Dora's ability to accurately detect whether a patient requires further assessment by a human clinician. Other factors relating to Dora's adoption, including patients' perceptions and implementation costs, were also assessed to determine the strength of evidence for broader implementation.

4.5.3.2 Study design

Evidence of impact is no guarantee of successful adoption and implementation of a healthcare intervention [338]. Implementation science emphasises the importance of gathering evidence not just of effectiveness but of contextual factors that will influence real-world use [339]. A mixed-methods implementation science approach [339] was used to evaluate Dora's potential impact as a postoperative assessment of cataract surgery patients (Figure 4-9). Research ethics approval was obtained from the Health Research Authority (21/PR/0767) and the University of Plymouth's Faculty Research Ethics and Integrity Committee (2863). The trial was registered on ClinicalTrials.gov (NCT05213390) and ISRCTN (16038063). The Developmental and Exploratory Clinical Investigations of DEcision support systems driven by Artificial Intelligence (DECIDE-AI) [382] and SRQR [362] reporting guidelines were

used to ensure that the mixed-methods study was reported transparently and comprehensively ([Appendix 4-J](#) and [4-K](#)).

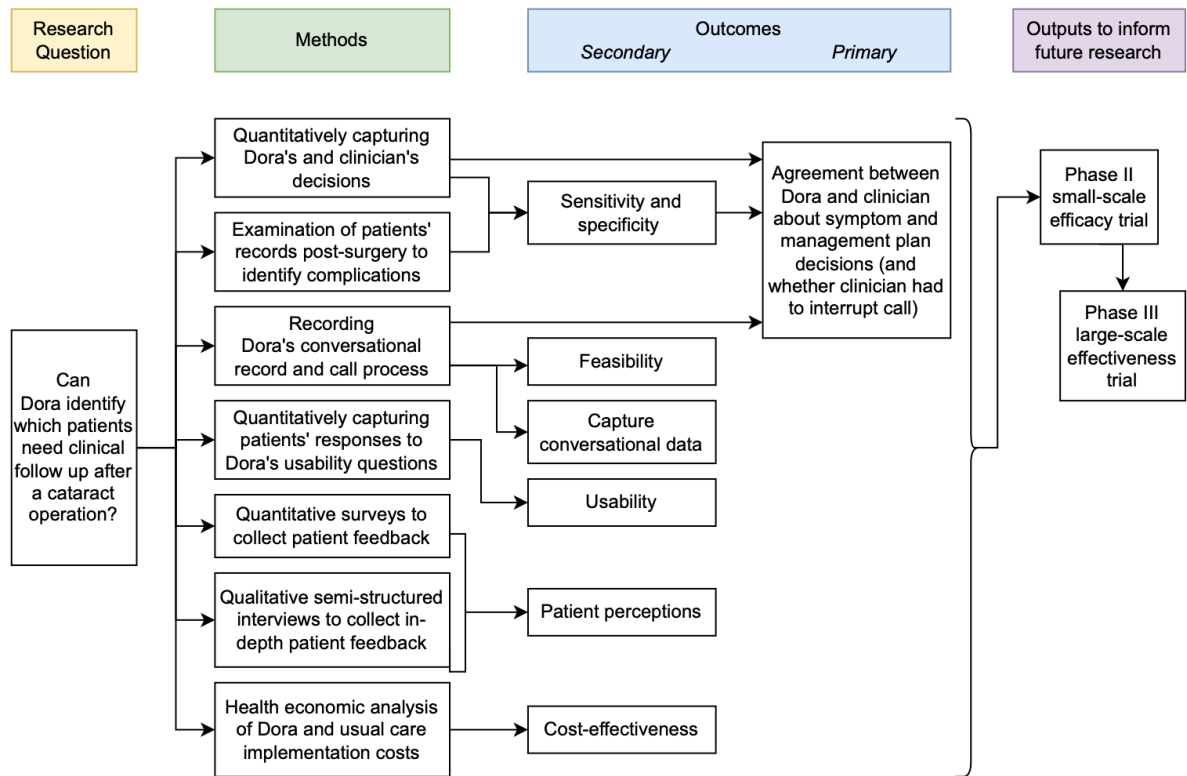


Figure 4-9. Case Study 3 logic diagram

4.5.3.3 Participants and recruitment

4.5.3.3.1 Recruitment and eligibility

Patients undergoing cataract surgery were recruited over five months (09.2021-01.2022) from two clinical sites: the Imperial College Healthcare NHS Trust (Imperial) and the Oxford University Hospitals NHS Trust (Oxford). These two sites represent different populations - ethnically diverse and urban at Imperial and primarily white British and more rural at Oxford. Patients scheduled for cataract surgery who met the eligibility criteria (Table 4-6) were invited to participate in and consented into the study by telephone by research nurses or in pre-assessment sites and post-operative discharge lounges. There were two stages of screening during enrolment: the first stage screened clinical notes to determine potential eligibility and the second involved contacting potentially eligible patients by phone to confirm eligibility and consent them for the study.

Table 4-6. Case Study 3 inclusion and exclusion criteria

Inclusion criteria	
1	Aged 18 years or older.
2	Subjects willing and able to provide consent
3	On the waiting list for routine cataract surgery. Cataract surgery as part of a combined procedure with other ocular surgery will not be included.
4	No history or presence of significant ocular comorbidities that would be expected to alter the risks of cataract surgery or normal post-operative follow-up schedule. Note that significant ocular comorbidities do not include stable, chronic, or inactive ocular conditions such as amblyopia, drop-controlled stable glaucoma or ocular hypertension, previous squint surgery, inactive macular pathology, previous refractive surgery, or previous vitreoretinal surgery with stable retina.
Exclusion criteria	
1	Individuals with any condition that could preclude the ability to comply with the study or follow-up procedures.
2	Presence of ocular or systemic uncontrolled disease (unless deemed not clinically significant by the Investigator and Sponsor).
3	History of current or severe, unstable or uncontrolled systemic disease (unless deemed not clinically significant by the Investigator and Sponsor).
4	Participants who are involved in current research or have been involved in any research within 2 months from enrolment.
5	Participants with cognitive difficulties, hearing impairment or non-English speakers

4.5.2.3.2 Sampling

Other members of the study team conducted an *a priori* power calculation to determine the minimum sample size necessary to achieve a power greater than 0.8 for a 0.05 level of significance for the primary outcome [383]. On the basis of a total population of 1014, a 95% confidence level, and a 3% margin of error, a sample size of 520 was determined to be required for statistical significance. An estimate of the potential pool of participants was generated by conducting an audit of pre-assessment appointment data at both sites. Based on historical and current data at the time (and including estimates of 70% of eligible patients consenting and 15% attrition), the expected number of recruited participants over the planned 26 week recruitment period was 591.

Stratified random sampling based on demographic data (gender, ethnicity, and income) was conducted to select a diverse sample of participants for interviews. Primary and matched

backup lists were created by randomly selecting a white male and female participant from each income bracket (n=12) and 8 patients of non-white ethnicity (there were insufficient participants to select from each income bracket).

4.5.3.4 Data collection and outcomes

4.5.3.4.1 Data collection procedures

All eligible patients received a postoperative call from Dora approximately 3-4 weeks post-surgery in addition to their default care pathway. The standard care pathway at Oxford is that patients without significant ocular comorbidities and have routine surgery are discharged on the day of surgery and advised to attend their community optometrist after 6 weeks (Figure 4-10). At Imperial, patients typically received face-to-face follow up at 4 weeks after surgery. The 3-4 week time point for the Dora call was chosen as this is prior to the standard 4-week follow-up at Imperial.

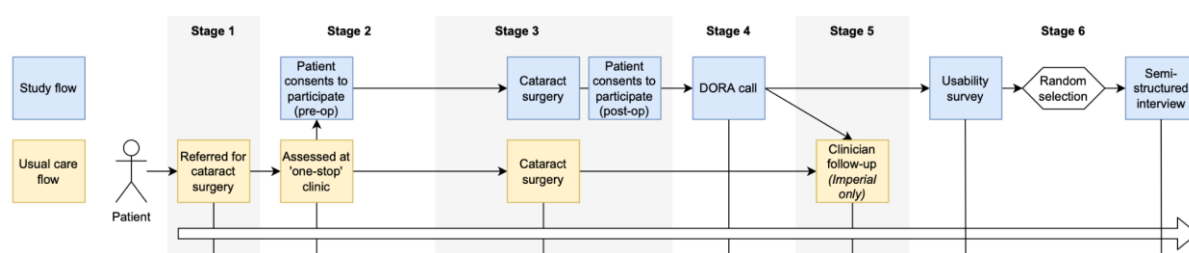


Figure 4-10. Case Study 3 study flow diagram [Note: patients at Imperial all have planned routine face-to-face reviews. At Oxford the majority of patients are discharged on the day of their surgery, although some patients have planned clinic review (for example, for existing ocular comorbidities)]

Dora calls were supervised in real-time by one of two ophthalmologists who could take over if needed. Dora and the supervising ophthalmologists independently made decisions about the clinical significance of five symptoms and the overall call outcome (either ‘recommend discharge’ or ‘recommend further review’; [Appendices 4-L](#), [4-M](#), and [4-N](#)). ‘Recommend discharge’ refers to discharge from hospital-based ophthalmology services; typically, all patients are advised to see their community optometrist for a sight test at 6 weeks. All patients recommended for further review by Dora received an immediate call-back by a clinician, which mirrors the real-world operational model.

The software automatically recorded Dora’s and the supervising clinicians’ decisions. Clinical notes were reviewed up to 3 months postoperatively. Demographic and usability data were collected via telephone or online Qualtrics surveys ([Appendix 4-O](#)) and through patients’

EHR (October 2021-March 2022). A subset of participants were invited to qualitative SSIs using stratified random sampling (February-March 2022). [Appendix 4-P](#) outlines the SSI topic guide, which was developed based on the Theoretical Framework of Acceptability (TFA) [320].

4.5.3.4.2 Outcomes

The primary outcome of the study was to establish preliminary safety evidence by evaluating the level of agreement between Dora and the expert human clinician on symptom and overall management plan decisions (Table 4-7). To evaluate agreement, Dora's ability to classify five key symptoms from the conversation was compared with a supervising clinician's classification of the same five symptoms: redness, pain, reduced vision, flashing lights, and floaters. For this study, the supervising clinician's decision - made based on the information from the call with Dora's decision masked - was considered the gold-standard for evaluation. The agreement between Dora and the clinician on the patient's management plan was evaluated in the same way as the symptoms.

Secondary outcomes included clinical safety, user perspectives, and cost impact. To examine safety, clinical notes were reviewed at three months to identify patients recommended discharge by Dora who subsequently had either unexpected management changes (UMC) or unplanned review. Performance of the call was assessed by examining the proportion of calls completed autonomously by Dora. Post-survey questionnaires, delivered via online or telephone surveys collected demographic (OxIS 2019 Questionnaire [365]) and usability data (System Usability Scale (SUS) [384] and Telehealth Usability Questionnaire (TUQ) [385]; [Appendix 4-O](#)). Users' perceptions of engagement with Dora - acceptability, satisfactoriness, and appropriateness - were assessed during qualitative interviews with a subset of the patients to capture more in-depth feedback about Dora ([Appendix 4-P](#)). Patients' willingness to recommend Dora to others was assessed at the end of the call using the Net Promoter Score (NPS): "on a scale of 1 to 10, how likely would you be to recommend this automated service to a friend or colleague." Cost-effectiveness was examined by collecting and analysing health economic data before and during the study.

Table 4-7. Case Study 3 outcome measurement

Outcome	Measure	Method type	Data collection
Accuracy and agreement	Kappa statistic, sensitivity, specificity, accuracy	Quantitative	Dora and clinician decision data recorded automatically by system
Safety	Presence of unexpected management change at planned follow up (i.e. addition of eye drop other than artificial tears, procedure other than suture removal, urgent or emergent referral) Presence of unexpected AnE visits within a 3 month period	Quantitative	Case note review, patients seen at Imperial face to face as part of the current pathway, 3-month post-operative audit of respective electronic health records blind to Dora's outcome
Satisfaction	Net Promoter Score (NPS) [386]	Quantitative	NPS question asked by Dora at the end of the call
Usability	System Usability Scale (SUS) [384] Telehealth Usability Questionnaire (TUQ) [385]	Quantitative	Questionnaires delivered by email or telephone by a researcher
Acceptability	Thematic analysis of SSIs	Qualitative	SSIs conducted by telephone by a researcher
Cost impact	Real-world economic analysis	Quantitative	Number of patients, follow-up clinician time, average NHS nurse salaries, Dora call costs

4.5.3.5 Data analysis

Quantitative data was analysed using Stata (Release 17) [387]. The primary analysis calculated inter-rater reliability (Cohen's kappa) between Dora and the clinician, accuracy, sensitivity, specificity, and area under the curve for the overall outcome and each symptom decision. Both Cohen's kappa and sensitivity/specificity metrics were included to provide a comprehensive evaluation of the performance of the CA compared to the supervising ophthalmologist. While sensitivity and specificity measure the accuracy of correct positive and negative classifications, respectively, the kappa statistic accounts for the agreement between observed and expected classifications, correcting for chance agreement. By incorporating both types of metrics, the analysis aims to assess not only the model's overall

accuracy but also its agreement beyond what could be expected by random chance, providing a more nuanced and robust evaluation of its performance.

Secondary analyses included calculating the same metrics comparing Dora's overall decision with unplanned hospital review and unexpected management changes, using descriptive statistics to examine feasibility and usability (SUS [384], TUQ [385], and NPS [386]), and conducting exploratory analyses to identify any associations between demographic characteristics and usability. Missing data was excluded from statistical analyses.

A codebook approach [367,368] was used to conduct a thematic analysis of the interview transcripts in QSR NVivo 12 [388]. The structure of the codebook was set based on the seven main constructs of the TFA (affective attitude, burden, perceived effectiveness, ethicality, intervention coherence, opportunity costs, and self-efficacy) [320]. As discussed in [Section 3.5](#), the constructs of the TFA can be mapped to different components of engagement: affective attitudes and ethicality largely relate to affective engagement, while burden, intervention coherence, opportunity costs, perceived effectiveness, and self-efficacy are more closely aligned with cognitive engagement (although this is not a clear division, there can be overlap among the factors). Within this structure, one author inductively generated an initial list of codes based on a semantic-level examination of the transcripts [371]. A second author independently coded the transcripts using this codebook, adding additional codes where needed. Codings evolved through intensive collaboration between the authors to challenge individual preconceptions and ensure alternative interpretations were considered. Both authors independently developed thematic maps [379], which were compared, discussed, and consolidated.

A cost analysis compared the direct costs of face-to-face (F2F) follow-up at Imperial with Dora (in Oxford, patients do not routinely have postoperative follow up booked). Assumptions included the annual costs for various healthcare professionals [389], the length of F2F follow-up appointments (estimated at 30 minutes).

4.6 Case Study Analytic Approaches

4.6.1 Selection of case study analytic approaches

As described in [Section 4.5](#), each case study had distinct main aims, with methods and outcomes designed to meet those particular aims: understanding barriers and facilitators, evaluating clinical effectiveness, and assessing accuracy and safety. These main aims were aligned to a greater or lesser extent with the thesis research question of how the different

components of engagement were related, but all of the case studies examined engagement with the intervention, enabling this question to be explored through secondary analyses of the data. Engagement in all three case studies was primarily assessed through thematic analysis of the qualitative interview data, triangulated with quantitative questionnaire and system use data, to explore how the different components of engagement were related. To address the research question, additional analytic techniques were needed to generate and synthesise mappings of the interactions between components over the process of engagement.

There are five main analytic techniques for analysing case study data: “pattern matching, explanation building, time-series analysis, logic models, and cross-case synthesis” [337]. Each of these techniques were considered in the process of determining how the case study analysis would be conducted to address the research question. The pattern matching technique “compares an empirically based pattern... with a predicted one” [337]. This basic process underlies most of the remaining analytic techniques, which are extensions or specific cases of pattern matching. For example, the explanation building (or hypothesis-generating) technique, builds on the pattern matching comparison to refine the proposed pattern and further test it against new data or other cases, an iterative process using theory and data to generate an explanation (or hypothesis) [337]. The time-series analysis technique also uses pattern matching, but over time, which may include different conditions or characteristics that influence the trend in question and the logic model technique maps out how particular events influence others in a complex web of interactions [337]. These first four techniques can be applied to individual case studies, while the last - cross-case synthesis - is restricted to multiple case study approaches [337]. This technique differs from other types of synthesis in its focus on a “case-based” analysis - it emphasises the importance of examining each case as a whole, including its context, rather than isolating and examining specific variables.

After considering these different analytic techniques, the pattern matching, explanation building, and cross-case synthesis techniques were determined to be most useful to address the research question. Each case was individually analysed to develop a process map based on the empirically-observed patterns of interaction between engagement components ([Section 4.6.2](#)). In this way, there were also elements of the logic diagram technique included in the analysis in the way that the data was visualised. The time-series technique was not relevant for this particular analysis.

4.6.2 Thesis-specific within-case analysis

Although each case study examined engagement with the DHI, engagement was not the primary outcome for all of them (Figure 4-1, Table 4-3). Each case study had relevant data for addressing the thesis research question and aim of exploring the associations between the different components of engagement and developing recommendations for how those could be supported, but this analysis was not conducted for the purposes of the main funded studies. To enable the cross-case synthesis to achieve the aims of the thesis, additional analyses needed to be conducted on data from each of the individual case studies to generate mappings that could be compared using pattern matching and cross-case synthesis techniques.

The thematic frameworks developed for each case study were re-evaluated to examine the codes and themes through the lens of the multifaceted conceptualisation of engagement. The definitions of the components of engagement generated from the literature in [Section 3.6.3](#) were used to guide the mapping of components of engagement. The three thematic frameworks variously listed codes, themes, and theoretical factors derived during the initial thematic analyses; these summarised the key themes but did not always provide sufficient context for determining which engagement component(s) the themes related to. To address this limitation, I re-engaged with the transcript data and extracted quotes that represented the generated themes. Together with the thematic frameworks, the quotes provided further information to shape my judgments about the components of engagement involved in each theme and how they related to each other.

Based on participants' descriptions of their engagement and the factors that affected it, patterns of how individual components of engagement influenced each other were developed. For example, if a participant described feeling anxious about using an intervention because they were not confident that their data would be safe, the pattern of engagement coded was cognitive → affective → behavioural. This is because cognitive engagement includes users' assessments of risk and benefits of the app - in this case, they are judging that there is a data security risk, which is causing them to feel anxious about intervention use (affective engagement) and could reduce the likelihood that they will behaviourally engage with the intervention. These patterns of engagement were based primarily on the data from the interviews with reference to the theme summaries and linked theoretical concepts.

Unlike the original case study thematic analyses, these thesis-specific within-case mapping analyses were conducted by one researcher, creating potential for bias. Efforts were made to mitigate this by using the same definitions of the engagement components ([Section 3.6.3](#)) to

structure the analysis and by conducting the analysis twice, separated by several months, to enable me to challenge my perceptions. Limitations are discussed in more detail in [Sections 5.2.6, 5.3.6, 5.4.5, and 7.3.6.1](#).

4.6.3 Cross-case analysis

The initial step in the cross-case synthesis was to examine the comparability of the cases, to determine whether it is valid to synthesise the cases and generate insights based on that analysis. Then the empirically-based models were compared to the theoretically-based predicted model ([Section 3.8](#)) and to each other using the pattern matching technique [337]. The cross-case analysis technique was used to synthesise the findings from the individual cases, identify similarities and differences, and determine if there is literal replication (between the similar cases 1 and 2) and theoretical replication (between cases 1 and 2 and the different case 3) [337,390,391]. The explanation building technique was then applied to aggregate the findings and develop a new hypothesised model for the pattern of relationships between the components of engagement. As with the thesis-specific reanalysis of the case studies, these cross-case analyses were only conducted by one researcher.

In addition to the pattern matching, explanation building, and cross-case synthesis techniques, a thematic synthesis was also conducted as part of the cross-case analysis [337,392,393]. Thematic synthesis methods commonly involve a three-stage process: 1) a line-by-line coding of primary studies, 2) categorisation of codes into descriptive themes, and 3) the development of analytical themes that generate additional insights beyond the primary studies [394,395]. A thematic synthesis was included in the analysis because the main data from each of the case studies was qualitative. Extending the analysis beyond the pattern matching of the generated models helped to ensure that the cross-case synthesis took a case-based approach that examined the contextual factors and how they all interacted to shape the patterns observed. It also had the potential to generate additional insights about factors that might influence the associations between particular components of engagement at particular times, which provided an additional contribution by generating hypotheses for future investigation. Together, the thematic synthesis and explanation-building techniques were used to iteratively develop a narrative around how and why users engaged with the DHIs and target behaviours and the roles of the components of engagement with DHIs (including what can be generalised and what may be specific to particular cases) [337,392,393].

The initial step of a typical thematic synthesis (line-by-line re-coding of the studies) was skipped because I conducted the initial coding and thematic analysis of the 3 case studies.

Although the codebooks used in the analyses were different, I had an intimate knowledge of the transcripts, codes, and themes from each study, unlike a thematic synthesis that aims to bring together qualitative data from studies previously published by different researchers. Therefore, the first step in my thematic synthesis was to analyse the original codings and themes to determine descriptive themes. The descriptive themes organised similar codes and themes from across the three case studies, based on the original codes without additional interpretation [394,395]. In the process of conducting this step it became clear that, as the case studies had different overall aims and research questions, not all codes and themes from the individual case studies were relevant for the thesis thematic synthesis. For the purpose of comparability, only codes and themes from each case study relating to factors influencing engagement with the DHI were included. For case study 1, this meant that themes relating to engagement with health behaviours were excluded from the synthesis; for case study 2, the organising theme ‘perceptions of potential usefulness in clinical pathways’ was excluded. Purely descriptive codes and themes (e.g. participants found the app engaging / not engaging) were excluded from all case studies. The descriptive themes were then interpreted to identify analytical themes that related to factors influencing engagement across the three case studies [394,395].

A proposed theoretical model of how engagement components interact over the process of engagement with a DHI was developed based on synthesis of the hypothesised and 3 empirical models and alternative explanations were generated and explored in the context of the data [337,396]. The refined conceptual model, positing how the different components of engagement influence each other and what factors can mediate these relationships, was combined with previous evidence from the literature to recommend strategies for supporting engagement with DHIs for future empirical evaluation.

Chapter 5: Within-Case Analyses

5.1 Chapter Overview

Chapter 4 presented the theoretical and methodological approach of the thesis. A multiple case study design was selected as the most appropriate method of addressing the research question and aims. Three cases were included to enable an exploration of how the process of engagement may be similar or different depending on the digital health intervention (DHI) and to allow for the development of a model positing how different components of engagement influence each other across DHIs more broadly. The three cases - NoObesity, Wysa, and Dora - were introduced and the full methods for each were presented. The chapter concluded with a description of the additional within-case analyses that were conducted to address the thesis research question and how the cross-case synthesis would bring these case studies together to generate additional insights.

The findings from these analyses have been divided into the next two chapters. Chapter 5 will present the individual case study findings, including the original case study thematic analyses (addressing specific case study objectives) and the re-analysis of the case study data for this thesis (to develop mappings of the associations between engagement components). Chapter 6 will describe the results of the cross-case analysis and thematic synthesis, which address the main thesis research question: *How are the different components of engagement (affective, cognitive, behavioural, micro, and macro) related in DHIs?*

The main sections of Chapter 5 include the results of each thematic analysis and other relevant intervention use, acceptability, or usability data (included in the published papers) and re-evaluations of those data conducted specifically to address the thesis research question. Each case study had distinct research questions and aims (Figure 4-1, [Section 4.2.2](#)); the full methods for each evaluation were presented in Chapter 4 to provide an overall understanding of each case study, but this chapter will only present the results that are relevant to the thesis. The full results for each case study have been published [77,345] or submitted for publication [129] elsewhere. Mappings of associations between the different components of engagement as indicated by the evidence were generated for each case study and will be synthesised in Chapter 6. Each section concludes with a discussion of the findings, strengths and limitations of the case study, and implications for future development and research. The final section of the chapter provides a brief summary of the key findings from each case study and sets the scene for the cross-case synthesis in Chapter 6.

Publications: Sections of Chapter 5 have been drawn from previously published and submitted papers; references and author contributions for each of the papers that were used in the preparation of this chapter are listed below.

Milne-Ives M, Rahman E, Bradwell H, Baines R, Boey T, Potter A, Lawrence W, van Velthoven MH, Meinert E. Barriers and facilitators to parents' engagement with and perceived impact of a childhood obesity app: A mixed-methods study. *PLOS Digital Health*. 2024 Mar 27; 3(3):e0000481. doi:[10.1371/journal.pdig.0000481](https://doi.org/10.1371/journal.pdig.0000481)

Authors' contributions: MMI collected data, carried out the initial analyses, and drafted the initial manuscript, and critically reviewed and revised the manuscript. ER, AP, WL, and MvV conceptualised and designed the study and critically reviewed and revised the manuscript. HB, RB, and TB, analysed data and critically reviewed and revised the manuscript. EM conceptualised and designed the study, supervised data collection and analysis, and critically reviewed and revised the manuscript.

Meinert E, **Milne-Ives M**, Lim E, Higham A, Boege S, de Pennington N, Bajre M, Mole, G, Normando E, Xue K. Accuracy and safety of an autonomous artificial intelligence clinical assistant conducting telemedicine follow-up assessment for cataract surgery. *eClinical Medicine*. 2024 Jul 4; 73:102692. doi:[10.1016/j.eclinm.2024.102692](https://doi.org/10.1016/j.eclinm.2024.102692)

Authors' contributions: NdP conceived of the study topic. The study design was developed by NdP, GM, EL, MMI, EN, KX, EM and AH. KX and EN led patient recruitment at the two study sites. AH and EL supervised Dora R1's calls and compiled the clinical data. MMI conducted the usability survey and MMI and EM completed participant interviews. Statistical analysis of the quantitative data was conducted by the University of Plymouth, MMI, SB and EM conducted the thematic analysis of the qualitative data. MB was responsible for the cost analysis. EM set the structure for the manuscript, MMI drafted the paper based on this framework, and it was reviewed by all authors.

Meinert E, **Milne-Ives M**, Taylor E, Inkster B, Paik A, Ananthakrishnan A, Orr M, Costelloe CE, Shankar R. Real-world testing of an artificial intelligence conversational agent as an early intervention and support tool in the mental health referral care pathway: A randomised controlled trial. *International Journal of Social Psychiatry*. 2024; Under Review.

Authors' contributions: All authors contributed to the study design.

Participants were recruited and screened and quantitative data was collected by AP. MMI conducted the participant interviews. CC and MO conducted statistical analysis of the quantitative data and MMI and AA conducted the thematic analysis of the qualitative data. EM set the structure for the manuscript, MMI and CC drafted the paper based on this framework, and it was reviewed by all authors with final revisions from EM.

5.2 Case Study 1: Barriers and Facilitators to Engagement with and Impact of a Childhood Obesity App

*Parts of this section have been previously published in [77] (including all or part of 5.2.2, 5.2.5-7).

5.2.1 Overview

The first case study that will be discussed is the evaluation of the NoObesity app (Figure 5-1). The aim of the overall study was to examine the factors associated with users' engagement with the app and to evaluate the app's impact on motivation, self-efficacy, and perceived behaviour change. As the first case study, this evaluation had a large impact on the overall thesis project (discussed in detail in [Section 5.2.3](#)). The analysis of barriers and facilitators to engagement highlighted the ongoing challenges in supporting engagement with DHIs and the need for further theoretical and empirical investigation to understand how this could be improved to support their impact.

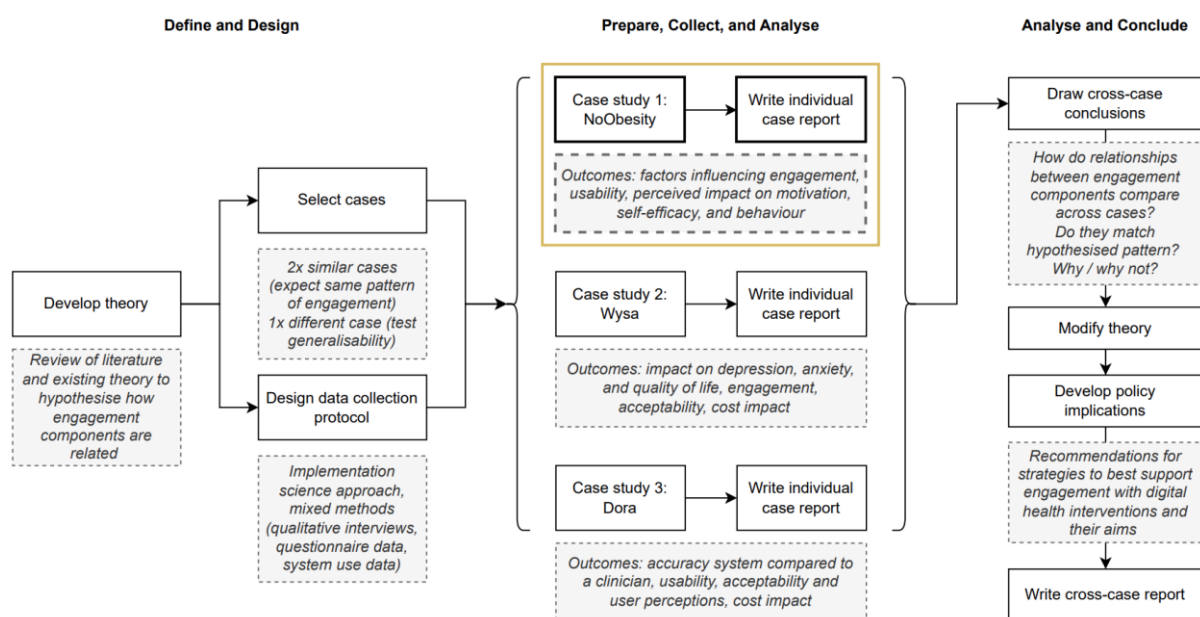


Figure 5-1. Case study 1 setting amongst the overall project

5.2.2 Results

5.2.2.1 Sample characteristics

Consented participants included 225 parents/guardians and 6 HCPs. 61 parents completed the demographic survey, 15 of whom were interviewed. Almost 40% of parents (85/225) completed the final questionnaire, but only 16% (35/225) met the criteria to be included in the quantitative analysis: completing both demographic and final surveys and passing an attention check. No HCPs completed the study (Figure 5-2). Most participants were female and white, with a mean age of 39 years (SD 6.0; Table 5-1). Of the participants initially invited for interview via stratified random sampling, 8 responded and 5 completed the interviews. After a second round of stratified random sampling, all eligible participants were invited to reach a sample of 15 interviewed participants. Participants who did not complete follow-up assessments may still have been using the app, as it was publicly available and usage was not controlled. Data was collected between September 2020 and April 2021.

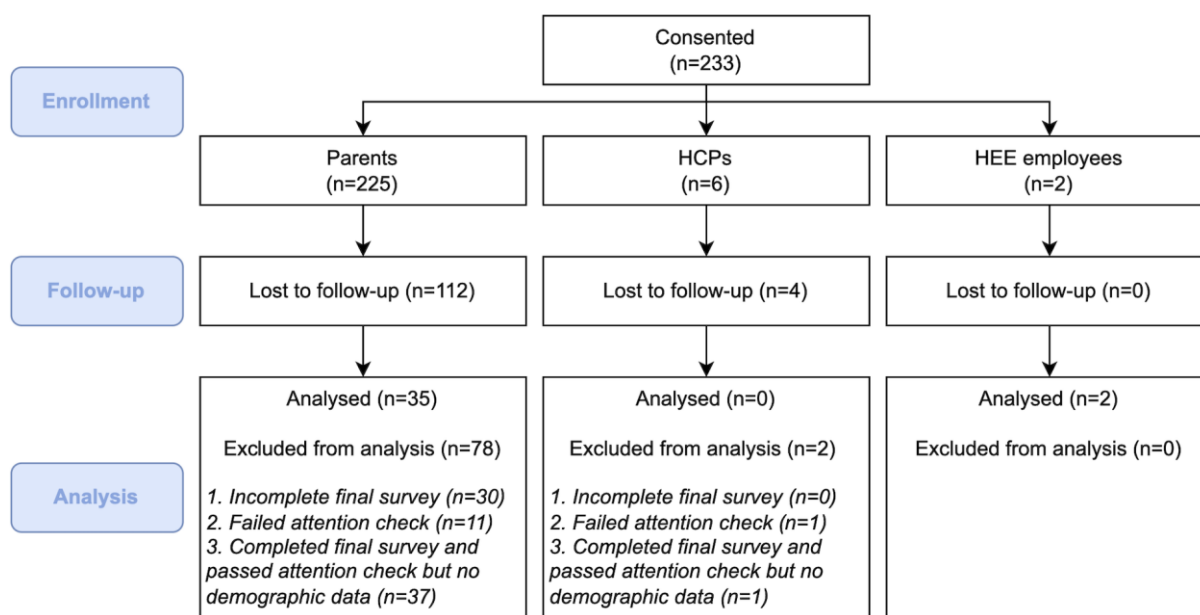


Figure 5-2. Case study 1 participant flow diagram

Table 5-1. Participant characteristics of the sample who completed the final quantitative survey (n=35) and interviewed sample (n=15)

Characteristic		Number (%) <i>final survey</i>	Number (%) <i>interviewed</i>
Gender			
	Female	31 (89)	14 (93)

	Male	4 (11)	1 (7)
Age (parent/guardian)			
	30-39	20 (57)	11 (73)
	40-49	12 (34)	3 (20)
	50-59	3 (9)	1 (7)
Ethnicity			
	White	32 (91)	12 (80)
	Asian	2 (6)	1 (7)
	Black	1 (3)	2 (13) ^a
Highest completed qualification			
	Bachelor's degree or higher	23 (66)	8 (53)
	Lower than Bachelor's degree	12 (34)	7 (47)
Annual household income before tax			
	£12,500 - 19,999	3 (9)	1 (7)
	£20,000 - 29,999	5 (14)	2 (13)
	£30,000 - 39,999	7 (20)	5 (33)
	£40,000 - 49,999	5 (14)	2 (13)
	£50,000 - 59,999	4 (11)	2 (13)
	£60,000 - 69,999	2 (6)	0 (0)
	£70,000 - 79,999	1 (3)	0 (0)
	≥ £80,000	6 (17)	1 (7)
	Prefer not to say	2 (6)	2 (13)
Employment			
	Working full time (30 hours a week or more)	12 (34)	3 (20)
	Working part time (8-29 hours a week)	15 (43)	7 (47)
	Doing housework, looking after children or other persons	4 (11)	3 (20)

	Student	2 (6)	1 (7)
	Permanently sick or disabled	2 (6)	1 (7)
Location			
	A country village	5 (14)	2 (13)
	A small city or town	24 (69)	10 (67)
	The suburbs or outskirts of a big city	4 (11)	1 (7)
	A big city	2 (6)	2 (13)
Number of adults in household			
	One	3 (9)	2 (13)
	Two	31 (89)	13 (87)
	Three or more	1 (3)	0 (0)
Number of children in household			
	One	11 (31)	8 (53)
	Two	16 (46)	5 (33)
	Three	6 (17)	2 (13)
	Four	2 (6)	0 (0)

^aNot all interviewed participants completed the final quantitative survey

5.2.2.2 Overview of thematic analysis

The initial thematic analysis grouped feedback into three main themes - facilitators and positive feedback, barriers and negative feedback, and suggestions for improvement - each including a variety of subthemes ([Appendix 5-A](#)). This write up of results and interpretation evolved substantially; initial drafts structured the analysis based on the RE-AIM framework and subsequent drafts used a more traditional thematic analysis structure [397–399] of organising the results by theme and subtheme and including a narrative and quotes for each. An additional analysis built on the initial thematic analysis by mapping the subthemes to the COM-B model to clarify how these factors could influence engagement with the intervention

and the target health behaviour. A revised thematic framework including this mapping is presented in [Appendix 5-B](#), which grouped factors into six key themes relating to engagement - motivation and usability - and perceived impact - on motivation, self-efficacy, health behaviours, and communication with HCPs. The following subsections are structured around the key factors, derived from the themes and subthemes generated in the thematic analysis, that were important for engagement (motivation and usability) and key barriers and facilitators to perceived impact. This structure was used to present results to align with the aim of the case study and to clearly highlight key factors for consideration in further research and intervention development.

5.2.2.3 Engagement with the intervention (adoption and use)

A large sample was initially recruited (n=225), suggesting that there was interest in the app, but there was substantial drop-out and discontinued app use (captured via self-report and app login data). Although most interviewed participants reported no plans to stop using NoObesity, only half of final survey respondents (19/35, 54%) reported still using the app and of the participants from whom app use data was collected (n=26), only a third (9/26, 34%) had recently logged in by or after the end of the study (April 2021). On average, there were 5.1 months between account setup and most recent login (95% CI: 3.9-6.3, range: 0.0-10.4 months). This section examines the key capability, opportunity, and motivational factors generated from the thematic analysis as influencing engagement with the intervention. As described in [Section 4.5.1.5](#), the factors were derived from the themes and subthemes of the thematic analysis and were associated with engagement or perceived impact based on the codes and associated participant quotes ([Appendix 5-B](#)). The key factors generated as influencing engagement with the intervention could be grouped into two themes: motivation and ability to engage (usability). These factors are summarised in Figures 5-3 and 5-4, which illustrate the directions of their association with engagement and their relation to the COM-B components; based on participants' feedback and the initial thematic analysis.

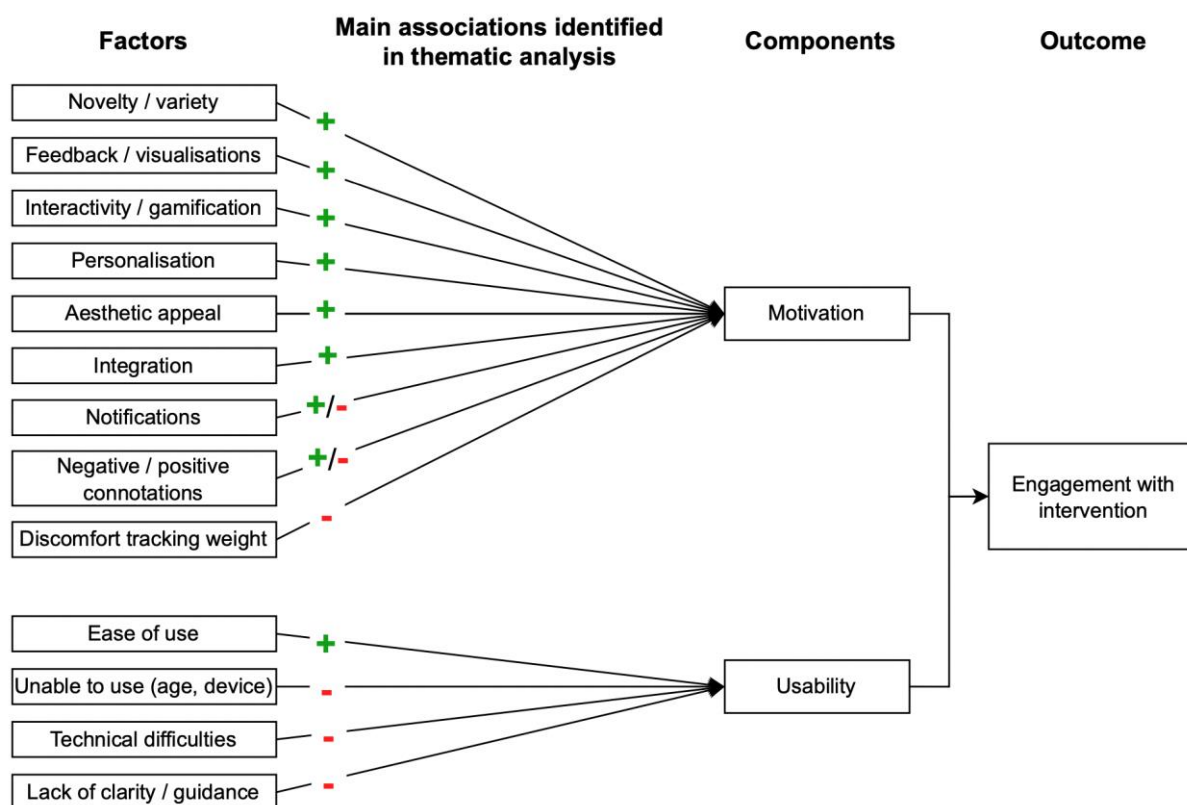


Figure 5-3. Summary of factors generated in the thematic analysis and how they were associated with engagement (*Note: + and - signs refer to whether the factor had a positive or negative influence on the outcome*)

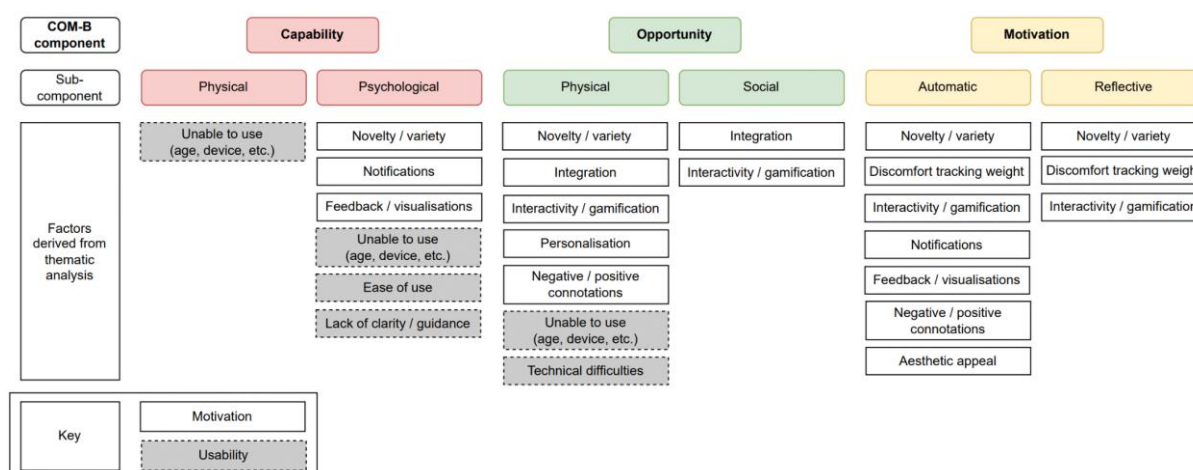


Figure 5-4. Summary of key factors for engagement mapped to the COM-B components

5.2.2.3.1 Factors influencing motivation to engage

Most of the factors that influenced participants' motivation to engage related to the automatic and reflective motivation components of COM-B, although several overlapped with psychological capability (eg. TDF domain: 'memory, attention, decision processes, and knowledge') and physical opportunity (eg. TDF domain: 'environmental context' offered by

the app). Novelty and variety were key for engagement and demonstrate this overlap; without something new to explore, participants felt bored, had reduced intentions to open the app, and would need increased cognitive effort to overcome these barriers (Textbox 5-1). Many participants, both in the interviews and the final questionnaire, reported finding the app repetitive and said that the limited content reduced their motivation to use the app. Participants all valued feedback, but several found the feedback provided by the app lacking or unhelpful. Visualisations of progress were suggested to help support motivation and more interactivity and gamification (e.g. daily challenges, tangible rewards, and friendly competition) to help engage children. Opinions about notifications were mixed: many participants appreciated the reminder but several found them annoying or poorly timed. A couple of participants did not like the app's aesthetic and felt this hindered use.

Positivity and a focus on behaviour rather than weight were important; some parents highlighted that while they might track calories or weight for themselves, they felt uncomfortable doing so for their children. This was also reflected in opinions of the name "NoObesity;" although most participants did not personally feel strongly about it, they thought it had negative connotations and that a more positive focus on health behaviours could improve new user engagement. The survey results supported this; participants reported somewhat disliking the name ($M = 2.6/5$, 95% CI: 2.2-3.0; Table 5-2).

Textbox 5-1. Sample of participant quotes from interviews about factors affecting engagement with the app

"Once you've gone through the sections it has, then there's nothing else you can really do with it. So there's not much interaction ... to keep you engaged." (ppt 2)

"Another reason why I didn't really do it is because... it feels quite repetitive having to click ... Tuesday, Wednesday, Thursday, Friday for every single challenge." (ppt 3)

"Had there been some new content on there, it would have encouraged me to come back more and more and go, "Oh, what more can I learn?"" (ppt 7)

5.2.2.3.2 Factors associated with usability

Factors associated with ability to engage were primarily related to psychological capability and physical opportunity. Overall, interviewees found the app "quite easy to use," in line with SUS scores ($M = 70.3/100$, 95% CI: 63.8-76.8), although ratings were slightly lower when parents completed the SUS from their child(ren)'s perspective ($M = 65.1/100$, 95% CI: 58.3-71.9). Barriers were primarily related to the technology (app freezing) and a lack of clarity and guidance in the app. Several parents reported not understanding the target audience, how

to report progress as a family unit if only some of the family members had successfully achieved the goal, or how the app's reward system worked, which hindered its potential impact on motivation (Textbox 5-2).

Textbox 5-2. Participant quotes from interviews about factors affecting usability

"Because of the iconography and how it looks, I think it's quite easy to use. It's quite self-explanatory." (ppt 4)

"It could've possibly done with an introduction page to start with." (ppt 6)

"There's a lot of external links. I think you could quite easily get lost looking and looking for things." (ppt 13)

"Two of us have gone out for a walk, but two of us haven't. Do I score it, do I not?" (ppt 11)

5.2.2.4 Perceived impact on health behaviours

Perceived impact on health behaviours was mixed and often conflicting; most participants reported some progress on their goals but many also said that the app did not have a major impact on their behaviour (Textbox 5-3). In-app self-reported goal progress indicated that families were successful at achieving any particular goal for a mean of 7.5 weeks (95% CI: 5.5-9.5), a median of 3 weeks, and a mode of 0 weeks (n=27). Factors relating to all three COM-B components were associated with the perceived impact on families' motivation, self-efficacy, and health behaviours (Figures 5-5 and 5-6).

Textbox 5-3. Participant quotes from interviews about perceived impact of the app

"The healthier meals ... last week we were able to tick off three or four, but I know we did more than that, it was just the actual ones I ticked off." (ppt 15)

"[I] don't necessarily think there were any changes. Let's say, we live very close to school, so it's easiest to walk. So that thing that I put in as our first goal about walking to school every day. We were going to use it every day, anyway." (ppt 4)

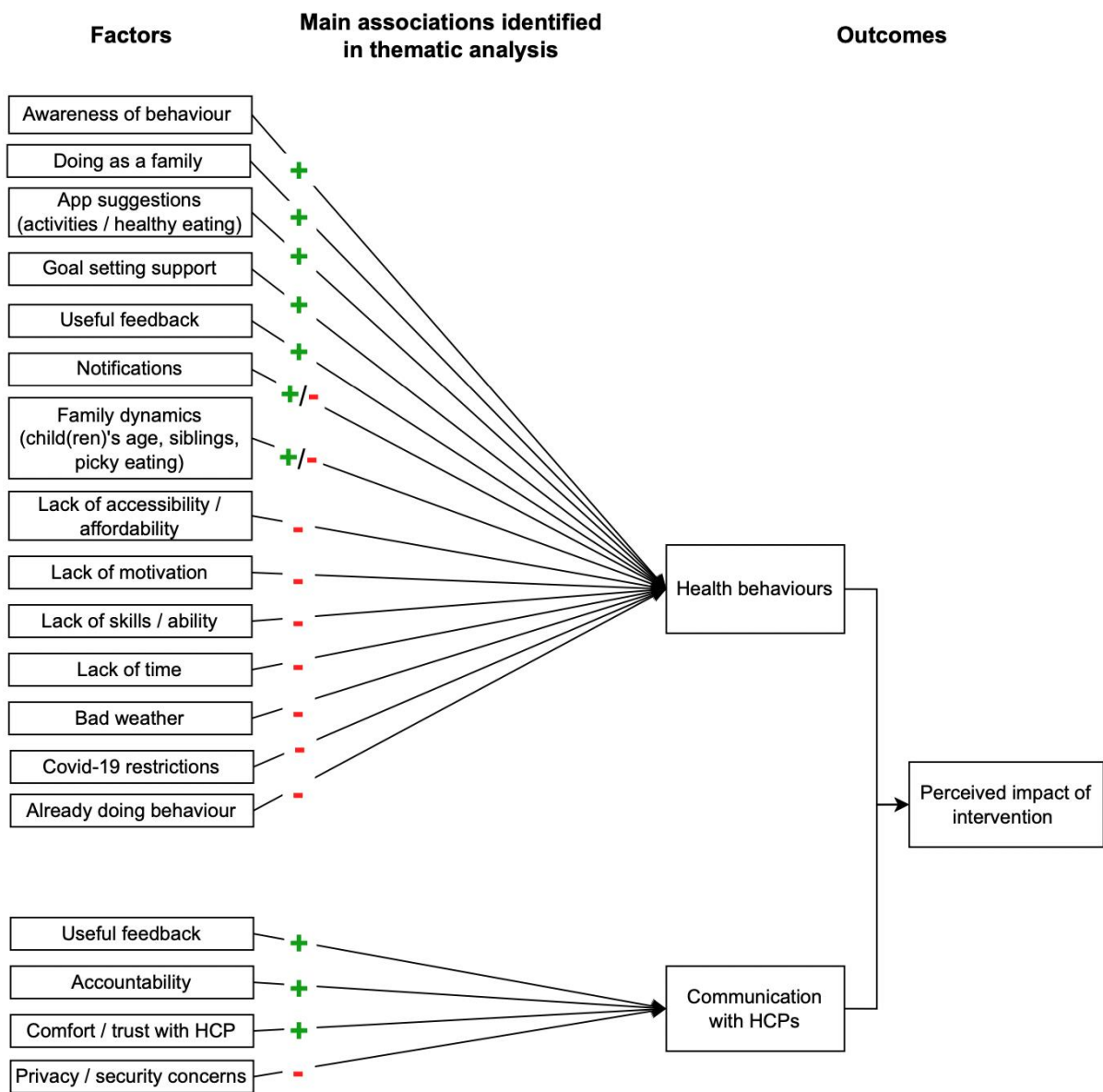


Figure 5-5. Summary of factors generated in the thematic analysis associated with the app's perceived impact (Note: + and - signs refer to whether the factor had a positive or negative influence on the outcome)

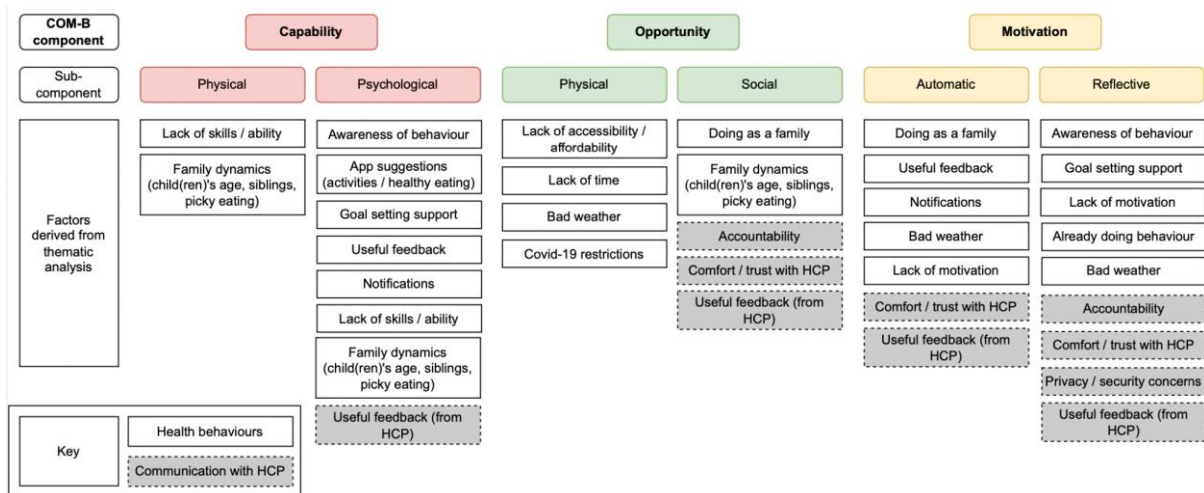


Figure 5-6. Summary of key factors for perceived impact mapped to the COM-B components

5.2.2.4.1 Facilitators for behaviours

Factors relating to the app's perceived impact on participants' motivation and self-efficacy were mostly related to motivation and psychological capability. Many participants found goal setting and progress recording motivating because it increased their awareness of their current behaviour and provided accountability (Textbox 5-4). Having “*achievable, attainable*” goals, “*being able to input data,*” and getting ideas for goal setting, physical activity, and healthy eating supported self-efficacy; which was largely aligned with the quantitative ratings of impact of app features on motivation, which were highest for ‘doing something together as a family’, ‘suggestions for activities and healthy eating’, and ‘goal setting’ (Table 5-2). Other facilitators included prompts and notifications (if they arrived at convenient times) and useful feedback (from the app or a linked HCP). The mean self-efficacy score per goal on the survey was 69% (SD = 20.4, range: 20-100%), similar to the mean confidence self-reported when setting goals on the app (3.9/6 stars, 65%).

Textbox 5-4. Participant quotes from interviews about factors affecting motivation and self-efficacy

“This kind of brings it forward for you to actually acknowledge what you're doing and what you're not doing, and to help you to correct it. So it was like, I'm not going to have a snack today because I want to say that I managed to not have a snack today.” (ppt 2)

“I would have liked to be able to see trends... it's nice to have a look at your graph, isn't it?” (ppt 1)

“If you've downloaded it and it's a family challenge, and it's telling you to do it all together, it does make it easier, I think” (ppt 3)

“I thought the useful links were good ... because if you're not an outdoorsy family, you might not even know where to start” (ppt 8)

“Without [feedback from the professional], it was completely lacking. [The app] didn't give me feedback on whether my goals were good ideas. It didn't give me feedback on how well I was doing meeting my goals.” (ppt 1)

Table 5-2. Participant ratings of the app and its features' perceived impact (n=35)

	Item	Mean (/5)	St. dev.	95% CI	Likert scale points meaning (1-5)
<i>App's support for behaviour</i>					
	Set goals	3·5	1·4	3·0-4·0	Strongly disagree - strongly agree
	Achieve goals	3·5	1·0	3·1-3·8	Strongly disagree - strongly agree
	Eat healthier	3·4	1·3	3·0-3·8	Strongly disagree - strongly agree
	Be more active	3·3	1·3	2·9-3·8	Strongly disagree - strongly agree
<i>Perceived impact of app on motivation</i>					
	Eat healthier	2·7	1·1	2·3-3·0	Not at all - extremely effective
	Be more active	2·6	1·1	2·2-3·0	Not at all - extremely effective
<i>Perceived impact of motivation strategy</i>					
	Doing something together as a family	3·2	1·3	2·8-3·6	Not at all - extremely effective
	Suggestions of activities	2·9	1·2	2·5-3·3	Not at all - extremely effective
	Suggestions for healthier eating	2·8	1·0	2·5-3·2	Not at all - extremely effective
	goal-setting	2·8	1·3	2·4-3·3	Not at all - extremely effective
	Self-monitoring	2·7	1·1	2·4-3·1	Not at all - extremely effective
	Points / trophies ^a	2·7	1·3	2·2-3·1	Not at all - extremely effective
	Notifications ^a	2·3	1·3	1·9-2·8	Not at all - extremely effective
	Feedback about current behaviour	2·2	0·9	1·9-2·5	Not at all - extremely effective
	Games	2·2	1·2	1·8-2·6	Not at all - extremely effective

	Family photo	1·9	1·3	1·5-2·3	Not at all - extremely effective
	Linking to a HCP	1·6	0·9	1·3-1·9	Not at all - extremely effective
Opinion of features					
	Points / trophies	3·5	1·1	3·1-3·8	Dislike a great deal - like a great deal
	Self-monitoring	2·7	1·0	2·4-3·1	Not at all - extremely useful
	Information feedback	2·6	1·1	2·3-3·0	Not at all - extremely useful
Opinion on healthcare provider link					
	Level of comfort	3·1	1·4	2·6-3·6	Extremely uncomfortable - extremely comfortable
	Perceived usefulness	2·8	1·4	2·4-3·3	Not at all - extremely useful
	Likelihood of linking	2·5	1·4	2·0-2·9	Extremely unlikely - extremely likely

**Sample was missing one data point (n = 34)*

5.2.2.4.2 Barriers to health behaviours

Most behavioural barriers were associated with physical opportunity, including Covid-19 restrictions, time, weather, accessibility, affordability and family dynamics such as children's age and abilities (Textbox 5-5). There were also barriers relating to social opportunity (children's willingness to try new foods and sibling teasing) and physical and psychological capability (barriers relating to childrens' age and skills). Some participants suggested potential capability barriers such as not knowing how to cook a healthy meal or not being aware of activities, but few reported experiencing these barriers themselves. A lack of motivation was also mentioned; sometimes related to physical opportunity barriers (e.g. outdoor activity in bad weather), but not always with a reason given.

Textbox 5-5. Participant quotes from interviews about factors affecting behaviours

"We don't have outside space for us. So, it's an effort to do it every day." (ppt 1)

"Some of the ideas were good, but there's just nowhere around here that we'd be able to do them." (ppt 10)

"Our time is very limited. I go out to work and I'm not back till half five/six. So once we've had a family meal it's dark, the children aren't wanting to go back out. I'm not really wanting to go back out. It's cold." (ppt 15)

"It's just easier to buy junk food, it's cheaper and easier to put... something in the oven rather than cook from scratch. And some people might not necessarily have the skills." (ppt 8)

5.2.2.5 Perceived impact on communication between healthcare providers and families

No HCPs completed the study and no participants connected with an HCP via the app (most had not seen an HCP during the intervention period because of the Covid-19 pandemic). Hypothetically, most parents felt comfortable communicating with HCPs via the app and that the HCP link could provide useful feedback and accountability (Textbox 5-6, Table 5-2), although a couple raised concerns about the app not reflecting true behaviour, bothering the HCP, discomfort discussing a sensitive topic, or data security.

Textbox 5-6. Participant quotes from interviews about factors affecting communication with HCPs

"I would have appreciated a real person saying, 'well, that's a good goal, or that's a bit wishy-washy. Perhaps you could make that better'." (ppt 1)

"It would be good because I think for me personally, if I'm accountable for something then, somebody checking in on you and going to look at what progress you've made...it would make me more determined to do it" (ppt 8)

"I don't really feel I've got a relationship with my GP. So for me, I think I felt a bit reluctant about doing that just because I don't know them." (ppt 11)

5.2.2.6 Organisational factors

Interviews with HEE employees reported organisational-level factors that could influence the app's sustainability. The app's focus fits within priority areas around childhood obesity and digital-first healthcare. Its holistic approach, combining workforce development and service delivery, spans the mandate of several governmental bodies (since the interviews, HEE itself has been incorporated into NHS England) [400]. The HEE employees interviewed felt that this could be a benefit, as inter-agency collaboration could provide access to more expertise, but also noted that it creates potential ambiguity around ownership, which could be a risk for longer-term maintenance.

5.2.3 Influence of Case Study 1 on the thesis

As the first case study conducted, and at an early stage in the process of my thesis, the findings from the evaluation of NoObesity had an important influence on the shape of the

thesis. The analysis of users' engagement with the NoObesity app provided a vivid demonstration of a common challenge in evaluations of DHIs: engagement with the intervention, beyond an initial period of exploration, was difficult to maintain. The COM-B framework was used in the analysis of the barriers and facilitators that influenced users' engagement with the mobile app and with the target health behaviours of physical activity, healthy diet, and communication between HCPs and families because it provided a method of exploring exactly how different barriers hinder behaviour and a theoretical means of mapping these barriers to specific BCTs that would be expected to help mitigate them. While the high attrition and low engagement with the NoObesity app was not an unusual observation, it made me curious about how people engage with DHIs and how that engagement could be better supported to improve their impact.

As described in [Section 7.5](#), the original plan for my thesis had comprised this case study evaluation, followed by an app redesign process and re-evaluation. This first case study identified the need to determine how to improve engagement with the intervention, and its impact on behavioural and health outcomes, which drove an examination of the recent theoretical conceptualisations of engagement in the literature and resulted in the question examined in Chapter 2 - how BCTs are related to the different components of engagement. The evaluation conducted in this case study informed my thinking around that review; the COM-B framework identifies capability, opportunity, and motivation as the main components of behaviour and highlights a potential parallel with the multifaceted conceptualisation of engagement as including affective, cognitive, and behavioural components. Although these components do not directly align, the observation in this case study that barriers to behaviour could relate to these different components and require different strategies of mitigation suggested that there may be barriers to engagement that relate specifically to different components and that engagement could be improved by identifying and addressing the specific barriers to engagement that a user was experiencing. The aim of the systematic review was to generate hypotheses for how to use specific BCTs to improve engagement by targeting the component of engagement that was blocked. Instead, I found a lack of consideration of the different components of engagement in the literature that highlighted a gap in our understanding of how users engage with DHIs and how this can be examined. Chapter 3 provided a theoretical analysis of the literature to suggest a model that could help fill this gap. Chapter 4 and [Section 7.5](#) outline how the plan for this thesis evolved to address this question around the interaction between components of engagement and how the three case studies were used to test and improve the proposed model.

Based on the research conducted in Chapters 2 and 3, the original analysis of this case study was re-examined in [Section 5.2.4](#) to explore how it could inform our understanding of how the different components of engagement were associated with each other over the process of engagement with the intervention and target health behaviours. This analysis was limited by the way in which the SSIs were conducted. As the topic guide had not been designed to elicit users' experiences of the different components of engagement and the process by which they engaged with the DHIs, the mapping of the relationship between engagement components was preliminary and based only on what users expressed about their engagement. For example, it was not possible to generate, based on the data, a model of how initial engagement with the intervention proceeded, because there was little discussion in the interviews about how users became aware of the intervention and what they felt and thought about it before trying it for the first time. This limitation shaped the way that the SSIs in the second and third case studies were conducted. The TFA was used to help structure topic guides for these studies because the way acceptability is defined in the TFA reflects "anticipated or experienced cognitive and emotional responses to the intervention" [320] and is well-aligned with the multifaceted concept of engagement, as discussed in [Section 3.5](#).

5.2.4 Mapping of relationship between engagement components

Based on the data collected in the SSIs and the analysis of the key themes and factors relating to engagement, an exploratory examination of how participants experienced the different components of engagement and how they interacted to influence overall engagement was conducted (Figure 5-7, [Appendix 5-B](#)). The semi-structured topic guide was not designed with the aim of identifying particular components of engagement, so the analysis was based on how participants described their engagement. There was little data about how affective and cognitive engagement influenced users' initial behavioural engagement with the intervention, but many of the factors influencing subsequent micro and macro behavioural engagement with the intervention could be mapped to cognitive and affective engagement.

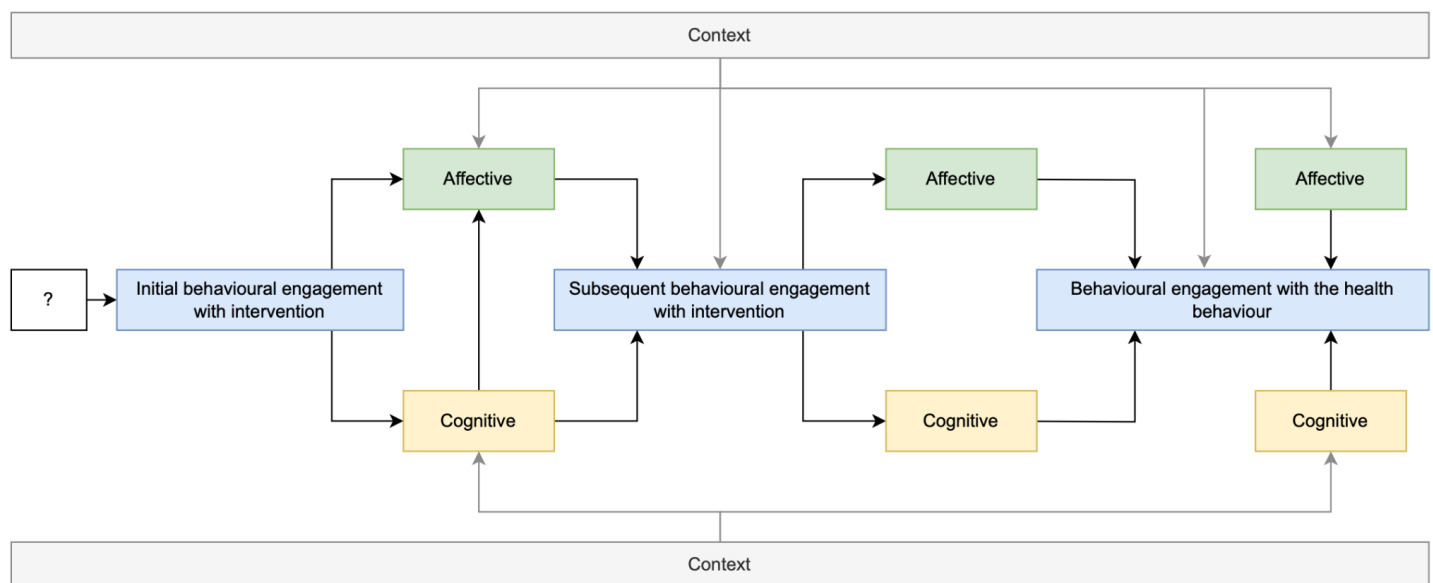


Figure 5-7. Mapping of engagement components derived from NoObesity SSIs

Affective engagement with the app was influenced by users' initial behavioural engagement with the intervention; their experiences using the app affected how they felt about it. The app's aesthetic design could influence the likelihood of further behavioural engagement - if users did not enjoy the visual experience of the app it could negatively influence their mood and reduce their enjoyment of the app. Likewise, the tone of the app - in terms of how information and content was presented - could influence affective engagement, with more positive affective engagement supporting further behavioural engagement with the app. The ability to personalise the app to their family increased users' sense of connection with the app, increasing affective engagement.

Some factors influenced both cognitive and affective engagement with the app. Some participants reported that the aesthetic design of the app looked interesting, which supported their behavioural engagement with it. Games and interactivity were noted as being fun, enjoyable, and providing a sense of competition and reward that boosted affective engagement, encouraging further behavioural engagement. Negative affective engagement could decrease the likelihood of further behavioural engagement: participants' expressed discomfort with the idea of tracking their children's weight and concerns about how that would affect their children, which made them unwilling to use that feature of the app. Likewise, feelings of irritation (e.g. related to poorly timed notifications) also decreased motivation to engage.

In the case of the notifications, the pattern of engagement is relatively clear; when participants were made aware of a notification, they were cognitively engaging with the app. If that

notification was drawing their cognitive engagement too frequently or at inconvenient times, it caused feelings of irritation or frustration (negative affective engagement), indicating a pattern of engagement of cognitive → affective → subsequent behavioural engagement. In other cases where both affective and cognitive elements of engagement were involved, it was often difficult to determine from the data captured the way in which the engagement components influenced each other. For example, motivation relating to gamification included elements relating to both cognitive and affective engagement, but not in a way that could be clearly ordered.

The cognitive effort needed to use the app was also related to behavioural engagement - most participants felt that the app was easy to use and navigate, facilitating behavioural engagement, but a lack of guidance and a repetitive process of entering data became frustrating (leading to affective and behavioural disengagement). There was both cognitive and affective engagement with the app content initially - participants' attention to the information was educational and they enjoyed the learning process - but once they had consumed all of the content, feelings of boredom and a lack of further content to cognitively engage with resulted in disengagement from the app. Cognitive engagement with risk-benefit judgments could also influence behavioural engagement; for example, participants who had data privacy concerns reported not wanting to use the app to share data with a healthcare professional.

Certain factors also influenced engagement with both the intervention and health behaviour. For instance, participants felt that feedback (from the app or from a HCP via the app) could help support cognitive and behavioural engagement by providing support for them to adapt their goals in the app and to generate new strategies for engaging with the behaviour, if what they were currently trying was unsuccessful. Notifications also facilitated cognitive and behavioural engagement by catching participants' attention and reminding them to use the app to record their progress or to engage in the behaviour. Behavioural engagement with the health behaviour also had the potential to influence behavioural engagement with the app; several participants reported already doing the behaviour that they had set as a goal and recognising the need to return to the app to 'level up' their goals.

Behavioural engagement with the intervention could also influence behavioural engagement with the health behaviour via cognitive and affective engagement. Several participants reported that using the app increased their awareness of their behaviour, which influenced their cognitive and affective engagement with the behaviour by helping them to assess whether their current behaviour was aligned with their goals and values. Likewise, the app's

suggestions for how to set health behaviour goals, make healthier food decisions, and engage in more and different activities helped support cognitive engagement with the health behaviour by providing ideas for how to improve it, which facilitated behavioural engagement with those goals.

Affective engagement with the behaviour could also influence behavioural engagement. Several participants mentioned that a sense of accountability could influence their behaviour, in both positive and negative ways. For some, this motivation to engage in the behaviour was because they did not want to feel judged by or guilty about their behaviour if HCPs were looking at their progress. Others enjoyed seeing their behavioural progress represented on the app and this encouraged them to hold themselves accountable and continue improving their behaviour. Positive affective engagement with the DHI - e.g. enjoying engaging with the intervention and goals as a family - could help increase motivation for sustained behaviour engagement. Negative affective engagement with the DHI may not hinder motivation for the health behaviour, but is unlikely to help support it.

Finally, there was a significant influence of context on engagement with the health behaviours. Many of the barriers to behaviour were related to opportunity - contextual factors such as time, family dynamics, work patterns, the availability of appropriate resources and environments, weather, and Covid-19 restrictions. The influence of several of these factors on behavioural engagement was mediated by affective or cognitive engagement. For example, engaging in physical activity outdoors in bad weather or early darkness (in the winter) was perceived as less enjoyable, reducing affective and behavioural engagement. The high price of healthy foods or the time required to prepare them might tip the cognitive cost-benefit assessment and reduce the likelihood of behavioural engagement. This emphasises the importance of including contextual factors in models of engagement.

5.2.5 Recommendations in the context of existing evidence

There are many DHIs that have been developed and evaluated to target childhood obesity, but while studies often assess user perceptions and acceptability, there is limited investigation of factors influencing the components of engagement with mobile health apps for childhood obesity [107,401]. A recent evaluation of the Aim2Be app for childhood obesity identified different patterns of engagement and associated them with demographic characteristics, but not user perceptions [402]. Another study conducted a formative co-design process with parents to identify preferences for engagement and barriers to the health behaviours, which were largely similar to those identified in [Section 5.2.2](#), but they were not analysed in depth in

the paper [403]. For this reason, the findings from this case study are examined in comparison to factors influencing engagement with DHIs more generally.

All three COM-B components were important for DHI engagement and impact [185,252,404–409]. Facilitators for engagement - a simple visually-appealing interface, feedback and visualisations of progress, guidance, customisation, reminders, access to a HCP, and positive messaging - were aligned with previous recommendations in the literature [122,201,202,225,410]. Both affective and cognitive engagement with the intervention were important in supporting continued behavioural engagement with the app. A lack of interest in the app, once all of the content had been consumed, was a key factor in micro behavioural disengagement; DHIs will need to find a means of maintaining either interest or enjoyment, or at the very least demonstrating clear benefits of using the app, to support sufficient engagement to enable the achievement of the intervention aim (in this case, establishing healthy behavioural habits).

Several factors, including common engagement strategies like gamification [411–413], notifications [407,414], and competition [268], could be facilitators or barriers and influence both affective and cognitive engagement. For example, points and trophies alone may not be sufficiently meaningful to support motivation [413] and that the impact of notifications on subsequent micro behavioural engagement can depend on their timing, frequency, and level of personalisation [407,414–416]. Friendly competition could help keep children affectively engaged with the intervention by making the experience of using the app fun and increasing their motivation to behaviourally engage with it [201]; however, parents were concerned about teasing and negative self-image that might be consequences of competitions and comparison between friends or siblings. Social comparison can be demotivating [268], but this could be mitigated through team-based competition that de-emphasises the individual [47] while maintaining affective engagement. For childhood obesity, enabling intergroup family-team-based competition could help support engagement and focus on positive behaviours rather than weight. Personalisation of interventions, rather than ‘one-size-fits-all’ strategies, could also support affective and behavioural engagement by increasing users’ sense of connection with the app. Although not directly evidenced in the interviews, it also seems likely that personalisation (of feedback or its mode of delivery) could support cognitive engagement with the intervention, as it could facilitate processing of the information [417] and interest and attention, as it is likely to be perceived as more personally relevant.

Key facilitators for behavioural impact included awareness of behaviour, accountability, motivation, prompts, suggestions, family characteristics, and accessibility. Features that

support a family focus, goal-setting, progress recording, positive messaging, feedback, and suggestions and inspiration could help support perceived motivation and self-efficacy for behaviour change in this context, in line with meta-analytic evidence in adults [418]. The relationship between motivation and accountability (e.g. via progress recording and HCP-link) has been previously observed [419,420] and theoretically linked to adherence [421]. Accountability has the potential to help support behaviours not done for enjoyment [421,422], so more active involvement of a linked HCP with families via the app may have helped increase cognitive and behavioural engagement with the intervention and behavioural goals; however, this may not work for everyone, as some participants worried that their reported progress did not reflect real behaviour and that they would be judged, which could reduce affective engagement.

Key contextual, structural-level barriers (e.g. time [406], affordable healthy foods [423], access to physical activity environments [410]) and individual-level barriers (e.g. knowledge, awareness, motivation, and self-efficacy) were also largely aligned with existing evidence [406,409,410,423,424]. Structural inequities can hinder weight management and contribute to feelings of stigma or blame that reduce motivation [425,426]; mitigating these barriers (e.g. by highlighting affordable healthy meal options or locating free resources personalised to the users' location) will be necessary to support macro behavioural engagement. Individual-level barriers were more aligned with motivation than capability [53], perhaps because families less aware of weight issues may have been less likely to use an app called "NoObesity." As a key facilitator for affective engagement was the app's family focus, developers are recommended to consider the psychology and motivation behind their users' behavioural engagement and how they can frame their intervention as something enjoyable rather than necessary but onerous [427]. Overall, digital intervention designers are recommended to research specific barriers (especially opportunity barriers) for their target populations and behaviours to identify the most appropriate means of mitigating them in that context. Table 5-3 highlights additional recommendations for intervention development, through the lens of the NASSS and RE-AIM implementation frameworks [341–344].

Table 5-3. Summary of implications for implementing similar interventions in the future

NASSS domains	RE-AIM dimensions	Insights for designing interventions for successful implementation
Domain 1: The condition or illness	N/A	• Social determinants of health - such as income, environment, employment, etc. - are important contextual factors that affect families' abilities to engage in healthy weight-related behaviours. • DHIs should account for these contextual differences in users by incorporating advice and behavioural suggestions that meet their varying abilities and financial and environmental circumstances. • For example, this could include providing advice or recipes on how to eat healthy on a tight budget and including suggestions of physical activities that can be engaged in within a house if outdoor spaces are unavailable or inaccessible.
Domain 2: The technology	Implementation	• The importance of novelty as a factor to support affective and cognitive engagement raises a potential issue for not-for-profit organisations that do not have capacity or funding for regular updates and uploading of new content for an intervention. • These types of limitations should be accounted for in the intervention design, for example, by setting up the intervention to provide content periodically over the course of the intervention or by enabling users to generate content (where appropriate). • Specific implementation strategies should be developed in collaboration with the target population, to ensure that the intervention meets their needs and circumstances.
Domain 3: The value proposition	Effectiveness	• To have a strong case to support public health interventions, evidence of positive impact is key - on an individual level, users should be able to see evidence of impact for themselves (eg. through progress monitoring and feedback from the intervention, or prompted self-reflection); on an organisational level, there needs to be significant evidence of positive impact to support widespread adoption. • Evaluations of effectiveness should include assessments of engagement as a prerequisite for impact (and intentionally look to capture affective, cognitive, and behavioural engagement outcomes) and should include end-user perspectives on what

		key outcomes they consider desirable (e.g. weight, mood, perceived energy, ability to do certain activities).
Domain 4: The adopter system	Adoption	- Given the high demand on HCPs' time and changes in practices and routines associated with implementation, perceived value, familiarity, and ease of use will be essential to clinical adoption [408]. - User-centred design processes should include a variety of clinical stakeholders, including those who would be directly engaging with the intervention and those responsible for managing adoption and implementation processes [428]. - This will be needed to identify and address challenges to adoption, whether they are related to perceived value and desirability of the intervention or structural barriers such as its lack of integration into existing routines.
	Reach	- Social media and app store advertisement is unlikely to be sufficient to reach the target population in many cases; to increase reach, organisations should diversify dissemination by engaging with community and clinical groups and settings, specific to their target population's clinical and demographic characteristics. - The use of user-centred design with the target population could also improve reach by ensuring that, as well as being aware of the intervention, the target population expects that it will add value.
Domain 5: The organisation	N/A	Within governmental and healthcare organisations, political changes could provide a potential challenge for continuity and responsibility of interventions; solutions for such issues will be case dependent, but should be considered from the outset to ensure there is a plan for sustainability and maintenance.
Domain 6: The wider context		
Domain 7: Embedding and adaptation over time	Maintenance	DHIs that aim to change behaviour contributing to health outcomes should use technical and social features that facilitate the incorporation of these behaviours into regular routines (see recommendations for features above).

5.2.6 Strengths, limitations, and future research

Strengths included the use of mixed methods (enabling triangulation and improving confidence in results), theoretical frameworks (which guided the investigation of a comprehensive set of factors influencing adoption and implementation), and independent and collaborative coding by several authors for a rigorous thematic analysis. Due to time and resource limitations, mapping to the COM-B model and mapping of engagement components were only completed by one author, which limits the robustness of that analysis. Several strategies were used to mitigate the potential bias of only having one perspective in the interpretation of the findings for these analyses. The COM-B mapping was structured around the theoretical COM-B and TDF frameworks; a coding framework was developed based on the descriptions and alignment of these two frameworks, with examples of their application in similar digital health studies included to provide further guidance of how factors should be coded against this framework ([Appendix 4-F](#)). The mapping of engagement components was conducted using the COM-B and TDF analysis and the definitions of affective, cognitive, and behavioural engagement in [Section 3.6.3](#). Additionally, this analysis was reviewed and revised several months after being initially conducted so that I could consider and challenge my initial interpretation. In doing so, I generated [Appendix 5-C](#) to provide evidence of how the mapping was conducted and to demonstrate how Figure 5-7 was generated from the analysis.

The main limitation of the overall case study was the sample; high dropout and non-representative demographics [429,430] created potential bias. Stratified random sampling was used to improve the diversity of the SSIs, but the limited overall sample and the need to invite all eligible participants to complete 15 interviews meant that interviewees' perspectives are likely to represent a particular set of lived experiences, reducing generalisability. The convenience method of sampling via social media also introduces potential bias, as participants who respond are likely to be those with the strongest opinions about the intervention and the reach of the recruitment materials may not reflect a representative sample of the population. Random sampling could reduce this potential bias, although it may not reflect real-world engagement. Alternate methods of recruitment may have more success in increasing the sample size; for example, recruiting families and HCPs via schools, children's centres, GP practices, and other community organisations. A larger-scale study with a more representative sample will be needed to evaluate the impact of the intervention on health behaviour change and childhood obesity. Future studies would likely benefit from capturing demographic data immediately after consent, to enable analysis of participant characteristics

in the initial and final samples; the lack of this data in this case study meant that potential associations between demographic characteristics and drop out could not be explored. As structural factors relating to accessibility and affordability were key barriers, future research could also benefit from exploring how engagement (particularly with the health behaviour) might be associated with socioeconomic status.

Other limitations included the exclusion of children (which was decided for this initial study to facilitate data collection and ethical approval) and the monetary reward for participants who completed the study (which might have incentivised app engagement but was not included as a factor in the analysis). Data about the age of the children was not collected in the demographic data, but this was highlighted as an important factor in the qualitative analysis for consideration in future design and evaluation. Additional factors that could influence engagement and perceived impact are initial behavioural- and weight-related characteristics; these were not assessed as weight was not an inclusion criteria, but may be important factors to consider in future research.

The lack of data from HCPs meant that their perceptions of the intervention or their reasons for not participating in the study could not be assessed. The intervention period took place during the Covid-19 pandemic; unusually high demands on HCP's time might have precluded engagement with the study and intervention, but lack of participation could also be related to a lack of awareness of the study or HCP perceptions of the app. This will be an important area for future investigation, as previous research in digital health has suggested that integrating human support can improve engagement and is important for childhood obesity management [431,432].

Our ability to triangulate qualitative and quantitative data was also limited; the app use data that the system could record was relatively minimal and did not enable detailed examination of the frequency, intensity, time, or type of engagement (it captured first and last login only, rather more detailed data capturing number of logins or use of intervention features) [44]. The app is being redesigned based on theoretical frameworks and the findings of this study; this process will also improve its ability to capture app usage data. These limitations described here will be addressed in future larger-scale efficacy and effectiveness studies.

5.2.7 Conclusions

The growing prevalence of childhood obesity and the overwhelming demand for healthcare resources more generally has resulted in a need for easily accessible interventions to empower families to manage weight-related health. The benefits of DHIs are often limited by

insufficient engagement, so understanding facilitators and barriers to engagement is essential. This case study highlighted how the use of a theoretical behavioural framework can help clarify key barriers to engagement with a DHI and its target behaviour(s) and suggest mitigations. It highlighted the importance of considering affective and cognitive, as well as behavioural, components when evaluating engagement with DHIs. Both affective and cognitive engagement influenced subsequent behavioural engagement with the intervention and with the target behaviour (directly and indirectly). While there was limited evidence to determine a strong order of association between cognitive and affective engagement, the qualitative data highlighted how both components can have relatively separate and closely integrated influences on behavioural engagement. Based on these analyses, recommendations for digital intervention designers include incorporating interactivity, novel content and suggestions, goal setting and progress monitoring, feedback and accountability, reminders, guidance on how to use the app, personalisation, and a positive and visually-appealing design and tone. The caveat is that developers must identify whether these suggested intervention strategies align with engagement barriers and patient or other factors in their particular context; if strategies address capability and motivation barriers but not opportunity barriers, the intervention is unlikely to succeed. In terms of implementation, the NASSS and RE-AIM domains highlight key, interconnected factors that can influence the success of an intervention - particularly important is the need to develop and demonstrate a value proposition that meets the needs and circumstances of the target users and clinical adopting systems, which can best be executed by adopting user-centred design practices to ensure that solutions to potential barriers are incorporated from early stages in development. Although the alignment of our findings with previous research indicates that these factors are important more generally, context can affect how they influence engagement and behaviour. Further research is needed to examine factors associated with clinical adoption and to better understand how engagement strategies can be tailored to address specific barriers, behaviours, individual characteristics, and engagement components.

5.3 Case Study 2: Engagement and Efficacy of a Conversational Agent App for Mental Health Support

*Parts of this section have been submitted for publication in [129] (including all or part of Sections 5.3.2, 5.3.4-7)

5.3.1 Overview

The second case study conducted an evaluation of Wysa, an AI conversational agent (CA) app for mental health support (Figure 5-8). The full study was a randomised controlled trial primarily comparing change in severity of symptoms of depression and anxiety and quality of life of patients waiting for mental health Talking Therapies between the intervention group and a waitlist control. For the purposes of the thesis, this case study will focus on how users engaged with the app, derived from qualitative interviews and app use data; full results of the RCT have been submitted for publication elsewhere [129].

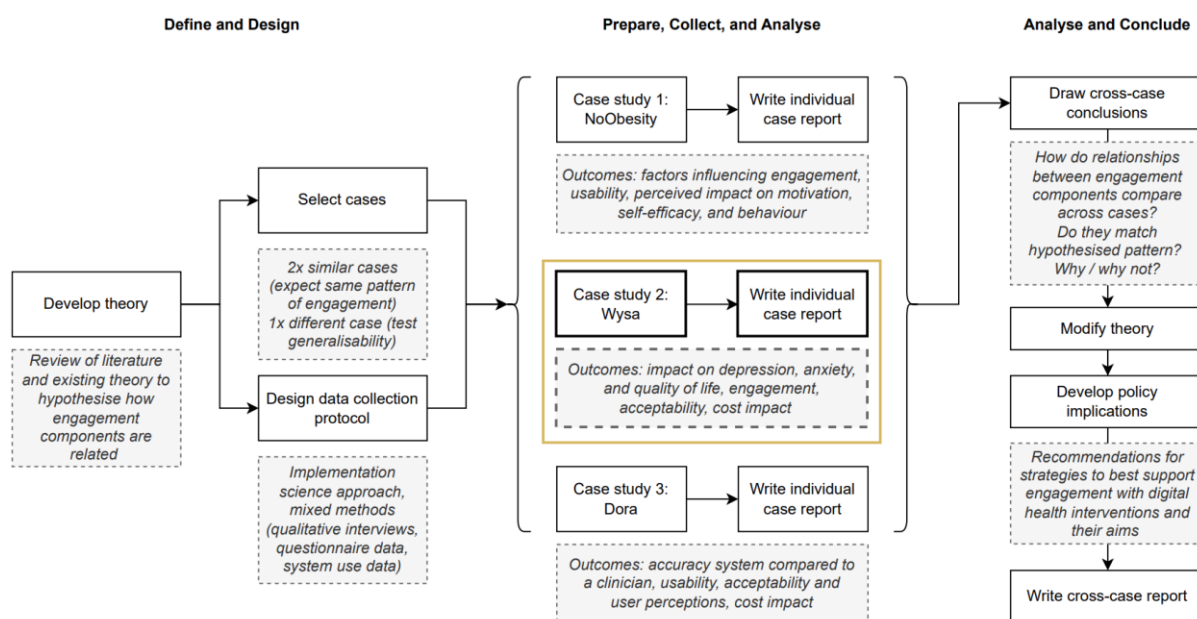


Figure 5-8. Case study 2 setting amongst the overall project

5.3.2 Results

5.3.2.1 Study flow and sample characteristics

2,161 patients were assessed for eligibility; 1,536 were excluded (Figure 5-9), mostly for PHQ-9 scores >15 (n=784) or being booked in for treatment (n=435). Many initially eligible participants were later excluded because they were booked into an appointment (no longer on waitlist). 99 patients consented; 76 completed baseline questionnaires and were randomised. Thirty participants did not complete the 12-week follow-up questionnaire, resulting in a sample of 46. Data was collected between 13 December 2022 and 31 July 2023. For the qualitative analysis, 10 out of 20 participants were reached and scheduled for an interview. Six participants were 'no-shows' and 4 completed an interview between 22 June and 25 July 2023.

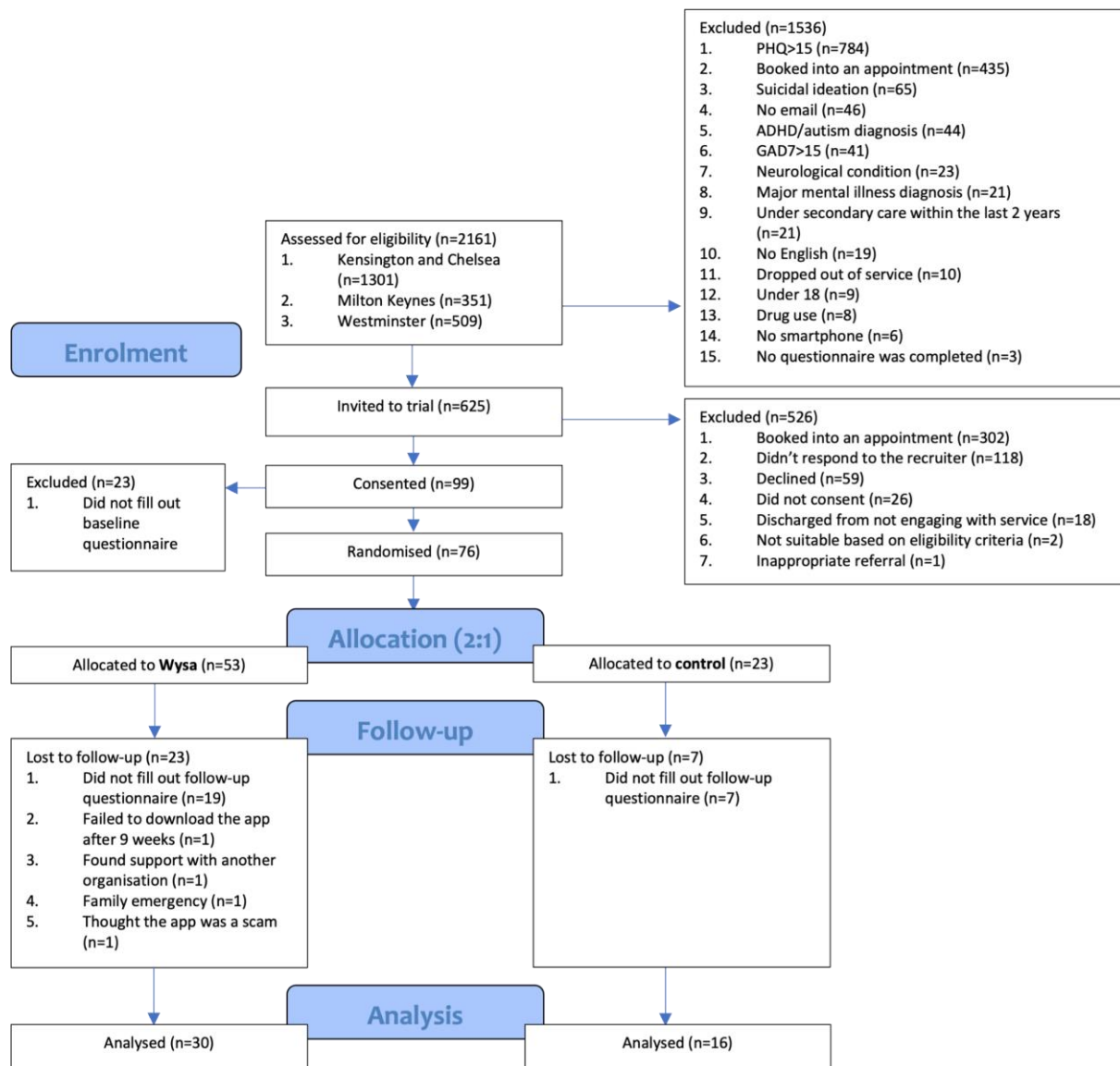


Figure 5-9. Case study 2 participant flow diagram

The 4 interviewed participants had a variety of demographic characteristics: gender (2 women, 1 man, 1 declined to respond), ethnicities (3 White, 1 Asian), age (3 were 25-64, 1 was 18-24), deprivation (2 each above and below median, based on postcode deprivation index), and initial PHQ-9 score (3 moderate, 1 mild).

There were 30 participants in the intervention group from whom app use data was collected. This group had a mean age of 36.0 years (SD = 14.4), was 63% female (19/30, including one transgender female), and 70% White (21/30). Baseline characteristics were similar between the groups for these characteristics except for ethnicity (control group was 50% White). There were a few other imbalances between groups which are submitted for publication elsewhere [129] but were not included here as the relevant analysis for this thesis primarily focuses on the subset of qualitative participants.

5.3.2.2 Overview of thematic analysis

The initial thematic framework developed in the case study analysis included three main themes: factors associated with engagement, perceptions of potential usefulness in clinical pathways, and how patients used the app. These themes, subthemes, and associated codes are presented in [Appendix 5-D](#). The final themes and subthemes presented below were refined through collaboration with another researcher ([Appendix 5-E](#)); participant quotes are integrated into the theme descriptions, but further examples are included in [Appendix 5-F](#). Several key themes relating to users' experiences with the app were generated from the qualitative interviews with the participants (Figure 5-10), relating to all of the components of engagement and most of the TFA constructs.

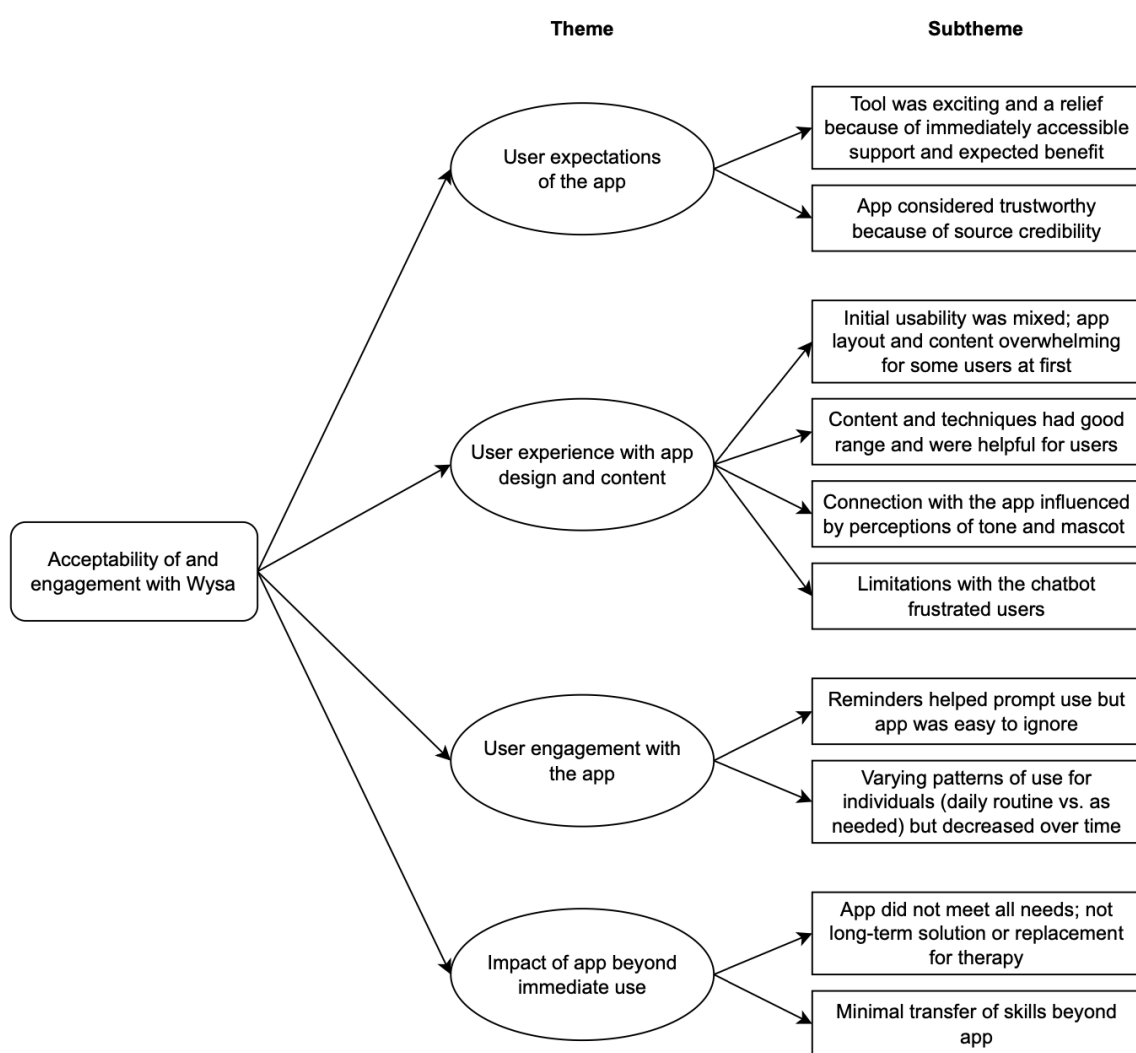


Figure 5-10. Thematic map of qualitative interviews to determine acceptability of and engagement with Wysa (Note: the thematic mapping involved two researchers; Figure 5-10 represents the final collaborative version of the themes and thematic map, the initial

independent thematic framework, maps, and alignment process are included in [Appendix 5-D](#) and [5-E](#))

5.3.2.2.1 Theme 1: User expectations of the app

Initial affective engagement with and attitudes towards the app were positive - participants were curious, excited, and relieved to have access to support (without “bothering anyone”) while they were on the waitlist and generally expected emotional benefits from Wysa. Participants generally considered the app trustworthy, for two key reasons relating to cognitive and affective engagement: 1) source credibility (the app was received via the NHS and cited evidence-based sources) and 2) perceived benefits from the app (“credible because it helps”). No participants reported ethical concerns - the only potential concern was privacy, but this depended on the extent of the personal details shared with the app.

5.3.2.2.2 Theme 2: User experience with app design and content

Beyond initial expectations, several factors influenced users’ experience. An important factor influencing affective engagement and attitudes was the app’s tone; though participants’ perceptions of the app’s tone differed (“it was very judgement free” vs. “when you’re in that very vulnerable state, you can take it as being sarcastic or criticising”), both perspectives highlighted the importance of a positive, genuine, supportive tone of communication from Wysa. Tone and connection could be shaped by design as well as content; participants also differed in their perspectives of Wysa’s penguin ‘mascot’, with one thinking it helps “keep people interested” and another that it “makes it too cartoony and just a bit less serious.”

Cognitive engagement was facilitated by the app’s ease of use - all participants found that the app required little effort to use, did not have any significant costs (low opportunity costs), and was relatively easy to navigate (low burden), although a couple felt that this could be improved as the “amount of things you get given at the very beginning ... was overwhelming.” This critique was limited to the layout - the amount of activities and content available on the app was noted as a positive: one participant said that a “highlight was ... the wide variety of stuff.” The key issue with users’ experience was the CA. This issue overarched all three components of engagement: all four interviewees reported that the CA’s ability to converse was limited and that the algorithm was predictable, repetitive, and recalled previous conversations inconsistently. This “became frustrating,” “annoying,” and “tedious” and reduced participants’ confidence in and use of the app.

5.3.2.2.3 Theme 3: User engagement with the app

Patterns of app use varied among the participants. Some incorporated the app into daily routines, completing morning and evening check-ins; others did not use the app regularly but only when they had the time and space or felt in need of support. Content length and format (primarily videos) was an issue for one participant, who felt that they did not have time to engage with this during the day when they might need, and instead suggested the app “just be like ‘do these breathing exercises’.” In terms of improving their engagement with the app, participants found that the reminders were useful prompts that helped “trigger [them] to use it” but thought that the app could better grab their attention and “was very easy to ignore”. Some participants suggested that enabling customisation of reminder times would give them “a lot of freedom”. One participant suggested that notifications or widgets that stuck to their screens and looked more interesting could help encourage them to open the app.

5.3.2.2.4 Theme 4: Impact of app beyond immediate use

Participants generally felt that Wysa was a good tool with some benefits (e.g. “the sleep sounds were the best, I used them quite a lot” and “the chats were really helpful”) and could be a useful introduction for patients, but that it did not meet all of their needs (eg. “I’d say about 30% of the time I got what I wanted from it” and “I would give it, on a scale of 10, maybe 6”). Overall, the sense was that it has potential and the resources could be “sufficient for [some users’] needs” but wouldn’t “replace someone getting one-on-one therapy.” Participants reported decreasing app use over time, but there was limited evidence to suggest that this was because they were incorporating skills they had learned from the app into their daily lives without needing to use the app. One participant reported that they “remembered [the breathing exercises] for later on and that definitely helped when I didn’t have my phone on me.” This was the exception; two other participants only reported engaging with the exercises when using the app.

5.3.2.3 *Broader engagement*

These perspectives only represent a subset of participants (n=4). Of the 30 patients who received Wysa, 40% (12/30) did not engage with it at all (Table 5-4). Participants who engaged completed an average of 7 exercises, 24.3 sessions, and sent 227 messages to the Wysa bot.

Table 5-4. Engagement with Wysa app

Engaged with Wysa	Overall (n = 30)
No (%)	12 (40)
Yes (%)	18 (60)
Wysa: sessions completed	(n = 18)
Mean (SD)	24.3 (29.8)
Median [min, max]	11 (1, 91)
Wysa: messages to CA	(n = 18)
Mean (SD)	227 (319)
Median [min, max]	99 (20, 1180)
Wysa: exercises completed	(n = 9)
Mean (SD)	7.11 (7.79)
Median [min, max]	3 (1, 22)

5.3.3 Influence of previous cases studies on Case Study 2

Case Study 2 was the last case study to be completed chronologically, which meant that learnings from Case Studies 1 and 3 could be applied to its execution ([Sections 5.2.3](#) and [5.4.3](#)). Case study 2 built on the topic guide structure and thematic analysis codebook approach used in case study 3 by incorporating TFA constructs and the different components of engagement. The purpose of developing the topic guide based on these theoretical models was to generate qualitative data that reflected how user engagement with Wysa proceeded from initial awareness of the app through their use over time. Participant responses were coded inductively and categorised into the TFA and engagement component codebook structure. As previously highlighted, there is conceptual overlap between the constructs of the TFA and the different components of engagement, codes could be assigned to multiple categories (including a component of engagement and a TFA construct or multiple components of engagement or TFA constructs). This was done to avoid eliminating the complexity of users' responses by forcing them to fit within the limits of one component or construct.

5.3.4 Mapping of relationship between engagement components

The revised structure of the topic guide enabled an exploration of the factors that influenced initial behavioural engagement, which was not possible in the evaluation of NoObesity (case study 1). The first association was that positive affective engagement with the app appeared to drive initial behavioural engagement. When asked about their initial thoughts and feelings about the app and what motivated them to try it, participants recalled emotions including curiosity, excitement, and relief. To understand this association better, I further explored the context that participants provided when reporting these emotions. The opportunity to use the app evoked relief because it provided participants with a means of meeting an unmet need for mental health support. One participant reported that they were going through a breakup and experiencing a lot of emotions; another had just completed a course of CBT and was worried about the sudden lack of contact with mental health professionals. The app provided a source of mental health support that did not require them to feel like they were “*bothering*” someone.

On the other hand, excitement and curiosity reported by the participants were because the app was a new and interesting tool that they wanted to learn more about. As discussed in [Section 3.6.2](#), interest - and curiosity - have been variously conceptualised in the literature as emotions or elements of cognition [433–435]; a recent framework argues that they are constructs of a reward-based process of knowledge acquisition that includes affective and cognitive elements [434]. The analysis of the DBCI-ES in Chapter 3 highlighted how some models (including Perski et al. [42]) conceptualise cognitive and affective components of engagement together because of their conceptual overlap. The value of this approach is clear in these circumstances, where a concept includes cognitive and affective elements that cannot easily be untangled. However, Chapter 3 concluded that while distinguishing between cognitive and affective engagement may be somewhat artificial, their consideration as separate components can be valuable to explore how various concepts - like interest and curiosity - fit into the overall process of engagement.

By examining participants’ explanations of why they were motivated to use the app, we can see that there was both affective and cognitive engagement with the idea of the app (Figure 5-11, [Appendix 5-F](#)). The feelings and motivations expressed - curiosity, excitement, and relief - were linked to the judgments that participants made about the potential value and expected benefits of the app. From participant responses, it can be inferred that relief was experienced because the app was perceived as meeting an unmet need. Therefore, we can posit that initial behavioural engagement was influenced by cognitive engagement via affective engagement.

This pattern of engagement might also hold true for excitement and curiosity - experienced because the app was perceived to be worth directing attention towards - however, this was not as explicit from interviews.

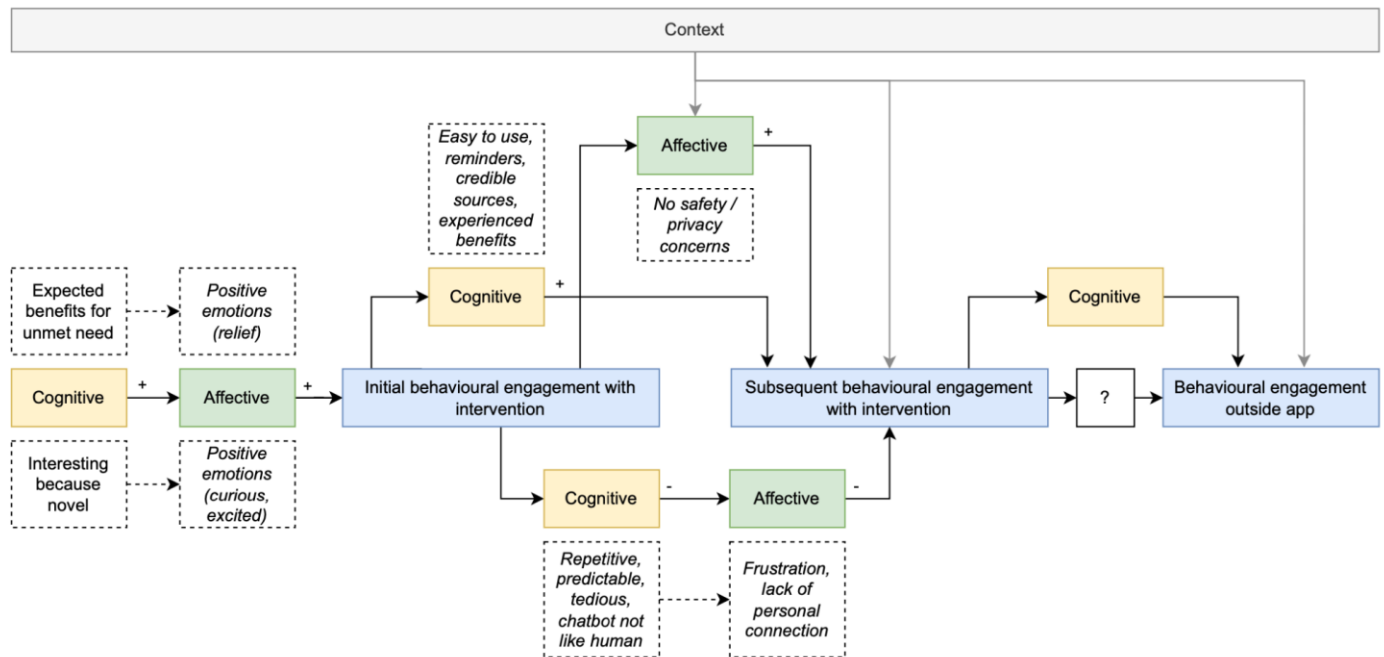


Figure 5-11. Mapping of engagement components derived from Wysa SSIs

Beyond the initial micro engagement, considering behavioural engagement was different for this case study compared to the previous case study (NoObesity). This is because the intervention design meant that behavioural engagement with the intervention and with the health behaviour overlapped. Wysa’s aim is to “[build] mental resilience and [promote] mental well-being” [358], which is achieved through conversations with the AI CA and engagement with the guided audios and videos, mindfulness exercises, and sleep stories. Key target behaviours of the intervention - eg. self-reflection, engaging in mindfulness - are behaviours that are conducted via the app. We can make a distinction between engaging with these behaviours via app and taking the skills learned through the app and applying the behaviours in daily life (without the app). In this case, however, participants did not report much use of the skills or strategies suggested by Wysa outside of the app. The mapping will therefore focus on users’ engagement with the app and with the intervention and target behaviours through their use of the app.

The AI CA - which was designed to be emotionally intelligent [358] - is a key element of the intervention’s strategies to support self-reflection. As one of Wysa’s central features, user engagement with the CA was an important part of their overall engagement with the intervention. As might be expected with a tool that is meant to emotionally connect with

users, the thematic analysis found that affective engagement with the CA was associated with subsequent behavioural engagement. Participants expressed frustration with the CA and a lack of personal connection with it. Their responses highlighted reasons for the negative affective engagement that were due to perceived limitations with the CA: users felt that it was repetitive, predictable, inconsistent, and not sufficiently advanced to feel like having a conversation with a human. Both cognitive and affective elements of engagement are reflected in users' frustration. Some participants felt frustrated when they invested effort in reflecting and typing out a long response to a question from Wysa and received a limited, templated response; as if the relationship was not sufficiently reciprocal to be emotionally rewarding and meaningful. From a cognitive perspective, it could be posited that the CA responses were not sufficiently interesting to justify the attentional effort users were putting into their inputs; effort that had to increase to overcome times when the CA did not understand the user or provided an inconsistent response. Although participants acknowledged that they weren't expecting the CA to have the same personal touch, depth, and empathy of responses as they would a human, these limitations made the CA tedious and frustrating to use. From this analysis, it appears that users' cognitive engagement with the CA was influencing their affective engagement, which in turn influenced their behavioural engagement with it.

Another concept worth considering in relation to users' engagement with the CA is that of the 'therapeutic alliance.' The therapeutic alliance is defined as the relationship between a patient and therapist that is shaped by the extent of shared goals (and defined steps to achieve them) and of a positive, genuine, emotional bond [436,437]. It has been positively associated with patient outcomes [436,437] and appears to share some commonalities with both cognitive and affective engagement. Users' therapeutic alliance with Wysa was recently evaluated in a separate study [438]. Its authors assessed users' therapeutic alliance twice (using the Working Alliance Inventory-Short Revised measure that includes three subscales measuring bond, goal, and task): once after 1-5 days of use and again 3 days after the initial assessment. They did not find a significant difference between the two assessments but noted that the mean score on the bond subscale were comparable to previous evaluations of other types of therapist-delivered CBT [438]. The study also conducted a content analysis of participants' conversations with Wysa, which identified that many participants expressed gratitude and positive impact and personified Wysa in their conversations. A smaller number of negative comments were also identified, which were similar to the feedback expressed in our case study - frustration with misunderstandings, repetitiveness, and limitations compared to a human [438]. The authors concluded that patients could form a therapeutic alliance with a

CA, particularly one that has the capability of interpreting free text responses and that participants generally felt comfortable and safe with Wysa [438].

Although none of the participants in our study expressed feeling uncomfortable or unsafe with Wysa, all four expressed some degree of frustration with the CA that affected their engagement with the intervention. It is possible this difference in findings is due to the small sample of participants interviewed in our case study. Alternatively, the therapeutic alliance study only took place over approximately 1 week - this might not be sufficiently long for participants to get frustrated with the CA's limitations. Additionally, there was a very high drop out from the therapeutic alliance study between the two assessments - the first assessment included 1205 participants while the second (non-mandatory) assessment only had 226 respondents [438]. Although the content analysis was conducted on available data from all participants who consented (n=950), participants who were frustrated with the app may have simply stopped using it rather than express their frustration to Wysa, which could explain why there was a much smaller proportion of negative comments identified.

Beyond the CA, there were a variety of factors that influenced users' behavioural engagement - both cognitive and affective. In terms of cognitive engagement, ease of use and reminders (low cognitive effort required) and positive perceptions of the app's credibility appeared to have a positive influence on behavioural engagement; however, users felt that the app was easy to ignore, which decreased their behavioural engagement. This was also reflected in users' affective engagement - one participant highlighted that there was no particular incentive to use the app, especially if they were in a bad mood. The tone of the app (including users' perceptions of the way the app communicated with them and the mascot design) was another factor influencing affective engagement. This suggests that improving affective and cognitive engagement could help improve behavioural engagement. There was also an influence of context on micro and macro behavioural engagement; participants highlighted factors such as user characteristics (eg. mental health issues, experience with therapy) and their schedules (eg. when they had a chance to use the app, and if that aligned with when it was most needed) as influences on behavioural engagement.

5.3.5 Recommendations in the context of existing evidence

Our finding that 40% of participants (12/30) did not engage with the app at all was in line with a recent meta-analysis of smartphone-based mental health interventions, which found that a mean of 41% of participants failed to download interventions targeting depression (8% for anxiety and 21% for general mental health)) [439]. This highlights the need for better

strategies for encouraging initial behavioural engagement with the intervention. Our analysis found that interest, curiosity, and expected benefits helped motivate users to try the app; advertising strategies could highlight these aspects to encourage users to engage.

Qualitative findings were also aligned with previous evidence suggesting that users have mixed opinions about CAs in healthcare [440] and observing an influence of the level of constraint on and repetitiveness of conversations on user experience [441,442]. The quality and capability of the CA was a major factor that influenced all of the components of engagement with Wysa. Limitations that could be addressed to improve this would be improving the depth and capacity of the CA, so that conversations are less repetitive, and improving model training to reduce inconsistencies in the CA's 'memory' and responses. A recent meta-analysis found that CAs that had the ability to respond empathetically and be personalised had significantly larger effect sizes in terms of efficacy than those that did not [443]. The challenge for the development of CAs for mental health will be how to support this emotional need while ensuring safety; providing mental health support is a complex and sensitive task and there is a risk of harm arising from using AI to provide support. The content library that Wysa's CA can draw responses from was designed and tested with human clinicians and is intentionally restricted in terms of potential responses to mitigate the risk of harm through AI hallucinations or inappropriate responses. Safety is a clear priority over engagement when it comes to the use of AI in healthcare, but both are important prerequisites for obtaining benefits from its use. Wysa's stated purpose is to provide evidence-based self-help tools and an opportunity for people to talk through their problems privately, not to diagnose or treat mental health conditions. However, the provision of guidance on standard therapeutic techniques and the conversationality of the CA may blur this line between digital and human therapeutic support, particularly for vulnerable users, which could have potential negative impacts such as withdrawal from social or clinical interactions [444]. An alternative could be to facilitate mental health support with clinicians via the digital tool [445], but this would reduce the value of using digital health tools to address the clinical capacity issue. This is an area in need of further research and development; the recommendation based on the findings can only be that conversationality is a key area influencing cognitive, affective, and behavioural engagement with mental health CAs and strategies for how to better support this within the bounds of safety will need to be considered.

Notifications did help remind users to engage with the app. This might especially benefit those users who incorporated the app at set points in their daily routines, but be less effective for those participants who used the app as needed or when they happened to have a quiet

moment. One participant suggested having notifications or widgets that were constantly present to remind users to engage when they had the chance and that were visually enticing, to provide an incentive for them to open the app.

5.3.6 Strengths and limitations

The in-depth examination of engagement and acceptability was one of the strengths of the study, but was limited by the number of participants. Of a planned 20 interview participants, only 4 completed interviews. Although there is no specific number of participants needed for qualitative studies, it is likely that a sample of 4 participants - even with different patient characteristics - was insufficient to achieve saturation and that additional opinions and insights could have been captured to improve the generalisability of the results. This was likely a symptom of the low sample size of the overall RCT, which also prevented the primary aim of determining whether the Wysa CA had a clinically-meaningful effect on symptoms of depression and anxiety. The recruitment period was significantly shorter than intended due to delays in approvals (4.5 months rather than 7) and monthly referral rates were lower than anticipated (~800 rather than 1100), with many patients being seen quickly and becoming ineligible. Of the ~800 monthly referrals, ~530 patients (65%) were screened and ~100 (12.5%) met the eligibility criteria. A subsequent analysis of the past year of waiting times at CNWL determined that around 95% of patients waited less than 29 days.

Options to increase recruitment included expanding eligibility criteria (removing PHQ-9 and GAD-7 cut-offs, which substantially reduced eligible participants), adding incentives, and adding sites. The first two options were not feasible - as PHQ-9 and GAD-7 scales are for screening (not diagnostic) purposes, increasing the eligibility criteria could change the study focus from mild-moderate to moderate-severe symptoms, unjustifiably increasing potential patient risk. Incentives have been associated with lower attrition [439,446] but their evidence in mental health research is mixed and risks exploiting a vulnerable population [447–449]. We were unable to use the third option (including additional Trusts or extending the study) due to funding restrictions. This is a common issue with digital health RCTs; one study found that only 40% of the NIHR-funded digital mental health trials investigated “met their original recruitment target in time” [450].

A strength of the qualitative analysis was that the SSIs were based on a synthesis of the TFA and a multifaceted conceptualisation of engagement. The topic guide was deliberately prepared to be able to address the thesis research question. It also provided an initial structure for the coding of qualitative data and facilitated the mapping of key themes to the components

of engagement and their associations. The thematic analysis was conducted by two authors, which strengthened the interpretation by providing alternative perspectives and challenging individual assumptions. I conducted the mapping of the engagement components alone as an additional analysis for this thesis; as such, it had the limitation of only including one researcher's interpretation of the data. This limitation was mitigated by using the definitions of engagement components in [Section 3.6.3](#) as a guide for associating the thematic analysis codes with engagement components ([Appendix 5-F](#)). To mitigate potential bias in the generation of the patterns of relationship between the components, the analysis was conducted twice with several intervening months. While this may not reduce bias as much as having independent researchers conduct the analysis, it did enable differences between the interpretations to be explored further before making final determinations.

5.3.7 Conclusions

The case study demonstrated that there is potential for a mobile CA to provide mental health support to supplement standard clinical treatment. It also provided evidence for a cognitive → affective → behavioural pattern of engagement both in terms of initial and subsequent behavioural engagement. Participants expressed relief at having a source of support while waiting for services; positive affective engagement appeared to be primarily (though not exclusively) related to positive interpretations of the app based on initial cognitive engagement with the information available before download. The content was generally found to be useful, though not sufficient to meet all mental health support needs. Key limitations with the CA's ability to provide a genuine sense of connection and hold in-depth conversations frustrated users and negatively influenced engagement. This appeared to be primarily driven by decreased cognitive engagement with the conversation, leading to negative affective engagement (frustration), and resulting in a greater likelihood of behavioural disengagement.

This case study was part of a larger RCT that experienced significant sample size limitations. Several lessons to improve future studies were learned from the execution challenges experienced, including the need to verify recruitment estimates with current data to ensure that the included sites have sufficient demand and to build in contingency to the study timeline and target recruitment to accommodate unexpected delays and higher than anticipated ineligibility or attrition. The issues experienced suggest that RCTs may not be the ideal study design for assessing DHIs, which require strong evidence from large populations in a timely manner to keep up with their development and implementation in healthcare. This thesis is providing hypotheses of patterns of association between the different components of

engagement; case study evaluations are a good study design for this stage, but future research will need to test these associations and strategies for supporting particular components of engagement. Are RCTs the best study design to answer these questions? This is a point of discussion within the digital health field that deserves further consideration [451–453]. Variations such as factorial RCTs have previously been used to evaluate the impact of individual features on engagement and impact [336], but future research should consider using mixed methods designs to capture context and facilitate real-world evaluation, which is critical for evaluations of engagement.

5.4 Case Study 3: Safety and Acceptability of an Autonomous Conversational Agent for Post-Surgical Follow-Up

*Parts of this section have been previously published in [345] (including all or parts of 5.4.2, 5.4.5-7)

5.4.1 Overview

Case study 3 is distinct from the first two case studies in the type of DHI that it evaluates. It was selected to provide a contrast to the mobile health apps examined in case studies 1 and 2, to assess how generalisable the generated patterns of engagement with mobile health apps are to DHIs more broadly (Figure 5-12). As an AI-based CA, the Dora post-surgical follow-up call requires only short-term engagement with the intervention and does not aim to provide health behaviour change support. The overall study aimed to evaluate Dora's safety and efficacy to assess its potential for adoption in the healthcare system. As patients' willingness to engage with the intervention is a key element of successful adoption, the study also aimed to explore Dora's usability and acceptability to patients, which is the relevant part of the overall study for this thesis and is presented below.

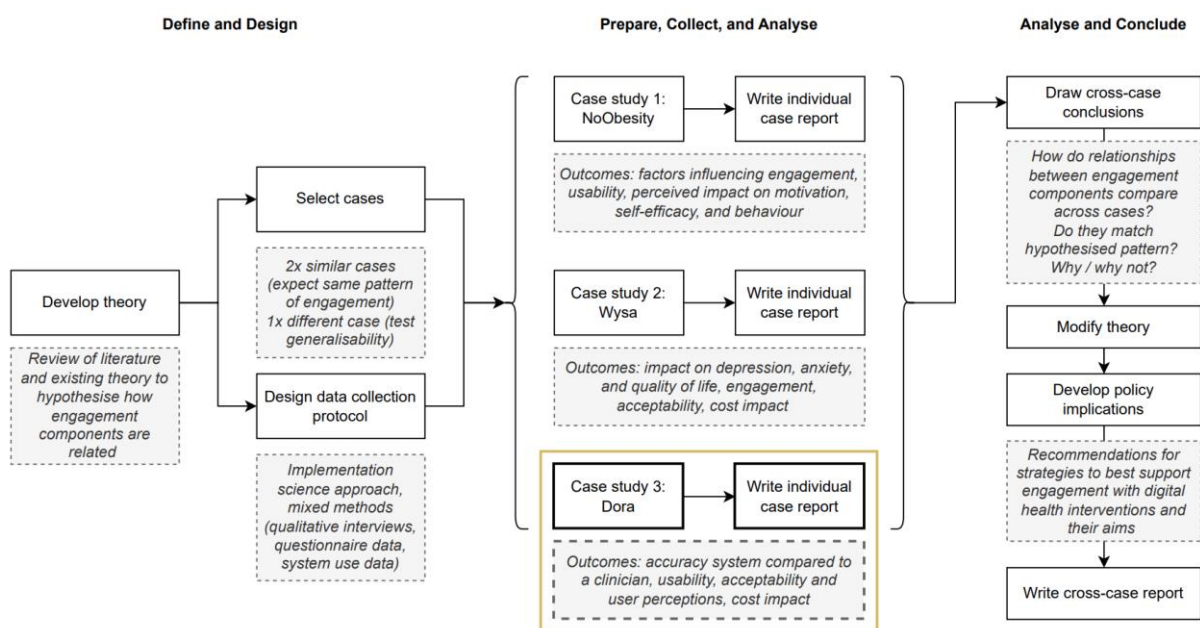


Figure 5-12. Case study 3 setting amongst the overall project

5.4.2 Results

5.4.2.1 Overview

Between 17 September 2021 and 31 January 2022, 767 patients were identified as eligible and 225 (30%) provided informed consent. 180 participants completed post-Dora call research surveys and 20 participated in SSIs (Figure 5-13, [Appendix 5-G](#)).

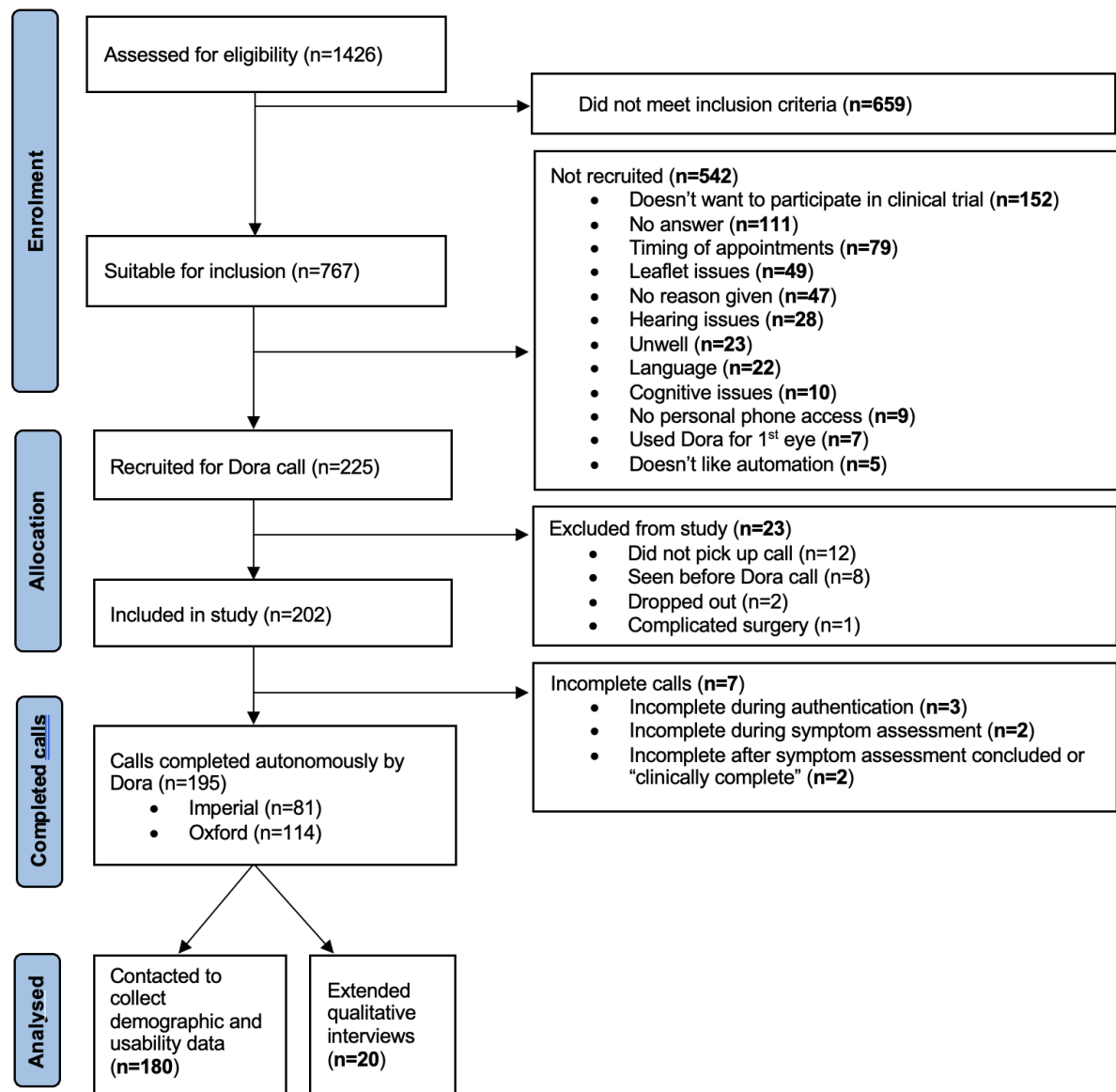


Figure 5-13. Case study 3 participant flow diagram

5.4.2.2 Participant characteristics

Demographic characteristics are summarised in Table 5-5.

Table 5-5. Participant demographic information

Demographics	Oxford	Imperial	Overall	Interviewed
Age ^a	n=119	n=83	n=202	n=20
Mean (SD)	75 (9.51)	72 (7.96)	74 (8.96)	74 (8.42)
Gender ^a	n=119	n=83	n=202	n=20
Male (%)	52 (44)	34 (41)	86 (43)	10 (50)
Female (%)	67 (56)	49 (59)	116 (57)	10 (50)
Ethnicity ^b	n=119	n=81	n=200	n=20
White (%)	107 (90)	53 (65)	160 (80)	13 (65)
Asian (%)	3 (3)	9 (11)	12 (6)	1 (5)
Black (%)	0 (0)	11 (14)	11 (6)	5 (20)
Other (%)	9 (8)	8 (10)	17 (9)	1 (5)
Income ^c	n=88	n=59	n=147	n=20
≤ £19,999/yr (%)	16 (18)	12 (20)	28 (19)	2 (10)
£20,000-29,999/yr (%)	13 (15)	11 (19)	24 (16)	2 (10)
£30,000-39,999/yr (%)	3 (3)	2 (3)	5 (3)	2 (10)
£40,000-49,999/yr (%)	3 (3)	2 (3)	5 (3)	2 (10)
£50,000-69,999/yr (%)	6 (7)	3 (5)	9 (6)	3 (15)
>£70,000/yr (%)	3 (3)	3 (5)	6 (4)	2 (10)
Undefined (%)	44 (50)	26 (44)	73 (50)	7 (35)
Education level ^c	n=81	n=50	n=131	n=20
Lower Than Bachelor's Degree (%)	52 (64)	36 (72)	88 (67)	9 (45)
Bachelor's Degree or Higher (%)	29 (36)	14 (28)	43 (33)	11 (55)

^aData retrieved from patients' electronic health record (EHR)

^bData retrieved from patients' EHR and supplemented with self-report questionnaire data if it provided more detailed information

^cData collected from self-report questionnaire

5.4.2.3 Overview of thematic analysis

The initial independent thematic framework for the analysis of Ufonia's acceptability and usability included 10 themes ([Appendix 5-H](#)). As in case study 2, the thematic map (Figure 5-14) was the product of collaboration between two researchers synthesising independent frameworks to present a clear overview of main themes (and subthemes) around acceptability and use of Dora (Figure 5-14). These themes are described in the following subsections incorporating participant quotes; additional example quotes are included in [Appendix 5-I](#). The results are presented in text based on the collaborative final framework, as this provides a higher level synthesis of findings and key points. [Appendix 5-I](#) uses the initial thematic framework to take advantage of the greater detail for the pattern of engagement mapping analysis.

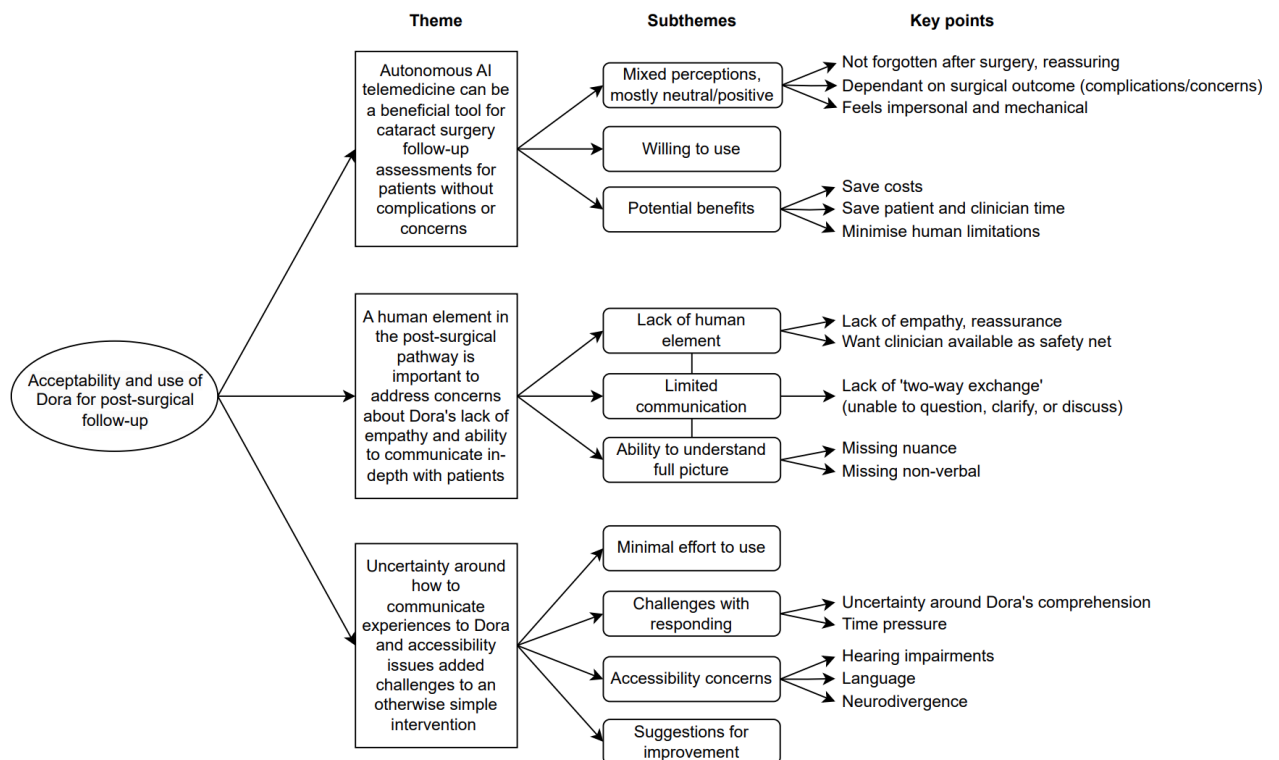


Figure 5-14. Thematic map of qualitative interviews to determine acceptability of Dora calls. (Note: the thematic mapping involved two researchers; Figure 5-14 represents the final collaborative version of the thematic map, the initial independent thematic map is included in [Appendix 5-H](#))

5.4.2.3.1 Theme 1: Beneficial tool for patients without complications

The first theme related to the general acceptability of the intervention as a tool with potential benefits for patients who had not experienced any complications with their surgery and did not have any concerns. Before the call, a couple participants worried it might be frustrating

and a few were interested, but most had no expectations, indicating that there was not a strong degree of affective or cognitive engagement prior to initial behavioural engagement for most participants. After the call, most interviewees' attitudes were neutral to positive; some found it "*fine*," while others felt "*comfortable*", appreciated the straight-forward approach, and "*enjoyed talking with Dora*." Many were confident in Dora's ability, with one caveat: around half the interviewees had concerns about how well Dora would work for patients with complications, particularly the AI's ability to handle emotional aspects or an unexpected surgical outcome.

Despite these concerns (discussed in Themes 2 and 3), many participants felt that Dora could benefit patients and clinicians by increasing convenience, saving time and costs, and providing reassurance by ensuring that all patients receive follow-up. A few participants even highlighted areas where Dora may be superior to a human: it "*gave [them] time to take a breath*" without being "*conscious that [the doctor is] busy and you don't want to look stupid*" and it wouldn't have "*off-days*." This aligns with quantitative data collected during the call about how likely patients would be to recommend Dora (Net Promoter Score, NPS [386]). The mean individual response was 8.59 / 10 ($SD = 2.05$) and the overall NPS score, which can range from -100 to 100 [386], was 51.06. Scores over 50 are generally considered to be above average [454].

Interviewees expressed approximately equal preferences for F2F or Dora consultations, with slightly more favouring F2F. Some indicated that it would depend on the situation or that they would prefer a F2F consultation but "*had no objection to [Dora]*." There was almost universal willingness to use Dora again; only one participant of the 20 interviewed was steadfastly against it, although a few others expressed some reluctance: "*well, if I have to use it, I have to use it*."

5.4.2.3.2 Theme 2: Importance of human element

The second main theme revolved around concerns about Dora's ability to manage complicated or emotional situations and the importance of the human element in care. Affective engagement with Dora was limited, leaning towards negative; many patients found Dora impersonal, with some saying that they felt "*a bit remote and lonely talking to a machine*" and others disliking the "*mechanical voice*". Some participants felt restricted by Dora's questions and the lack of a "*two-way exchange*" through which patients could ask questions or expand on their answers. One of the biggest concerns for Dora compared to a clinician was the potential to miss nuances and non-verbal information: "*closed questions...*

will only get a certain answer. That may not be... the whole picture, but a computer won't pick up on that." Many participants were "happy to talk to AI, if there's nothing wrong," but "would not feel reassured" if they had concerns. Several stressed the importance of "a two-step process" whereby patients would be called by Dora but "if [they] have any concerns at all, [be put] through to somebody [they] can speak to."

5.4.2.3.3 Theme 3: Simple intervention with some usability and accessibility challenges

The third theme focused on usability and accessibility. Almost all interviewees (mean age = 74 years, SD = 8.42, range: 56-90 years) found Dora easy to use and that "[they] could understand everything it was saying". This was in line with the quantitative assessment, which found a mean SUS score of 77.76/100 (SD = 17.55) and average TUQ item score of 3.79/5 (SD = 0.89). An SUS score of 77.76 equates to a 'B' on the Sauro-Lewis curved grading scale and is considered relatively good [384]. Cognitive engagement was important for successful behavioural engagement; usability depended to some degree on users' efforts to provide simple answers to avoid confusing Dora and having to revise their answers if Dora did not understand them. Potential accessibility barriers such as hearing, understanding Dora's accent, and "older people... [finding] it quite difficult to communicate with a computer" were raised. The intervention was also potentially unsuitable for people with autism; one participant with autism "was extremely distressed by the whole experience" because they "find it ... difficult to process information [without] the opportunity to ask relevant questions" and were not confident in Dora's "ability to identify... a problem that needs further investigation." Another participant felt their brother, who was "dyslexic [and] mildly autistic" would struggle with Dora. To improve usability and accessibility, participants suggested further training so Dora can better understand more detailed responses, clarifying the objective of the call, providing questions in advance or an alternative paper version, and facilitating scheduling and reminders.

5.4.3 Influence of Case Study 3 on the thesis

Case study 3 was the second case study to be conducted, chronologically. Based on the objectives for the evaluation of Dora (which focused on accuracy, safety, usability, and acceptability) and the learnings from case study 1 that highlighted the limitations of conducting SSIs without a theoretical framework to ensure that the relevant data was being collected, the topic guide for Dora was developed based on the TFA. This benefitted the analysis by capturing various elements of acceptability, but I found that it was still somewhat limited in terms of capturing detailed and specific data about how users engaged with the

intervention. This may have been related to the type of intervention - Dora did not require prolonged engagement - but led me to include specific questions shaped to capture the different components of engagement as well as the TFA in the final case study to be conducted (Wysa). This helped to facilitate a more comprehensive discussion of how users' engaged with Wysa and what factors influenced that engagement.

5.4.4 Mapping of relationship between engagement components

As in the first two case studies, the thematic analysis was revisited to generate hypotheses for potential associations between components and patterns of engagement based on the interview data ([Appendix 5-I](#)). The following section discusses these various patterns of interaction, which are summarised in Figure 5-15. Although participants were asked about what they expected talking to Dora to be like before they received the call, it was difficult to identify important effects of cognitive or affective engagement from the responses. Most participants reported lacking any expectations of the call; of the ones who did, both positive and negative affective expectations were expressed. For example, some participants felt regret about not being able to speak to a person or worry that they would get frustrated with the call, but one participant felt positively about the call because they expected that it would provide a source of information (affective engagement tied to cost-benefit judgments, an aspect of cognitive engagement) and another was interested in seeing what Dora was capable of doing (cognitive engagement). This suggests that both cognitive and affective engagement could influence initial behavioural engagement with the intervention. Although it was limited to one participant, it was interesting that positive affective engagement with the idea of the intervention was reported as a result of the cognitive engagement (expectation of benefit from the call), as this was the pattern generated in case study 2, but is opposite to what the theoretical model predicted based on the literature in [Section 3.8](#). This difference will be discussed further in the cross-case synthesis in Chapter 6.

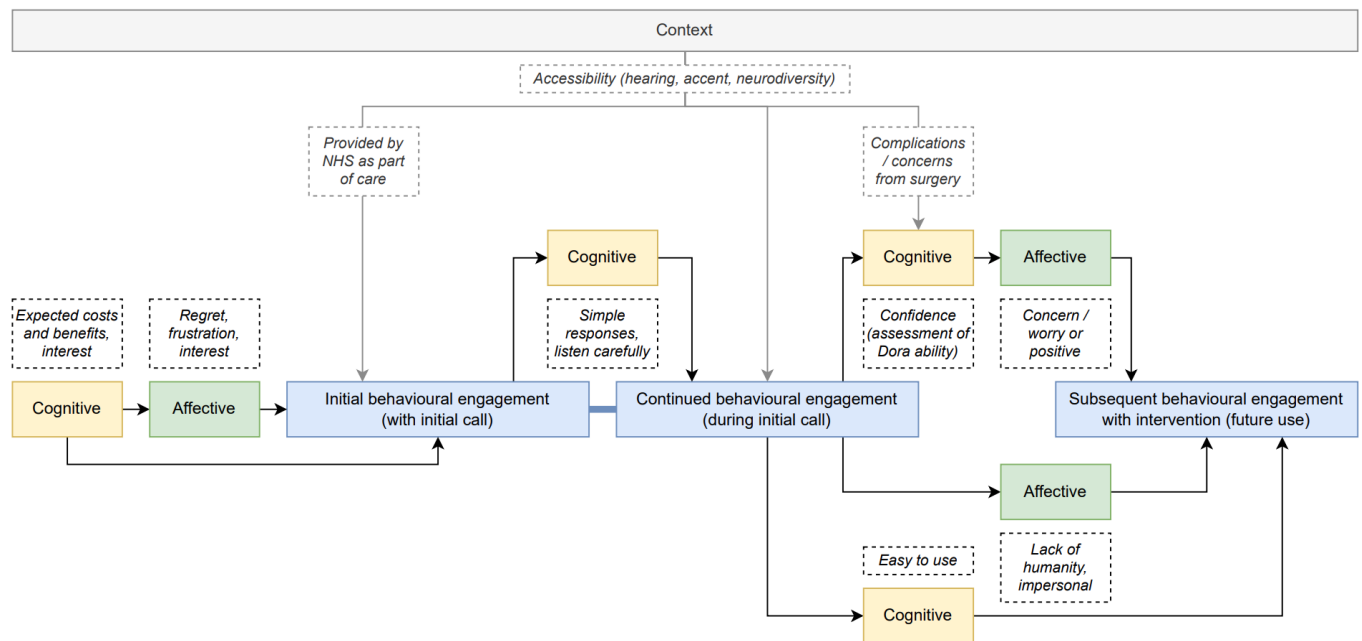


Figure 5-15. Mapping of engagement components derived from Dora SSIs

Although this suggests that cognitive and affective engagement can influence initial behavioural engagement, the fact that most of the participants did not have strong opinions or expectations about the intervention before its use suggests that contextual factors, rather than cognitive or affective engagement, likely had the strongest influence on initial behavioural engagement. A key contextual factor can be inferred from the delivery of the intervention and is highlighted by one participant's response to the question about expectations: "*I'm taking the position that anybody from the NHS contacting me with a thing like this is starting from a high bar. So I would have no reason to be suspicious of it really.*" Social trust was one of the affective factors highlighted as having a potential influence on acceptability and behaviour in Liu *et al.*'s examination of acceptance of self-driving vehicles ([Section 3.7](#)) [62]. In their study, Liu *et al.* define social trust as users' trust in the organisations or people who are responsible for the technology and cite previous literature that indicates that this trust can act as a heuristic to simplify the process of assessing potential risks and benefits and making a judgement about how they feel about the technology [62]. There is not sufficient data from the interviews to determine if trust in technologies provided via the NHS was a major factor influencing participants' decisions to engage with the Dora call, but it would help explain why most participants did not have any particular expectations of the call. Participants were recruited by NHS Research Nurses in clinical settings - although participants were all provided with information explaining that the purpose of the study was to evaluate the safety and efficacy of the technology, the fact that it was being assessed in the context of standard healthcare delivery may have engendered some level of trust in the technology. Additionally,

patients at Oxford are not routinely booked for postoperative follow up assessment, so the option to use Dora may have been considered as a bonus or better than nothing. At Imperial, patients are routinely booked for follow-up appointments, which took place shortly after the Dora call, so - as the call was brief and could be completed at home - they may have felt like they had little to lose by participating, as they would be seen in a face-to-face follow-up appointment anyway. These reasons are all speculative, as the study did not capture reasons for choosing to participate in the study; the purpose of including them here is to illustrate how various potential contextual factors could play a role in users' initial behavioural engagement with the intervention even in the absence of affective or cognitive engagement.

Another contextual factor that could affect engagement was whether the patient had any concerns or complications. In the majority of cases, where there were no issues with the surgery or recovery, several interviewees thought Dora was "*more than adequate*", and in some cases even "*superb*" because it saves time when everything is fine. Participants expected concerns to influence their engagement with Dora; many were less confident in its ability to conduct the assessment in cases of complications. Both cognitive and affective engagement with the intervention shaped this perception. Through their experiences talking to Dora, participants made judgments about how and what they could say to Dora, with some noting that they deliberately "*kept [their] answers simple*." Others were concerned that Dora's inability to enable the flexible provision of further contextual information about their answers might result in Dora missing the nuances or not being able to capture their full experience. Dora's reported "*impersonality*" and lack of "*humanity*" could also influence patients' affective engagement with Dora, especially in cases where they had complications, were worried or upset, and wanted reassurance. Although participants had different affective engagement with Dora - some participants felt positively about the conversation and even "*enjoyed talking with Dora*" - it is clear that user's affective engagement influenced their overall perceptions of Dora and their willingness to engage with Dora in the future.

Beyond the immediate affective experience of engaging with the intervention, participants were split on whether the intervention aligned with their values - some reported having no concerns or issues with talking to Dora, while others did not like talking to machines and felt that this was not the best direction to be moving in to deal with patients. Cognitive engagement with the intervention did "*not [take] much effort*" for most participants and, because the call was brief, no costs of engaging with the intervention were mentioned other than the lack of a human element and the need to consider how to talk to Dora. Despite the potential concerns raised, most participants reported being willing to engage with Dora again.

Many stated that they would prefer face-to-face appointments, suggesting that - at least for some - future behavioural engagement with Dora would depend on contextual factors such as whether they are given an option and how they are feeling after the surgery as well as their experiences with their initial engagement.

5.4.5 Strengths and limitations

The overall study findings provided preliminary evidence that Dora is a safe, generally acceptable, and potentially cost-effective means of delivering follow-up assessments after cataract surgeries. The qualitative analysis described in this case study was a strength because it highlighted certain barriers that will need to be addressed to successfully implement technologies such as Dora in cataract surgery follow-up and other clinical pathways.

Limitations included potential issues around selection bias. The two clinical sites were located in a similar area in the south of England. Although the Imperial site serves a diverse population, these sites are unlikely to have captured the variety of accents and other demographic characteristics present in the UK, which may influence successful use of Dora. The qualitative analysis purposefully aimed to include diverse participants: 20 interviews were conducted with participants of different ages, genders, ethnicities, income levels, and education (Table 5-5). Including participants with different characteristics and backgrounds helped to improve confidence that a diverse set of perspectives were being captured in the interviews. Despite these efforts to mitigate selection bias in the interviews, patients with strong opposition to AI telemedicine may not have consented to participate in the overall study, so the results may reflect a more positive response than could be expected if Dora was implemented as a standard part of clinical service. Likewise, participants with cognitive difficulties, hearing impairment, or a lack of fluency in English were excluded from the study, reducing generalisability. Several participants felt that - if they had experienced complications - they would be less comfortable using Dora, but as cataract surgery has a high level of success and safety, most interviewees did not experience complications. This meant that the feedback around the use of Dora in the case of complications was largely hypothetical; the analysis would have been strengthened if it had also captured the experience of patients who did have complications.

A strength of the thematic analysis was that the initial coding and thematic analysis was conducted by two researchers. This enabled independent and collaborative reflection on and interpretation of the data, resulting in the final thematic map included in this case study (Figure 5-14). The subsequent analysis of the mapping of engagement components was

conducted specifically for this thesis by one researcher. This introduces potential bias in the interpretation. As with the other case studies, I made an effort to mitigate this by using the same definitions of the engagement components ([Section 3.6.3](#)) to guide the analysis and by revisiting and reanalysing the data several months after the initial analysis with fresh eyes. This helped strengthen the findings by enabling me to compare my two analyses and question and resolve any differences with reference to the data.

Another potential limitation of the engagement mapping analysis in this case study is that the type and purpose of DHI was distinct from case studies 1 and 2, particularly because it did not involve any subsequent macro engagement with a health behaviour. The implications of this for the comparability of the case studies will be discussed in Chapter 6.

5.4.6 Recommendations in the context of the literature

Despite most patients being willing to use Dora, at least when they had no significant concerns or complications, they provided suggestions for improvement. Although the call was generally considered easy to use, patients' uncertainty about Dora's ability to understand what they say and the time pressure of answering questions live increased the difficulty of responding to Dora's questions. A suggestion from participants that could potentially reduce this burden would be to provide the questions in advance. Providing additional information and clarity about the purpose of the call and what to expect from it could also help improve cognitive and affective engagement by reducing uncertainty and providing reassurance about the continued involvement of human clinicians in the pathway. Participants also suggested continuing the training and improvement of the CA; the impersonality and mechanical sound of the call and concerns about its ability to understand more subtle issues hindered affective and cognitive engagement. Contextual factors that could be modified to improve engagement included adding reminders and the ability to schedule or reschedule calls and increasing accessibility by offering the call in other languages and via a paper version for people with hearing difficulties or those who did not want to use the CA.

A recent integrative review examined CAs in healthcare using the Unified Theory of Acceptance and Use of Technology to assess factors that could influence acceptability and adoption [455]. Factors identified were largely in line with this analysis, such as "performance expectancy," "perceived risk" (expected benefits vs. costs), and "effort expectancy" (perceived ease of use) [455]. Additionally, the ability of the CA to enable human-type interaction (anthropomorphism) was generally found to have a positive influence on acceptability and continued use, although having the CA be too life-like could cause

negative emotional reactions for some users (i.e. “uncanny valley”) [455]. To address this, they recommended CAs clearly identify themselves as a machine [455]. Further trust could be built by providing simple explanations of the system with information about prior evaluation, performance, reliability, and data security and privacy to address users’ potential concerns [455]. Another key factor, especially in healthcare, was the ability to provide free-text, rather than predetermined, answers [455] - this paralleled our participants highlighting the desire to provide further detail and context around their responses to ensure accurate assessment. Developing the language skills of CAs to better comprehend such responses will likely improve affective and cognitive engagement; alternatively, in the case of a CA such as Dora, patients could be given the option during the interaction to flag that they have concerns that have not been addressed and provide a verbal summary that could be automatically transcribed and sent to a clinician for review.

The thematic analysis emphasised that a key factor related to patients’ attitudes towards Dora was the outcome of their surgery, which was also identified in Wutz *et al.*’s review [455]. It will be important to ensure that the introduction of automation into healthcare does not result in the dehumanisation of care by removing patient-clinician interaction, but rather increases the time available for clinicians to interact with patients who have complications [456,457]. Studies implementing similar systems in other contexts should consider the likelihood of complications and patients’ post-surgical emotional experience; mitigations such as more patient information, a clear procedure for what to do if concerned, and tailored call timelines might be necessary for acceptability and engagement.

The study also highlighted the importance of digital equity and accessibility considerations. Issues of comprehension and hearing are common for natural language technologies and implementation will require equitable alternatives for non-English speakers and people with serious hearing difficulties [458]. Another digital equity issue that arose was related to the acceptability of this technology for people with autism; a participant who self-identified as being autistic was deliberately included in the qualitative analysis to capture perceptions from a diverse sample. Although Dora was generally acceptable for most of the sample, it was very unacceptable for this participant. This case raises two key points: potential issues of acceptability for people with autism using automated CAs, which should be investigated further and accounted for in the implementation of autonomous telemedicine, and the issue of inclusivity in research. If the participant had not reached out to the research team to express their feedback, it is likely this patient would have been missed from the randomly-selected interview sample. This highlights the importance of making deliberate efforts to include

diverse perspectives and user needs through the digital health development and evaluation process [456] to ensure that diverse needs are understood and can be accounted for in the implementation of such technologies. For instance, digital equity and accessibility issues could be facilitated by including a brief question in pre-surgical assessments to determine whether patients are willing and able to receive the automated call and by delivering digital services in multiple languages.

5.4.7 Conclusions

The case study findings provided preliminary evidence of Dora's acceptability, engagement, and potential for adoption within clinical pathways. The analysis of users' engagement with Dora highlighted the importance of ensuring that there is still a point of human contact available for patients when digital technologies are incorporated into routine clinical pathways. Mixed-methods analyses were conducted to better understand issues that could affect the adoption and use of Dora in a real world context. Scores on the SUS and TUQ were above average, indicating an acceptable level of usability, and system data demonstrated that most calls were completed autonomously without issue. The quantitative data aligned with the qualitative interviews, where most participants found Dora relatively acceptable and easy to use; however, this depended on the context in which they were receiving the call. Having concerns or complications influenced users' affective and cognitive engagement with Dora and their confidence in Dora's ability to capture the nuances of their experience and to support their emotional need for reassurance. These findings have implications for Dora's adoption within cataract surgery and more broadly for future applications of AI-enabled telemedicine in healthcare. In terms of efficiency, Dora successfully recommended discharge for over half of patients (58%) [345]; in real-world operation, these patients would not be seen for an in-person follow-up assessment or for further clinical review, reducing clinical follow-up workload and costs of service delivery but also reducing routine patient-clinician contact. The analysis also emphasised the importance of clear patient information regarding the role of Dora and human clinicians within their overall care pathway (how to seek clinician-delivered care if needed) to support positive affective and cognitive engagement with the intervention. More broadly, our findings highlight the potential for automated systems to support other repetitive, low-skill healthcare tasks and reduce burden on healthcare professionals but emphasise the importance of considering contextual factors when implementing the intervention - in cataract surgery and other contexts - to support engagement and ensure that all patients receive equitable, humanising care.

5.5 Summary of Within-Case Analyses

Chapter 5 has presented the results and analyses of the evaluations of three different DHIs, with a focus on users' experiences of engagement. This section will outline the value of these three case studies for addressing the overall research question and aims of the thesis. The thesis asked "*How are the different components of engagement (affective, cognitive, behavioural, micro, and macro) related in digital behaviour change app interventions for weight management and mental health and in an autonomous telemedicine post-surgical follow-up assessment?*" In addressing this research question, it aims to 1) **map the different components of users' engagement** with digital health interventions and how they influence each other, 2) **compare engagement with different types of digital health interventions** to determine if identified processes of engagement are generalisable, and 3) generate **evidence-based recommendations for strategies** to optimise engagement at different stages.

The three case studies presented here were selected for inclusion to help address aim 2 - they represent distinct DHIs with different characteristics and clinical contexts. NoObesity and Wysa are both mobile health apps, but have different means of addressing particular health needs. NoObesity focuses on supporting goal setting, education, and health behaviour change - particularly external behaviours of physical activity and healthy eating. In contrast, Wysa aims to provide mental health support through therapeutic exercises and conversations with a CA in the app. The third intervention, Dora, is also a CA but with a very different purpose and mode of delivery. Dora provides a brief, post-surgical follow-up assessment conducted over a regular telephone call; unlike NoObesity and Wysa, there is no behaviour change element or need for sustained engagement. The benefit of including such distinct DHIs is that the cross-case synthesis can explore the similarities and differences between the findings and how those might be related to the specific characteristics or aims of the interventions. The analysis will support the potential generalisability of the final hypothesised framework of how different components of engagement influence each other over the process of engagement with different DHIs. The potential limitation of such distinct cases is whether they are sufficiently similar to be meaningfully synthesised; this will be examined in detail in [Section 6.3](#).

The initial, within-case analyses have already generated useful insights into engagement with these different types of DHIs. Case study 1 (NoObesity) found that both affective and cognitive engagement could influence each other and behavioural engagement over the course of engagement with the intervention and the target health behaviours. Although both were important in this overall process, affective and cognitive engagement were closely linked and

it was difficult to provide evidence for a particular order of their influence from the available data. The case study also highlighted the influence of context on micro and macro behavioural engagement (and on affective engagement with the health behaviour). Key recommendations for supporting engagement in this type of DHI were to design the app so that it could provide novelty, interactivity, and feedback; support goal setting, self-monitoring, and accountability; and prompt engagement through notifications and a visually appealing, fun, and positive interface.

Case study 2 (Wysa) also provided evidence of the critical role of cognitive and affective engagement in behaviour. From users' descriptions of their experiences of engaging with the intervention, the pattern generated was that of cognitive engagement influencing affective engagement, which then influenced behavioural engagement. This pattern occurred with both initial and subsequent engagement with the intervention, although cognitive and affective engagement also appeared to be independently associated with behavioural engagement. A challenge with case study 2 was that the boundary between micro and macro engagement was blurred because engagement with the behaviour was facilitated by the app. As such, this is an example of what may be a good case for the use of the TWEETS measure of engagement, which explores engagement with the intervention in the context of achieving a goal ([Section 3.5.3](#)). The main factor influencing engagement was the conversational ability of the CA; however, recommendations for modifying this to improve engagement risks increasing the potential for harm, an unacceptable alternative. Further research will be needed to find a means of better supporting engagement with CAs while maintaining strict safety bounds.

Affective and cognitive engagement were also found to have potential associations with behavioural engagement in the case study 3 (Dora); however, compared to the first two cases, there was a particularly strong influence of context on behavioural engagement. Despite this, a similar pattern of influence from cognitive → affective → behavioural engagement was generated from the data. As with Wysa, the communication ability of the CA was a particular point of concern, especially for patients who had not experienced uncomplicated post-surgical recoveries. Other recommendations to support engagement included providing sufficient and simple information to alleviate user concerns about the system, ensuring the role of human clinicians in the system was clear, and addressing digital equity and accessibility issues in its implementation.

Chapter 6 will explore how these three different cases can be synthesised to generate an evidence-based hypothesis for a model of associations between the different components of engagement (aim 1) that is potentially transferable to other types of DHIs (aim 2). Achieving

these aims will provide a framework that can guide further research and testing of these proposed relationships. Chapter 6 will also explore potential key, transferable factors influencing engagement, which will support the development of evidence-based recommendations ([Section 7.4](#)) for strategies to incorporate in DHIs to support these different components throughout the process of engagement (aim 3).

Chapter 6: Cross-Case Synthesis

6.1 Chapter Overview

Chapter 5 provided the relevant results and analyses of three case studies that evaluated three distinct DHIs (a childhood weight management behaviour change app, a CA mental health self-help app, and a CA post-surgical follow-up assessment telephone call). For each case, Chapter 5 provided a summary of patient characteristics, presented the primary thematic analysis executed to address each case studies' own objectives, and re-analysed the qualitative data to develop a mapping of associations between different components over the process of engagement.

Chapter 6 presents the results of the cross-case analysis and thematic synthesis and aims to address the main research question of the thesis: *How are the different components of engagement (affective, cognitive, behavioural, micro, and macro) related in DHIs?* This question has been examined in each of the case studies to generate empirical models mapping engagement with each DHI. This chapter compares these mappings to each other and the hypothesised model that was generated in Chapter 3. The comparison interprets the similarities and differences between these mappings in the context of the cases they derived from and explores potential explanations for those similarities and differences. A refined model mapping the relationships between the components of engagement with DHIs was developed from this synthesis [337,459].

The chapter starts with a brief summary of the findings from each within-case analysis and an exploration of the comparability of the cases. As each of the cases examines a different type of DHI, with a different clinical aim and application, determining the validity of comparison is an essential prerequisite to the cross-case synthesis. It also provides context for the interpretation of the individual case study findings, which is a key element of case study research design. Chapter 6 then presents the results of the pattern-matching analysis of the mappings of engagement and the thematic synthesis of the case studies. The thematic synthesis explores factors that can influence the various components of engagement and generates insights to inform recommendations for strategies to support engagement with DHIs. The chapter concludes with a discussion of the principal findings, strengths and limitations of the cross-case synthesis, and a summary of the revised model for the associations between the components of engagement.

Publications: Chapter 6 synthesised data and analyses that have previously been published, but none of the analysis conducted in this chapter had been published when the thesis was

submitted; see [Section 5.1](#) for the references and author contributions for the results synthesised in this chapter. Sections of Chapter 6 have since been submitted for publication; references and author contributions for these papers are listed below.

Milne-Ives M, Homer S, Andrade J, Meinert E. Mapping the process of engagement with digital health interventions: A cross-case synthesis. *Mayo Clinic Proceedings: Innovations, Quality & Outcomes*. 2025; Under Review.

MMI conceived the topic, conducted the analysis, and drafted and revised the manuscript. SH, JA, and EM supervised the research and contributed revisions.

Milne-Ives M, Homer S, Andrade J, Meinert E. Exploring factors influencing engagement across different digital health interventions: thematic synthesis. *Nature Scientific Reports*. 2025; Under Review.

MMI conceived the topic, conducted the analysis, and drafted and revised the manuscript. SH, JA, and EM supervised the research and contributed revisions.

6.2 Summary of Individual Case Findings

[Section 5.5](#) summarised findings from the within-case analysis; they are included here briefly to recap study characteristics, key findings, and mappings of engagement components before delving into the cross-case synthesis.

6.2.1 Case Study 1: NoObesity

The first case study examined ‘NoObesity’, a mobile health app intended to prevent and mitigate childhood obesity [77]. The app was developed to achieve two main aims: 1) to provide support for weight-related behaviour change for families and 2) to deliver training to improve HCPs’ skills and confidence in communicating with families about childhood obesity concerns. The case study explored factors associated with users’ engagement with the app over a 6-month intervention period. Key facilitators included novelty, interactivity, supported goal setting and self-monitoring, positively-framed content and feedback, and reminders. Boredom and forgetting to use the app were key barriers, as were contextual factors that shaped families’ opportunities, motivation, and ability to engage with the behaviour (e.g. weather, time, available resources, children’s age, etc.). Component mapping began after initial use of the app, as no data was captured about cognitive or affective engagement with the idea of the app before initial behavioural engagement. Qualitative data suggested that behavioural engagement with the intervention shaped affective and cognitive

engagement with it, which in turn influenced subsequent engagement with both the intervention and the target health behaviours.

6.2.2 Case Study 2: Wysa

The second case study evaluated ‘Wysa’, a CA app designed to provide mental health support [129]. Wysa aims to improve mental health through guided self-help exercises and self-reflection. The primary objective of the overall RCT was to evaluate Wysa’s efficacy at improving symptoms of depression and anxiety compared to a waitlist control, but low sample size prevented a statistical determination of whether Wysa clinically-meaningful effect. This analysis focused on acceptability and engagement with the app over the 3-month intervention period. There was overlap between micro and macro engagement because engagement with the target behaviours primarily occurred while users’ were using the app. Positive cognitive engagement with the idea of the app was associated with positive affective engagement and initial behavioural engagement. Subsequent behavioural engagement (with the app and behaviour) was also influenced by cognitive engagement via affective engagement, although both cognitive and affective components could also influence subsequent micro behavioural engagement independently. A key factor influencing overall engagement with Wysa was limitations in the conversational capacity of the CA, which caused frustration and reduced motivation to use the app. The variety of content (in terms of therapeutic exercises) influenced cognitive engagement; it was viewed as helpful, but overwhelming to navigate. There was little evidence to suggest that engagement with the intervention supported engagement with the therapeutic techniques outside of the app.

6.2.3 Case Study 3: Dora

The third case study examined ‘Dora’, an autonomous telephone CA for conducting post-cataract surgery follow-up assessments [345]. The evaluation aimed to assess Dora’s accuracy and safety as well as its acceptability and usability for (primarily older) patients. Unlike NoObesity and Wysa, there was no target health behaviour, so the case study focused on micro engagement. Contextual factors (e.g. provision of the call via the NHS, presence of post-surgical complications) had an important influence on engagement. Although the call was generally perceived as acceptable and potentially beneficial, the lack of human element and concerns about the CA’s capacity to understand nuances in patient responses influenced affective and cognitive engagement. As with Wysa, the interpretation of the evidence indicated a pattern of cognitive engagement influencing behavioural engagement via affective

engagement, although cognitive and affective engagement could also influence behavioural engagement independently.

6.2.4 Within-case mappings of engagement components

The mappings of the engagement components developed for each case study in Chapter 5 are included to facilitate their comparison in advance of the pattern-matching analysis in [Section 6.4.1](#) (Figures 5-7, 5-11, and 5-15).

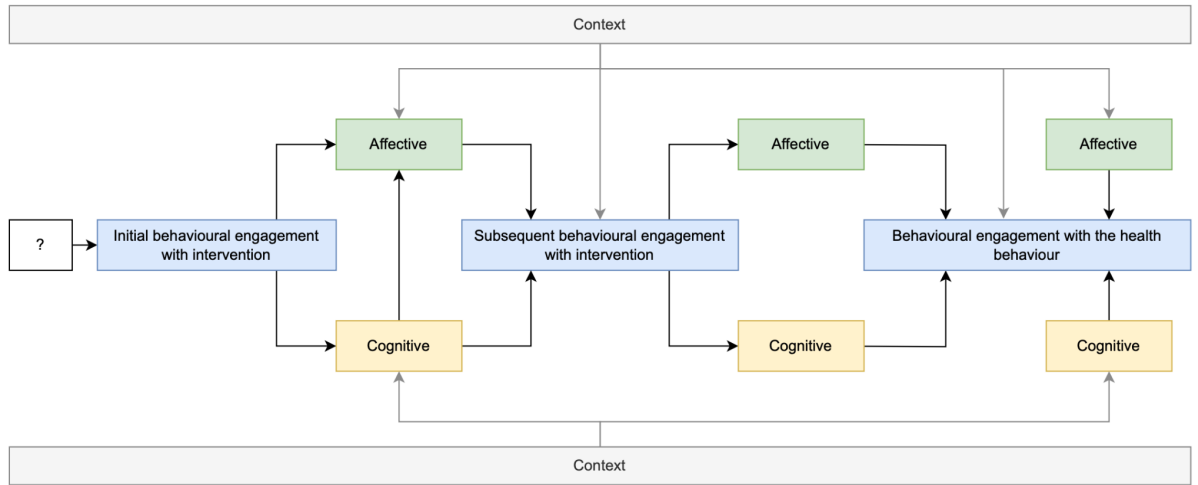


Figure 5-7 (duplicate). Mapping of engagement components derived from NoObesity SSIs

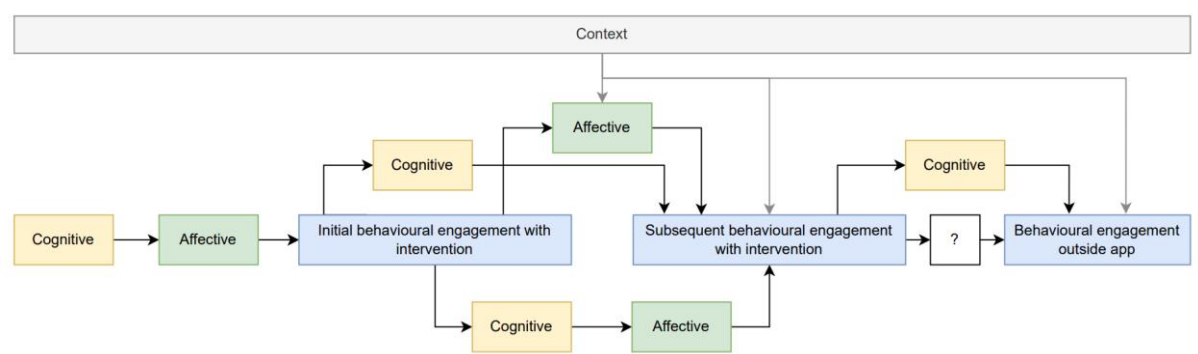


Figure 5-11 (simplified). Mapping of engagement components derived from Wysa SSIs

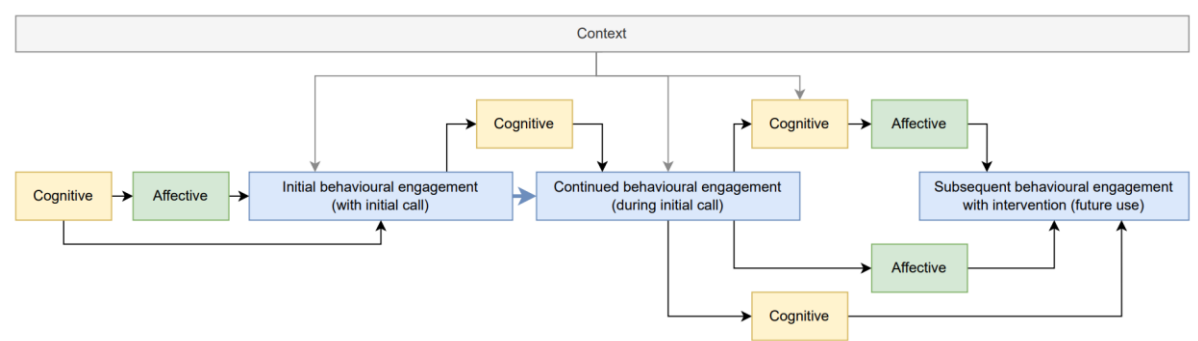


Figure 5-15 (simplified). Mapping of engagement components derived from Dora SSIs

6.3 Comparability of Cases

Although all case studies examined engagement with DHIs, their characteristics are different. This section will highlight key differences and argue why the cases are sufficiently comparable for the cross-case analysis to generate useful insights. The potential for these differences to undermine the cross-case findings and produce rival interpretations will be examined in detail throughout the chapter (e.g. [Sections 6.4.3](#) and [6.5](#)).

The first difference was the clinical applications: childhood obesity, mental health, and post-surgical follow-up. While the reason for using a DHI is likely to influence how users think and feel about the intervention and their engagement with it, this context would be expected to shape the individual influences of the components on each other, rather than the overarching pattern in how the components of engagement interact. For example, a person might find that their mental health has a more significant influence than their weight on their quality of life, and thus might have higher expectations of, and feel more motivated to engage with, a mental health app. In this case, they would likely have more positive cognitive and affective engagement with the idea of the intervention and be more likely to try it. On the other hand, they may not enjoy exercise and therefore feel reluctant to use an app that will try to increase their physical activity, even if they believe that it will be beneficial to them. These differences, however, would not be expected to change the pattern of influence, for example, of an assessment of costs and benefits (cognitive engagement) on users' feelings about the intervention (affective engagement) and their likelihood of trying it (behavioural engagement). This example demonstrates why, although the clinical use case is part of the context that shapes engagement, it should not influence the structure of how the components of engagement interact and thus does not present a problem for comparability.

Beyond clinical application, intervention type and aim are other key differences between cases: the first two examine mobile health apps for behaviour change over a prolonged period, while the third is an AI CA, delivered via telephone, for a brief clinical assessment. These cases were selected to enable two different types of replication logic: literal and theoretical ([Section 4.4.1](#)). Literal replication refers to similar cases predicting similar results while theoretical replication refers to different cases predicting different results, but in line with expected reasons based on their differences [337]. The first two cases are similar types of DHI, despite the different clinical application and target behaviours, so patterns of engagement should be largely similar; case 3 is a different type of DHI with a different aim - clinical assessment rather than behaviour change - so how users engage with Dora can be hypothesised to be different than with NoObesity or Wysa. This offers an opportunity to test

theoretical replication [337] - does the way the engagement components interact differ in the way that would be expected based on how Dora differs from the behaviour change apps?

DBCIs - represented by mobile apps in this analysis, but not limited to that modality - are generally longer-term interventions, as forming new, healthier behavioural habits requires time. Although this time can vary depending on a number of factors relating to the individual, context, and behaviour [460,461], previous research has found that it takes approximately two months for behaviours to become automatic [462,463]. DHIs like Dora only require a brief period of engagement to achieve their intended purpose, but in an increasingly digitised healthcare system, users' experience with such interventions should be sufficiently acceptable that they would be willing to use such an intervention again.

[Section 3.8](#) laid out hypotheses of how different components might influence each other in DHI engagement based on previous theory and literature. The hypothesised model was not specific to a particular type of DHI, but was based around a process of micro engagement with the intervention and macro engagement with a health behaviour. DHIs that are not focused on behaviour change still require micro engagement to achieve their intended purpose. Therefore, the way that the components of engagement are related up to the point of engagement with the health behaviour can be expected to be similar. As with the different clinical cases, the relative importance of various factors might be different - for example, affective engagement with the intervention may have less of an impact on behavioural engagement for brief interventions, as patients are likely to be more willing to engage in something that is not intrinsically enjoyable if it is short (limited behavioural engagement) and has perceived value (cognitive engagement). However, a sufficiently unpleasant experience could still trigger behavioural disengagement, suggesting that all components are still relevant. Again, the differences in how and to what extent the components might influence each other in particular cases does not imply that the structural pattern of how they influence each other should vary. Therefore, cases 1 and 2 should be comparable to case 3 for, but not beyond, micro engagement. This assumption will be investigated in the cross-case synthesis. The synthesis will also explore how factors influencing affective, cognitive, and behavioural engagement were similar or different between the cases and how this might relate to the clinical area, intervention type, or individual characteristics and context.

The final major difference that could affect comparability is not in the cases themselves but in the study methodologies. Although all three examined acceptability and engagement primarily with qualitative thematic analysis within mixed-methods evaluations, the case studies were conducted sequentially and the structure of the SSIs evolved over the course of

the studies. The NoObesity topic guide was not based on a theoretical framework but focused on examining engagement with and perceptions of specific app features. The case study of Dora, the second evaluation conducted, structured the SSI topic guide using the Theoretical Framework of Acceptability (TFA). By the time the last case study (Wysa) was conducted, an explicit focus on examining how the different components of engagement were related was included in the study design, so the topic guide was based on both the TFA and the multifaceted conceptualisation of engagement. Incorporating the learnings of the previous case studies into the design of the subsequent ones improved my ability to capture the relevant data and interrogate the relationships between different components of engagement; however, it raises the question of how comparable the data generated from the interviews in these different case studies are.

The initial analyses were done with the particular aims of that case study in mind; for this thesis, each thematic analysis was revisited and the codes and themes were re-interpreted through the lens of the research question: *how are the different components of engagement (affective, cognitive, behavioural, micro, and macro) related?* This re-interpretation means that the within-case engagement mapping results are aligned around this question. However, the extent to which the whole process of engagement could be examined for each case was limited by the available data, as was referred to briefly in [Section 6.2](#). In some senses, this cross-case synthesis is similar to the execution of a qualitative meta-analysis, which aims to synthesise findings from studies addressing a particular topic, which may have different research questions or methodologies, to develop themes, insights, or theory that goes beyond the individual studies [464–466]. Differences between the studies can be addressed by examining how they influence the interpretation, which is how differences in methodology will be addressed in this cross-case synthesis. The benefit of a cross-case synthesis is in the focus on the cases as a whole, rather than reducing the analysis to an examination of variables [337]. This focus on analysing the data in context helps mitigate the potential limitation of the different SSI topic guide prompts, as the analysis is based on interpreting participants’ descriptions of their experiences rather than specific data points, as might be captured via a survey or other quantitative measure.

6.4 Pattern-Matching of Mappings of Engagement Components

6.4.1 Overview

This section details the comparison between the hypothesised and empirical models mapping the relationships between engagement components and highlights key similarities and

differences. All models demonstrated that affective and cognitive engagement influence and are influenced by behavioural engagement throughout the process of engagement. In this way, empirical results supported the hypothesis of the close interaction between these three components. The comparison of the case studies demonstrated that the relationships between engagement components were largely similar, despite differences between the cases ([Section 6.3](#)). A key difference between the hypothesised and empirical mappings was the order of influence of affective and cognitive engagement on behavioural engagement. The case study analyses did not support the hypothesised pattern of affective engagement influencing behavioural engagement via cognitive engagement; the opposite pattern was generated from the data - cognitive engagement influencing behavioural engagement via affective engagement (Figures 3-3 and 6-1). This discrepancy will be explored further in the following subsections, which discuss the findings in the context of the hypotheses.

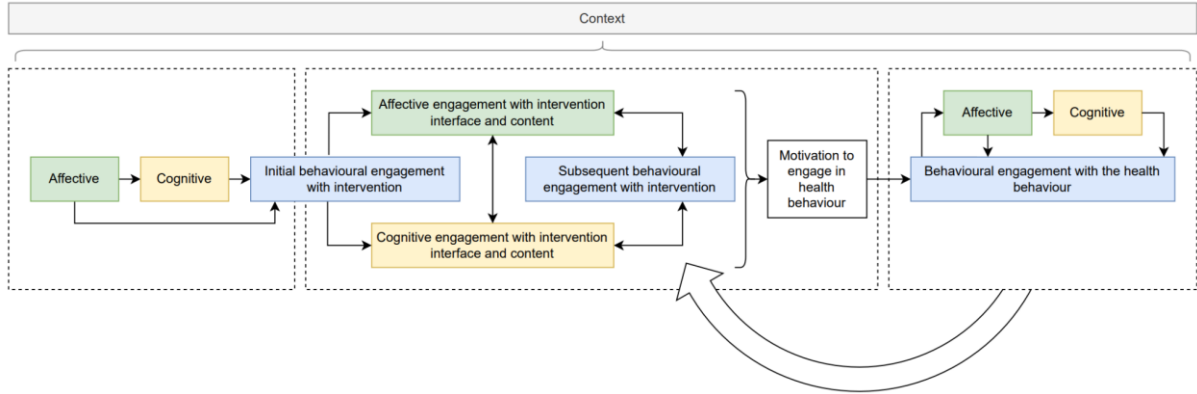


Figure 3-3 (duplicate). Proposed model of the process of engagement with a DHI and its target behaviour

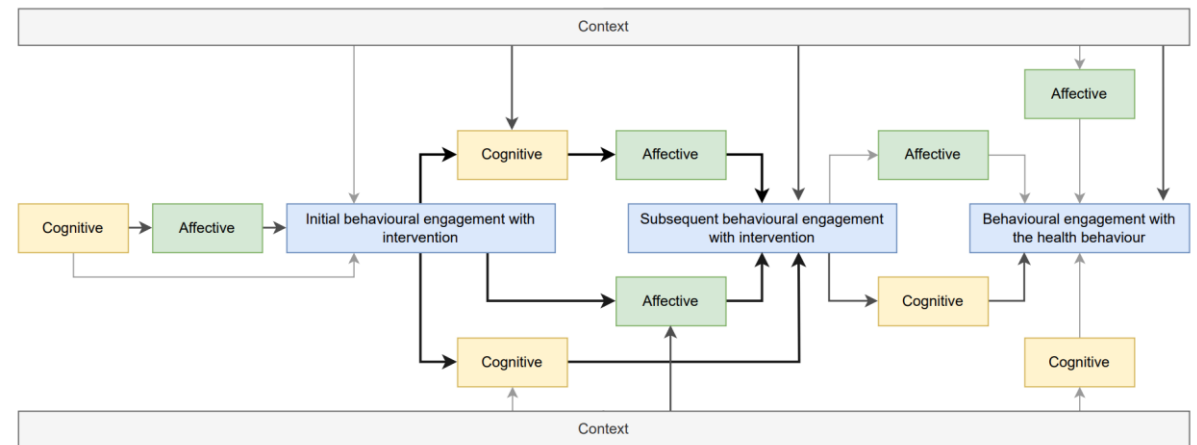


Figure 6-1. Combined mapping of NoObesity, Wysa, and Dora (*Note: the colour and weight of the arrows indicates whether a particular relationship was generated in 1, 2, or 3 cases; this is for illustrative purposes only, it does not represent a statistical count of evidence*)

All of the patterns of engagement captured in the hypothesised model and the three empirical case studies are detailed in Table 6-1. This table is intended to facilitate comparison of the patterns (mapped in Figures 5-7, 5-11, 5-15, and 6-1) and interpretation of how and why particular engagement components are related, what that means in the context of the cases, and what broader insights for engagement with digital health can be generated. Neither Table 6-1 nor Figure 6-1 should be interpreted as a numeric tally indicating the strength of particular patterns of relationships; there are insufficient cases for such a tally to be meaningful [337]. It is also important to note that the blank cells do not indicate that that particular pattern of engagement cannot occur, just that there was insufficient evidence to support it from these three particular case studies.

Table 6-1. Pattern-matching theoretical and empirical mappings of engagement components

Patterns of engagement	<i>Hypothesised (Chapter 3)</i>	Case 1: NoObesity	Case 2: Wysa	Case 3: Dora
Evidence for components affecting initial micro behavioural engagement				
Affective → Initial Behavioural (micro)	X		X	
Cognitive → Initial Behavioural (micro)			X	X
Affective → Cognitive → Initial Behavioural (micro)	X			
Cognitive → Affective → Initial Behavioural (micro)			X	X
Context → Initial Behavioural (micro)				X
Evidence for components affecting subsequent micro behavioural engagement				
Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro)	X	X	X	X ^a
Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)	X	X	X	X
Initial Behavioural (micro) → Affective → Cognitive → Subsequent Behavioural (micro)	X			
Initial Behavioural (micro) → Cognitive → Affective → Subsequent Behavioural (micro)		X	X	X ^a
Context → Subsequent Behavioural (micro)		X	X	X ^a
Context → Affective → Subsequent Behavioural (micro)		X	X	
Context → Cognitive → Subsequent Behavioural (micro)	X ^b	X		X ^a
Context → Affective → Cognitive → Subsequent Behavioural (micro)	X ^b			
Context → Cognitive → Affective → Subsequent				X ^a

Behavioural (micro)				
Evidence for components affecting translation from micro to macro behavioural engagement				
Behavioural (micro) → Affective → Behavioural (macro)		X		
Behavioural (micro) → Cognitive → Behavioural (macro)		X	X	
Behavioural (micro) → Affective → Cognitive → Behavioural (macro)				
Behavioural (micro) → Cognitive → Affective → Behavioural (macro)				
Evidence for components affecting macro behavioural engagement				
Affective (macro) → Behavioural (macro)	X	X		
Cognitive (macro) → Behavioural (macro)		X		
Affective (macro) → Cognitive (macro) → Behavioural (macro)	X			
Cognitive (macro) → Affective (macro) → Behavioural (macro)				
Context → Behavioural (macro)	X ^b	X	X	
Context → Affective (macro) → Behavioural (macro)	X ^b	X		
Context → Cognitive (macro) → Behavioural (macro)	X ^b	X		

^aAs the Dora call comprised one telephone call, 'subsequent behavioural (micro)' in this table can refer both to how initial engagement or context changes subsequent behavioural engagement during the same call and to how engagement with the first call might hypothetically influence engagement with another call, if offered (see Figure 5-15 and [Appendix 5-I](#))

^bThe hypothesised relationships between context and engagement were primarily drawn from the AIM-ACT framework; as this did not break down engagement into the same stages included here, the hypothesised patterns have been selected based on interpretation of the description of 'context' as a component of the framework in the paper [49]

6.4.2 Pattern-matching analysis in relation to hypotheses

Following the initial pattern-matching (Table 6-1), this section will explore similarities and differences between the case studies in the context of the theoretical hypotheses. Table 6-2 summarises the hypotheses outlined in [Section 3.7](#) and the extent to which they are supported by the findings. Each hypothesis is discussed in greater detail in the following subsections.

Table 6-2. Summary of hypotheses and supporting evidence

Hypothesis	Evidence
<i>Initial micro engagement (Section 6.4.2.1)</i>	
H1a: Given that an individual is aware of and able to access a digital health intervention, initial micro behavioural engagement will be driven by affective engagement with the idea of the intervention either directly or indirectly by influencing cognitive engagement (the individual's appraisal of the expected benefits, risks, and demands of intervention).	The analysis supported the hypothesis that initial behavioural engagement was driven by both affective and cognitive engagement; however, the patterns of engagement developed from the data suggested that cognitive engagement influenced behavioural engagement via affective engagement, rather than affective engagement influencing behavioural engagement via cognitive engagement as hypothesised.
H1b: Both positively- and negatively-valenced emotions can support affective engagement and initial micro behavioural engagement.	There was evidence for positively-valenced affective engagement supporting initial behavioural engagement across the studies. Evidence for the influence of negatively-valenced emotions was more limited but aligned with H1b.
H1c: Only positive cognitive engagement will support initial micro behavioural engagement.	This hypothesis was supported, but cannot be confirmed, as there were no findings suggesting negative cognitive engagement (ie. lack of attention or interest, judgements that costs outweighed potential benefits, etc.) increased behavioural engagement.
<i>Subsequent micro engagement (Section 6.4.2.2)</i>	
H2a: An individual's initial behavioural engagement with a digital health intervention will influence their affective and cognitive engagement with the intervention by providing further experiences which generate subsequent affective responses and cognitive assessments.	This hypothesis was supported by evidence across the studies indicating that participants' use of the DHI shaped their subsequent affective, cognitive, and behavioural engagement with it.
H2b: The intervention needs to be sufficiently affectively OR cognitively engaging to maintain behavioural engagement with the intervention interface and content.	This hypothesis was somewhat supported by the findings from the three case studies, as affective and cognitive disengagement were associated with behavioural disengagement, but their influences could not be sufficiently distinguished to determine if one alone would be sufficient to sustain behavioural engagement.
<i>Translation of micro to macro engagement (Section 6.4.2.3)</i>	
H3a: Affective or cognitive engagement with the intervention content will increase	This hypothesis was supported by findings from case studies 1 and 2 suggesting that

an individual's motivation to engage in the target health behaviour.	ffective, cognitive, and behavioural engagement with the intervention were associated with engagement with the target behaviour (physical activity/healthy eating and use of Wysa's mental health resources).
H3b: The translation of motivation into behaviour is dependent on sustained cognitive engagement (attention) and influenced by the individual's context.	All findings highlighted the importance of reminders for prompting DHI engagement, but attention could not be specifically linked to motivation with the available evidence. Context played a role across the process of engagement in all cases and was associated with motivation, primarily in case study 1.
<i>Macro engagement and influence on subsequent micro engagement (Section 6.4.2.4)</i>	
H4a: As the individual engages with the health behaviour, affective, cognitive, and behavioural components of engagement with the health behaviour will all influence each other to determine whether macro behavioural engagement is sustained.	There was insufficient evidence to determine whether this hypothesis could be supported.
H4b: High macro engagement with the health behaviour may result in decreased micro engagement with the intervention as the health behaviour becomes more habitual.	There was insufficient evidence to determine whether this hypothesis could be supported.
H4c: Decreasing macro engagement with the health behaviour may prompt an increase in micro engagement with the intervention, if there is still macro affective or cognitive engagement with the health behaviour goal.	There was insufficient evidence to determine whether this hypothesis could be supported.

6.4.2.1 Initial micro engagement

As the NoObesity case study did not capture any data about influences on initial behavioural engagement, hypotheses relating to initial micro engagement could only be assessed through the synthesis of case studies 2 and 3. While the case studies supported the hypothesised importance of affective and cognitive engagement, there was a key difference between **Hypothesis 1a** (based on previous literature) and the case study findings: interpretation of the data generated a reversal of the predicted order of influence (affective → cognitive). In both case studies, the data suggested that cognitive engagement influenced initial behavioural engagement both directly and indirectly via affective engagement. Potential explanations for this difference will be explored in [Section 6.4.3](#).

The two cases were largely aligned; the main differences were that case study 2 highlighted a potential direct association between affective and initial behavioural engagement while case study 3 highlighted the role of context on initial behavioural engagement. Care should be taken in over-interpreting these differences, which could relate to specific case characteristics - the key finding was that both affective and cognitive engagement could influence initial behavioural engagement. Determining whether this influence was direct or indirect was more vulnerable to bias in how participants reported their experiences and how I interpreted their descriptions of their experiences.

Hypothesis 1b predicted that both positively- and negatively-valenced emotions could support affective engagement and initial micro behavioural engagement. In both cases curiosity and interest were reported as encouraging initial engagement; whether these can be considered ‘emotions’ is debatable [434,435], but as highlighted in [Sections 3.6.2.2](#) and [5.3.4](#), they have affective as well as cognitive aspects that make them relevant for the assessment of this hypothesis. The role of negative emotions was more complicated. Negative emotions (worry about a lack of mental health support) were associated with positive affective engagement (feelings of relief) and behavioural engagement with Wysa, supporting H1b. On the other hand, some participants highlighted negative emotions expected with use of Dora - namely regret and expected frustration at communicating with a machine. This did not appear to support positive affective engagement, but neither did it prevent interviewed participants from using Dora; however, such emotions may have contributed to patients declining to participate and thus avoiding engagement ([Section 6.5.3.2](#) discusses this with reference to recruitment data).

Hypothesis 1c posited that only positive cognitive engagement would support initial micro behavioural engagement. This was difficult to test with the data available. While there was no indication that negative cognitive engagement supported initial behavioural engagement, this lack of evidence cannot prove the hypothesis. Considering the definition of cognitive engagement established in [Section 3.6.3](#), negative cognitive engagement with the idea of a DHI would either be represented as a lack of attention, a lack of interest, or a judgement of the intervention as too difficult to use or not likely to be of sufficient benefit for the effort or costs required. An association between cognitive disengagement and behavioural engagement would not align with theoretical expectations [42].

6.4.2.2 Subsequent micro engagement

In line with **hypotheses 2a and 2b**, patterns generated from all case studies indicated that initial behavioural engagement influenced affective and cognitive engagement, which in turn influenced subsequent micro behavioural engagement. Each analysis suggested that continued micro behavioural engagement could be independently influenced by both cognitive and affective engagement, which is aligned with theoretical expectations. As with H1a, the main difference between the hypothesised and empirical mappings was the order of influence of affective and cognitive engagement. In contrast to the hypothesised order (affective → cognitive), the pattern generated from all three case studies was that, when there was an indirect relationship, it was cognitive engagement influencing affective engagement. Potential explanations will be explored in [Section 6.4.3](#).

Hypothesis 2b predicted that, to support subsequent micro behavioural engagement, the intervention would need sufficient affective OR cognitive engagement. Themes from case studies 1 and 2 indicated that both affective and cognitive disengagement could increase the likelihood of behavioural disengagement. For example, once all content in the NoObesity app had been explored, users became bored and experienced less motivation to use the app. Likewise, Wysa's CA limitations led to frustration and decreased motivation to engage. However, it was difficult to distinguish between the influence of affective and cognitive engagement, as participant feedback largely blurred these together or included both.

Context had an important influence on the likelihood of subsequent micro behavioural engagement in all case studies, but it appeared to be relatively more important for Dora. For example, most participants completed Dora's brief intervention, regardless of their attitudes towards it. As discussed in [Section 6.3](#), this difference in the relative importance of factors was expected. As a brief intervention with a single clear objective, Dora does not need to be especially interesting or enjoyable; it only needs to be sufficiently usable and acceptable that users do not hang up mid-assessment. NoObesity and Wysa, however, need to sustain engagement for longer periods of time to support behaviour change. Another significant contextual factor in case study 3 was the presence of complications or concerns after the surgery. While it was unclear whether this would influence behavioural engagement, the analysis suggested that it would have a significant influence on cognitive and affective engagement. The intended context of the intervention delivery - as part of the standard post-surgical care pathway - might be sufficient to support behavioural engagement even lacking affective and cognitive engagement (i.e. if this is the only follow-up offered, participants may engage with it even if they find it upsetting or insufficient). It is important that evaluations of

engagement take such contextual factors into account - an operational evaluation of such an intervention might record a high call completion percentage and assume the intervention is working successfully without capturing affective or cognitive components. Patient experience and comfort should not be disregarded in a quest for improved efficiency, so it is essential that researchers are aware to look for such outcomes.

6.4.2.3 Translation of micro to macro engagement

As the Dora intervention did not involve any health behaviour change, the analyses of the following two sets of hypotheses are limited to Cases 1 and 2. **Hypothesis 3a**, that affective or cognitive engagement with the intervention would increase motivation to engage in the target health behaviour, was supported by both. In case study 1, it was more clear how cognitive and affective engagement with the intervention translated to macro engagement with the health behaviour. For example, parents reported that cognitively engaging with content that highlighted the health benefits for children would increase their motivation to do the behaviour and that framing the behaviours as a family activity made their perceptions more positive. While engagement with the intervention and behaviour were conflated in Wysa because users were guided through therapeutic exercises on the app, the same pattern was observed (e.g. positive experiences with exercises encouraged further engagement). This supports the theoretical expectation that both affective and cognitive engagement can help support behavioural engagement with interventions and target health behaviours.

Interestingly, there was little evidence that Wysa users extrapolated their in-app behaviours to their everyday lives (without the app). This could be related to the context in which users engaged in mental health behaviours compared to obesity-related behaviours. Self-reflection or meditation behaviours generally occur in a quiet moment alone - assuming users have their phone with them, there is no particular reason to engage in that behaviour without the app facilitating the process. On the other hand, obesity-related behaviours require some type of interaction with the environment. While an app might facilitate this (e.g. by providing a recipe or list of exercises), the behaviour is not conducted via the app.

Hypothesis 3b predicted that the translation of motivation into behaviour would be dependent on sustained cognitive engagement (attention) and influenced by context. The themes and patterns generated were in line with this hypothesis, although the link between attention and the translation of motivation into health behaviour could not be definitively supported.

Themes from both NoObesity and Wysa highlighted that lack of attention was a key barrier to

subsequent engagement; reminders were an important facilitator to help prompt micro and macro engagement.

As discussed above, the role of context in the whole process of engagement was established across the case studies. In terms of translation to macro behavioural engagement, contextual factors were key in case study 1 (for example, weather, daylight, and availability of resources influenced motivation, opportunity, and capability to engage with target health behaviours). This case study had the strongest evidence associating context with motivation, but case study 2 also highlighted how available time and current circumstances could influence engagement with Wysa.

6.4.2.4 Macro engagement and influence on micro engagement

The final set of hypotheses (4a-4c) focused on the final stage of engagement - sustaining macro engagement with the target health behaviour - and the feedback of that engagement on micro engagement with the DHI. Unfortunately, there was insufficient evidence to draw conclusions about these hypotheses. The best evidence came from case study 1, which found that, in line with **hypothesis 4a**, macro affective and cognitive engagement were associated with macro behavioural engagement. However, no data was captured that suggested experiences engaging with the health behaviour then revised affective or cognitive engagement with the intervention or behaviour. The hypothesis is not rejected; this was not a primary focus of the studies.

There was evidence from case studies 1 and 2 that engagement with the intervention decreased over time, but no evidence was captured suggesting that this was because of high engagement with the target health behaviour, as predicted in **hypothesis 4b**. Likewise, conclusions could not be drawn about **hypothesis 4c**, that decreasing engagement with the health behaviour could prompt re-engagement with the DHI, as this was not reported by participants.

6.4.3 Potential explanations for differences between hypothesised and empirical mappings

The main difference observed in the pattern-matching analysis was the order of influence of affective and cognitive engagement. Although a full systematic search was not conducted to derive the hypothesised model, the sources reviewed consistently proposed patterns of affective → cognitive → behavioural engagement [49,62,66,173]. One potential explanation is that the context of the studies from which the hypothesised model was developed varied

from the current context. Two of those studies were based on a framework of neural activation related to decision-making [49,173] and one proposed a model of acceptance based on previous literature about how affective heuristics influence cognitive risk and benefit judgments, which was tested using a questionnaire [62]. The study that was most comparable to the current context was a mixed-methods evaluation of engagement with a DHI for diabetes, which concluded that affective components influenced initial engagement, which was followed by cognitive engagement (familiarisation) with the DHI [66]. However, in their discussion of the initial affective step, the authors describe participants' motivation to engage as driven by "the thought of avoiding medication and health risks associated with [type 2 diabetes]" and wanting "to take advantage of the opportunity" to be part of the program [66], which has some alignment with the cost-benefits assessments included in my definition of cognitive engagement.

As discussed in Chapter 3, the concept of motivation can include both affective and cognitive elements. For example, in the Wysa case study, participants expressed being curious, excited, and relieved. Interrogating the reasons behind these emotions determined that participants felt curious and excited because they were interested in trying something new and felt relieved because they perceived that the tool would meet an unmet need. It is possible that a similar interrogation of the diabetes study would find that participants' felt worried because of their judgments around the risks associated with type 2 diabetes or that they wanted to take advantage of the program because they expected it would benefit them. These cognitive elements could well be components of the affective response identified in the diabetes study, but were not discussed in the paper [66].

Previous research has identified a very close relationship between affective and cognitive processes [326,329]; perhaps there is an ongoing feedback loop between cognitive and affective experiences and appraisals at such a close level that it is difficult to disentangle them into an orderly process. This could explain why the hypothesised and empirical mappings propose different orders - because of the context in which the previous models were developed, they may be picking up on different stages of the process compared to my thesis. For example, one of the findings from case study 2 relating to initial behavioural engagement was that participants were worried about a lack of mental health support while they were on the waitlist for treatment. It could be argued that this represents affective engagement, which leads to cognitive engagement (consideration of options to address this negative emotional state and the selection of an alternative, such as Wysa, that could meet the users' need). This would align with the hypothesised affective → cognitive → behavioural pattern derived from

the literature and contradict my analysis. Making such judgments when generating the patterns of engagement was difficult; in this case, I would argue that while this worry does represent a negative affective state, it does not represent affective engagement with the intervention, which was defined in [Section 3.6.3](#) as relating to *perceptions of and experiences engaging with the intervention and the behavioural goal*. The worry experienced is related to the patient's context - long wait times and lack of resources - which might drive cognitive engagement (interest and attention) with the intervention, but is not directly related to the intervention itself.

Alternatively, the difference could be due to the way the data was captured in this thesis. The data analysed for each case study was qualitative, provided retrospectively. Data was limited to what participants remembered experiencing and shaped by their subsequent interpretations - for instance, it is possible that they experienced affective engagement before cognitive engagement, but were unaware of it, or in recounting the experience ordered the components in a way that made more logical sense to convey. For example, "I felt X because Y" would reflect the cognitive → affective pattern generated; however, if participants experienced an affective heuristic that they were unaware of (as Liu suggests [62]) that shaped their cognitive judgement, and a subsequent affective response that they were aware of, only the cognitive judgement and subsequent affective response would be reported. Participant recollections are a different type of data compared to neural activity, which may be able to shed more light on the order of these quick, integrated responses. This will be an interesting area for future research ([Section 7.4.1.2](#)).

For the purposes of digital health engagement, theoretical and empirical evidence agree that both affective and cognitive components are important influences on each other and on behavioural engagement. To determine how and why these components shape behavioural engagement, it is necessary to dive deeper into the factors influencing these components and how those factors are shaped by the context and characteristics of the interventions.

6.5 Thematic Synthesis of Factors Influencing Engagement

6.5.1 Overview

The pattern-matching of the hypothesised and empirical mappings of the engagement components provided a high level structure of how they can influence each other. To design DHIs that can support sufficient engagement to achieve their aims, it is essential to understand the relative importance of each relationship and common barriers that interfere with them in particular contexts. This section describes the results of a thematic synthesis of

the qualitative data and analyses from the case studies. Four main themes were generated: connection with the intervention; attention, interest, and incentive to use; usability for the target population; and confidence in the intervention. The following subsections summarise each theme and explore how the findings compared across the case studies and how specific case characteristics might have influenced the salience of their association with engagement. The thematic synthesis framework is included in [Appendix 6-A](#) and summarised in Figure 6-2 and [Section 6.6.1](#).

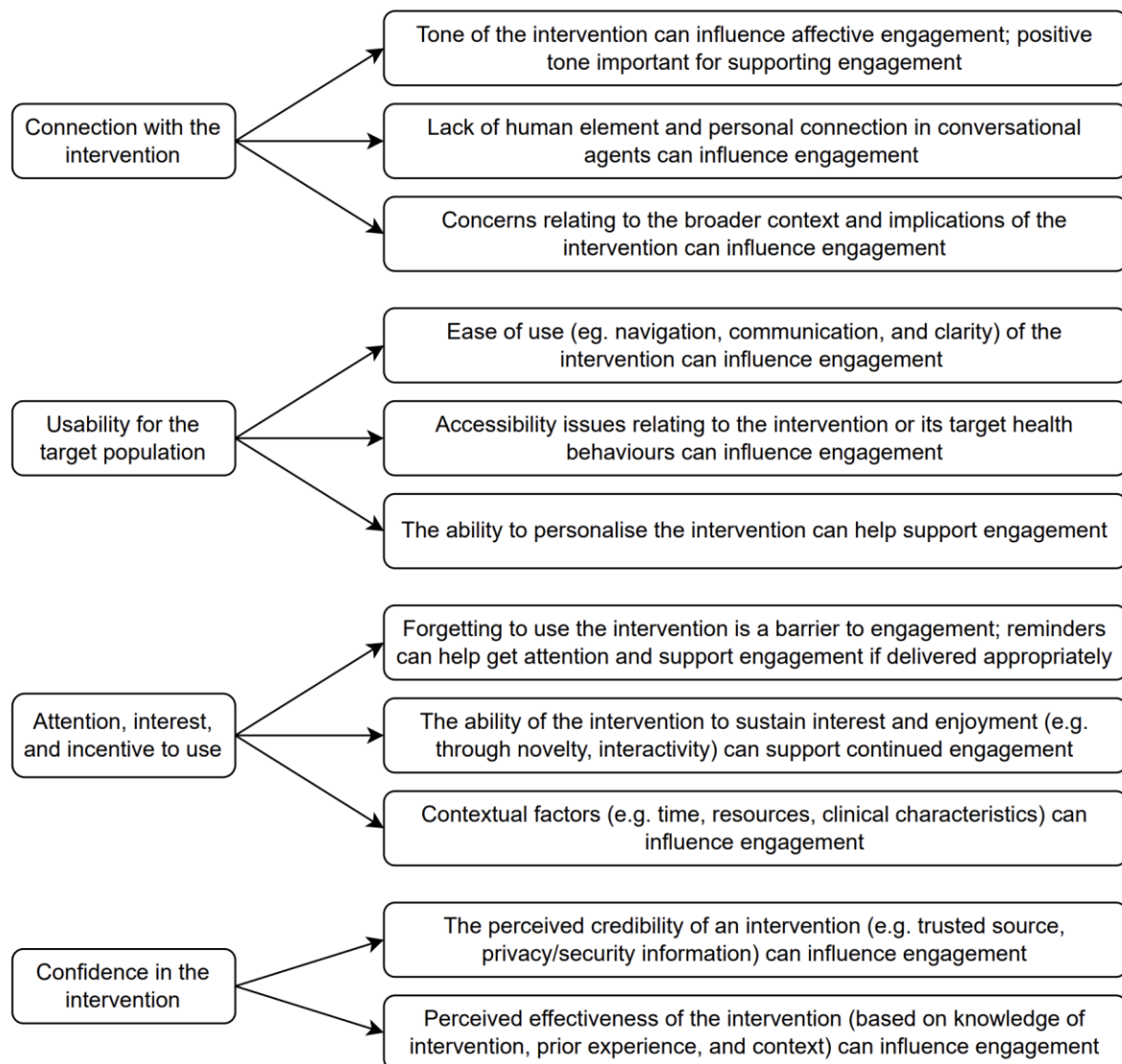


Figure 6-2. Summary diagram of main themes generated from thematic synthesis

6.5.2 Theme 1 - connection with the intervention

6.5.2.1 Summary

The first main theme generated was the importance of users' sense of connection with the intervention. This connection influenced a number of factors and shaped affective, cognitive,

and behavioural engagement. In NoObesity and Wysa, tone was a factor influencing affective engagement; a positive, non-judgmental tone could improve users' feelings about and connection with the intervention, whereas perceiving the tone as negative or condescending reduced motivation to continue using it. Tone can be conveyed by various features (generally language and design) and can be perceived differently by individuals - for example, users had varying reactions to the Wysa penguin avatar mascot.

Perceptions of Wysa's CA tone as genuine and non-judgmental supported users' engagement, while perceptions of Dora's 'voice' as brisk, remote, and impersonal negatively influenced affective engagement. Discussion with the CAs was felt to be different from discussion with a person and participants in both cases desired that missing empathetic, human touch. This theme primarily related to affective engagement with the intervention, but also influenced cognitive judgments about the interventions' capacity and quality. Limitations with the CAs - eg. repetition, limited ability to understand nuances or to support two-way discussion - decreased users' confidence in the interventions and were associated with frustration, negatively influencing affective engagement.

6.5.2.2 Potential explanations for differences between cases

There was less evidence for the importance of connection in NoObesity compared to Wysa and Ufonia. There are a couple potential explanations for this difference. First, unlike the other cases, NoObesity did not include a CA - participants interacted with the app, but not as if it was a 'person'. A sense of connection may be more important for engagement with interventions that mimic (to some degree) human interaction. Second, the difference could be due to the different characteristics of the clinical application: users of Wysa might be in more vulnerable or emotional states and users of Dora might be experiencing worrying symptoms. This explanation is supported by the fact that user assessments of Dora's suitability and effectiveness were influenced by whether they had complications. In both cases, a sense of connection appeared to be particularly important for users experiencing emotions while interacting with the intervention. It is possible that users who engaged with NoObesity, which focused primarily on encouraging health behaviours, were less likely to be experiencing, or expect the intervention to help manage, emotions. Despite this, there were some emotions associated with NoObesity; several users mentioned feeling discomfort at the idea of tracking their own, or their children's, weight. More broadly, obesity has been associated with mental health and self-esteem issues, suggesting that it may be an emotional topic for many users [467], which would not align with this potential explanation.

A third explanation is that the sense of connection with the intervention is important for all types of DHIs but is expressed differently, depending on intervention characteristics. In psychotherapy, a ‘therapeutic alliance’ between patient and therapist is an important mediator of engagement and positive outcomes [437]. As highlighted in [Section 5.3.4](#), the most common conceptualisation of the therapeutic alliance includes three elements: an emotional *bond* of trust and confidence, agreement on the *goals* of the therapy, and co-operation on the *tasks* required to achieve those goals [468–471]. While there is so far limited conceptualisation and evidence of an association between therapeutic alliance and positive outcomes in a digital context [471], it has been positively associated with engagement and adherence to treatment [470–472]. There are similarities between the aim of a therapist and a DHI - many DHIs, like types of therapy such as CBT, aim to change thoughts, feelings, and ultimately behaviour, which requires alignment between user and intervention on the goal and the steps required to achieve it. The role of a bond between user and DHI is less clear [470], but aligns with the theme of ‘connection’, suggesting a potential parallel for future exploration.

Assuming that these elements of the therapeutic alliance (bond, goal, tasks) can be translated from psychotherapy to digital health, what is still unclear is to what extent each is important for the development of an alliance. Certain characteristics of an intervention or its clinical case may influence the relative importance of the ‘bond’, as highlighted in the second explanation above. Alternatively, the way that the bond is understood and expressed by users may differ depending on the intervention, even if it is equally important for engagement. For example, in the CA interventions, the ‘connection’ or ‘bond’ may be more obvious to users and relevant to discuss in an interview, because it is a large component of the intervention. How NoObesity made users feel was important, but they may have been less likely to consider this in terms of a ‘bond’; compare this to Wysa, where one user was relieved because the intervention offered a source of support without the concern of burdening a friend with their problems. Essentially, the way a ‘bond’ or ‘sense of connection’ is conceptualised and discussed by participants may vary, despite it remaining an important element in supporting user engagement.

6.5.3 Theme 2 - attention, interest, and incentive to use

6.5.3.1 Summary

The process of users’ attention and interest in the interventions varied over the course of their engagement. Interest and curiosity - and expectations of potential benefit - supported users’

engagement. This was particularly observed in NoObesity and Wysa; in both cases, factors such as boredom, repetitiveness, and a lack of new content or other incentive to draw users back to the app were significant barriers that were associated with frustration and decreased engagement. Although a couple participants also expressed interest in Dora, most had no particular expectations before use. A likely explanation for this is that the app-based interventions involved longer-term, more self-motivated engagement whereas Dora simply required participants to consent and answer a brief telephone call. A related explanation is that the apps had a variety of content to explore, while Dora was simply an automated call, which may have been viewed as a task to complete rather than something to be interested in. For short interventions with direct clinical purposes, interest may be less important for engagement.

Reminders were a key factor for supporting attention and engagement in all cases. NoObesity and Wysa users reported that notifications helped prompt app use and some Dora users suggested reminders as a way of increasing engagement by ensuring that patients knew to (and when to) expect the call. Timing and other contextual factors could have an important influence on engagement. Several of the issues with the apps' notifications revolved around when participants received them - participants engaged with the apps at particular times and settings; notifications outside of those times and places were not as helpful at prompting use.

6.5.3.2 Potential explanations for differences between cases

Findings from the Dora study did not provide much evidence for the importance of attention, interest, or enjoyment for engagement. Only a couple participants mentioned being interested in Dora, and while some participants experienced frustration communicating with Dora, none reported disengaging because of this. Here frustration refers to the negative affective state experienced when a goal is thwarted [473,474]; the experience of frustration does not necessarily lead to disengagement - in some circumstances it can motivate further engagement (e.g. gaming, gambling) [475] - but in this case, was associated with decreased cognitive and affective engagement. There are a couple potential explanations for why these participants remained behaviourally engaged without strong affective or cognitive engagement: 1) Dora was short enough and served a sufficiently discrete purpose that patients were willing to tolerate a brief period of frustration, 2) participants who became frustrated disengaged from the call, or 3) patients who lacked interest did not engage at all. The second explanation can be refuted by the participant flow information - of the 202 patients Dora called, only 7 calls were considered 'incomplete'. The third cannot be definitely refuted, although the recruitment data collected does not provide strong evidence for it; of the 542 patients who were eligible

but not included, only 5 specifically mentioned dislike of automation as their reason for non-participation and only 12 of the 225 recruited did not answer the call. This leaves the first explanation as the most likely.

6.5.4 Theme 3 - usability for the target population

6.5.4.1 Summary

Another theme that was important across the studies was usability for the target population. Interface design (aesthetic and functional) was an important factor affecting engagement - lack of clarity about intervention aims or guidance on how to use it confused users and increased frustration. Perceived capability of the intervention was another important factor that is also related to theme 4; in the CA interventions, several users reported purposefully providing simple inputs to be sure that they would be understood. For Dora, accessibility factors were also raised, including physical (e.g. hearing impairments) and cognitive barriers (e.g. time pressure of the call and uncertainty in how to respond). A couple participants also raised the issue of accessibility for people with neurodivergence (e.g. autism). This was seen as a significant barrier for use of Dora - a participant with autism was extremely distressed by the intervention because the format did not enable the two-way discussion, context, and clarification needed to accurately capture their experience.

One factor that could support usability and accessibility, and improve users' experience, is personalisation. This was raised in all of the studies, although suggestions for how it could be achieved varied. For example, clarity and simplicity in Wysa could be improved by enabling users' to save a personalised collection of favourite resources for quick access and accessibility issues for Dora could be improved by enabling scheduling, capturing additional contextual information, or providing questions in advance.

6.5.4.2 Potential explanations for differences between cases

Compared to Dora, there was little discussion of the accessibility of NoObesity and Wysa. Accessibility issues raised in the NoObesity intervention related primarily to engagement with the target behaviours (eg. affordability of healthy food, availability of safe outdoor spaces) rather than the intervention. For Wysa, participants liked that it increased accessibility of mental health support while waiting for treatment. Accessibility and health equity are central issues for digital health, so the difference in these cases must be examined. Although not mentioned explicitly by participants, the equivalent accessibility factors regarding Dora and hearing impairments would exist for NoObesity and Wysa for people with visual

impairments; this was likely not raised because people with visual impairments would be unable to use the intervention to participate. Although people with cognitive impairment and hearing impairments were ineligible as participants for the Dora study, it is possible that the older population of participants had minor hearing difficulties or were more aware of that as a potential barrier for widespread intervention use. Another difference that could explain this discrepancy in results is how Dora is positioned as a tool compared to NoObesity and Wysa. The two apps, even if they were incorporated into standard clinical pathways, would be supplementary resources. On the other hand, if Dora was incorporated into standard clinical workflows, it might be, or appear to be, mandatory. While alternative pathways could be incorporated for patients with hearing impairments or neurodivergence even if Dora was widely adopted, patients may not be aware of this and feel they had no choice to interact with Dora, which could make accessibility concerns more salient. Although ensuring accessibility for every individual may be beyond the scope of some interventions, factors relating to accessibility are essential to consider and, where possible, to address (e.g. by including voice access in health apps).

6.5.5 Theme 4 - confidence in intervention

6.5.5.1 Summary

The final theme is closely related to the other main analytical themes discussed above, but is discussed separately to highlight its importance and how it derives from them. There are several factors that can influence confidence and trust in an intervention: connection, credibility, experience of use, perceived quality, privacy or other concerns, and clinical or other contextual factors. Credibility can be an important element for connection [470], but interestingly, it was raised more explicitly by participants in the Wysa case study than the others. We would expect credibility to be an important factor for all types of DHI, which are often dealing with personal or sensitive data and providing health guidance. So what case characteristics might explain why credibility was generated as a factor in Wysa but not NoObesity or Dora?

6.5.5.2 Potential explanations for differences between cases

For NoObesity, with its focus on health behaviours, credibility may have been less salient because participants generally perceived themselves as knowing how to behave healthily and being capable of doing so, just needing a nudge to encourage action on that knowledge. NoObesity's purpose was primarily to encourage behaviours that most participants were familiar with (e.g. a suggested goal to eat more vegetables); however, the app provided advice

for specific challenges, which may have been new information. An alternative explanation is that the goals were not perceived as being 'high stakes' - for example, following a piece of advice that was unsuccessful or inaccurate might be perceived as unlikely to have significant negative effects - and thus credibility was less important. This is countered by participant feedback about feeling uncomfortable tracking their weight and concerned about causing distress or self-consciousness, which highlights the emotional significance of weight-related issues. Potential privacy issues were raised as a factor in both Dora and NoObesity, but did not appear to be a major concern for interviewed participants. A third potential explanation is that the focus of the qualitative interview on the app features did not provide an opportunity for discussions of credibility. I would argue that it is likely a combination of the first and third explanations; credibility likely is important, but the development of the app by a governmental body (Health Education England), its alignment with users' prior knowledge, and the structure of the SSI meant that this was not discussed in depth.

Wysa participants referred to in-app source citations and the delivery of the app via the NHS as factors that supported its credibility (as well as their own positive experiences with it). Although most participants did not discuss credibility in Dora specifically, it seems likely that its delivery as part of the NHS care pathway supported perceptions of the tool as credible, at least from a clinical standpoint. The issues with credibility and trust in Dora revolved instead around technical capability, rather than the clinical assessment (which was derived from standard procedure). Context was particularly important for users' confidence in Dora - most participants felt confident in Dora's ability to conduct the assessment if they had no issues, but were less confident in its ability to manage concerns or complications. This highlights how an intervention's perceived quality and capability influence users' confidence in it, which was in line with Wysa findings - participants in both CA cases felt the need to provide simple inputs to be sure that they would be understood.

6.6 Discussion

6.6.1 *Principal findings*

The cross-case synthesis emphasised the importance of affective, cognitive, and behavioural components of engagement and their interconnectedness over the process of engagement. Patterns of engagement generated from the three case studies were largely similar; differences were mainly due to data collection methods and case characteristics that influenced which stages of behavioural engagement were captured. The main difference observed between the hypothesised and empirical models was the order of influence of cognitive and affective

engagement on behavioural engagement. The literature-base hypothesised model posited that affective engagement influenced cognitive engagement, while the opposite trend was observed in the case studies. This was discussed in [Section 6.4.3](#) and will be explored further in [Section 6.6.2](#).

The thematic synthesis generated a number of common factors that influenced cognitive and affective engagement (in varying degrees) and thus were important for behavioural engagement with DHIs (Figure 6-3). Context - whether clinical, personal, or situational - could influence all of the factors as well as behavioural engagement directly (depending on whether the user had the opportunity to engage with the intervention). Users' sense of connection with the intervention was primarily associated with affective engagement, but this could influence behavioural engagement via users' perceptions of an intervention's credibility and their confidence in its ability to meet their needs (cognitive judgments). Usability and accessibility were primarily related to users' ability to cognitively and behaviourally engage, but barriers to usability and accessibility negatively influenced affective engagement and confidence in the intervention (related to cognitive engagement). Attention and interest were primarily related to cognitive engagement, but as with usability, a lack of interest could also influence affective engagement by triggering feelings of boredom or frustration. Specific case characteristics influenced the importance of each theme in influencing behavioural engagement with the intervention, but all appeared relevant to some degree. For discrete, short-term interventions, low interest may not be a significant barrier to engagement. For CA interventions, connection and confidence appear to be especially important. For longer-term, DBCIs, sustaining connection, attention, and interest is likely to be necessary to support behavioural engagement over time. Usability is a critical prerequisite for behavioural engagement with all types of intervention.

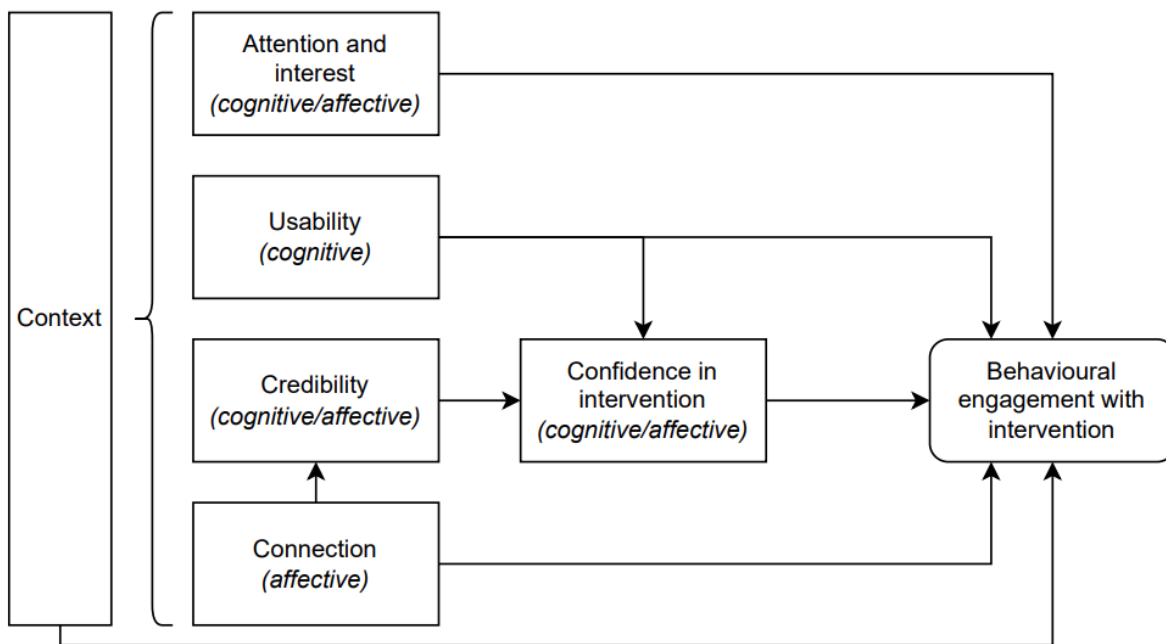


Figure 6-3. Thematic map of factors influencing micro behavioural engagement

6.6.2 Proposed model of how engagement components are related

This cross-case analysis adds to the body of evidence for potential patterns of relationships between engagement components but has also raised questions about the validity of attempting to map and order distinct processes. There was general consensus between the case studies and previous literature that cognitive and affective components could influence and be influenced by the behavioural component at various stages of engagement. The main aspect of debate is the process by which affective and cognitive engagement influence each other. The literature reviewed in [Section 3.6](#) proposed that affective engagement comes first, with various explanations - motivating initial behavioural engagement that was then followed by cognitive familiarisation [66], influencing cognitive risk-benefit judgments that shaped behavioural intentions [62], or providing an initial interpretation that was then contextualised by a cognitive interpretative process [49,173]. In the context of digital health engagement, this implies that emotional reactions to the intervention come first and are then included in cost-benefit evaluations of the value of behaviourally engaging in the intervention [49].

In the case studies, participants typically reported emotional experiences as occurring due to cognitive engagement. For example, frustration was associated with limitations in users' ability to communicate with a CA and relief was associated with expectations that the intervention would meet a need. This data suggested that cognitive engagement can influence behavioural engagement via affective engagement, rather than the other way around.

The thematic synthesis conducted in [Section 6.5](#) provides support for an alternative explanation raised in [Section 6.4.3](#), that both cognitive and affective engagement can influence each other, depending on the context and stage in the process of engagement. For example, cognitive engagement with intervention content that is hindered by poor interface design, app freezing, hearing difficulties, or other usability issues could negatively influence affective engagement (by causing frustration), which would in turn decrease the likelihood of continued behavioural engagement. On the other hand, as a second example, feeling a positive emotional connection with the intervention could influence participants' cognitive judgments about its quality and their confidence in it, which could also influence subsequent behavioural engagement.

If we take the next step back in the chain of reasoning, using the first example above, you could suggest that without initial positive affective engagement, the user would not be cognitively engaged to experience usability issues and feel frustration. In the other example, we could ask how the connection with the app emerged; users likely developed a bond with the intervention based on affective and cognitive engagement. The difficulty with trying to unpick the order of affective and cognitive processes is their close integration. As the debate on the nature of emotion in psychology has demonstrated (highlighted briefly in [Section 3.6.2.5](#)), our understanding of how affective and cognitive processes are integrated on a neural and subconscious level is still developing. The key learning from this analysis is that both affective and cognitive engagement play a critical role in behavioural engagement and closely influence each other.

How can this debate be resolved to produce a final model positing the associations between different components across the process of engagement with DHIs? The value of the multiple case study design and cross-case analysis techniques is that insights can be generated beyond a single case and iteratively incorporated to build a refined model [337,459]. The hypothesised model (based on the literature) and the empirical models developed in this thesis were refined into a new model that takes into account the synthesis of the literature and empirical findings and is shaped by the investigation into how factors influence these various components. This is visualised in the revised model in Figure 6-4. The final model does not include a definitive answer to the question of order of influence of affective and cognitive engagement; given the disparity in the literature and case study findings, and the potential explanations generated from consideration of these differences in relation to theory and methodology, this order cannot be definitively determined with the evidence here. This model

highlights instead the close integration of components; [Section 7.4.1.2](#) discusses in more detail the implications of this unanswered question for the thesis and future research.

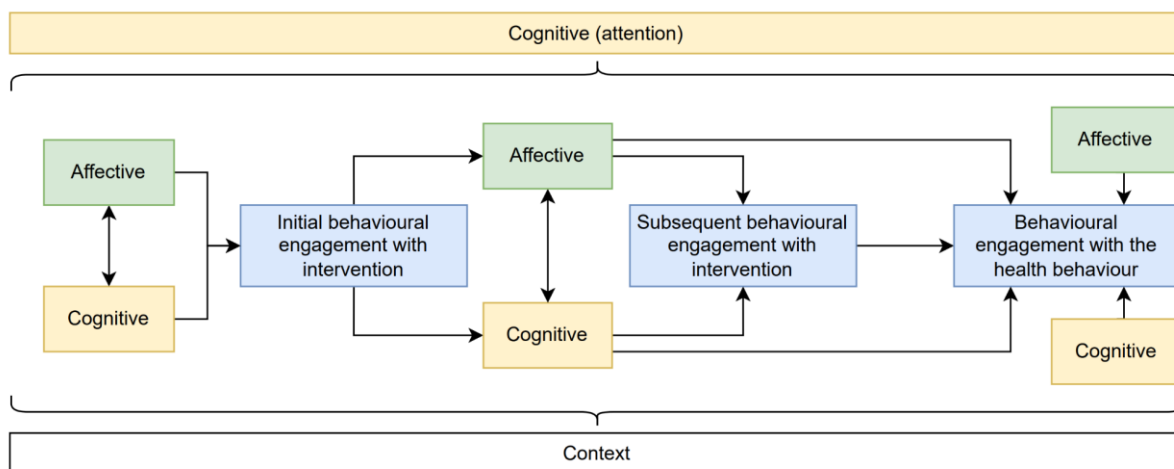


Figure 6-4. Proposed model of relationships over the process of affective, cognitive, and behavioural engagement with digital health interventions

6.6.3 Strengths and limitations

The main limitation of the cross-case analysis is the scope of comparability. The differences in methodology, research questions, and SSI topic guides meant that outputs varied. Although there was sufficient overlap to compare the cases with the data available, the only aspect of the cross-case synthesis that drew on findings from all three case studies was the exploration of the components influencing subsequent micro behavioural engagement. The analysis of hypotheses 1a-1c was limited to data from case studies 2 and 3, as the SSIs conducted in case study 1 were not designed to capture factors influencing initial engagement. The designs of case studies 2 and 3 were adapted to address this limitation and enable hypotheses 1a-1c to be assessed, but the synthesis would have been stronger had findings from case study 1 also been available.

This highlights how differences in the case characteristics limited the comparability to some degree - for example, engagement with the intervention and health behaviour were largely conflated in Wysa (case study 2). Likewise, as there was no longer-term behavioural engagement with Dora (case study 3), the analysis of hypothesis 3a-3b and 4a-4c was limited to case studies 1 and 2. This could not have been avoided unless a different intervention was selected for the third case study. Had a third DBCI been included in Dora's stead, this limitation would have been mitigated; however, the benefits of including different types of DHI would also have been reduced. Having cases with different characteristics was a strength

for the cross-case synthesis because the differences between the cases provided valuable insights and enabled an exploration of how the characteristics of different types of DHIs might influence the process of engagement. Although the number of cases is not intended to be statistically generalisable, synthesising a larger number of case studies would have provided more opportunities for literal and theoretical replication, which could have improved confidence in the conclusions and their transferability to other types and applications of DHIs.

As discussed in the limitations of the individual case studies in Chapter 5, another limitation is that the cross-case synthesis was only conducted by one researcher. This is a particular limitation for qualitative research, as interpretation is inherently influenced by researchers' own experiences and perspectives. Ideally, the cross-case synthesis would have been conducted initially by two independent researchers and then discussed collaboratively, providing an opportunity for assumptions to be challenged and explored further. Efforts to mitigate this limitation included revisiting the cross-case analysis again after a couple months to attempt to identify and challenge my own assumptions and by identifying and discussing potential alternative explanations for the key differences between hypothesised and empirical findings.

6.6.4 Conclusion

The cross-case synthesis provided the opportunity to investigate the generalisability of the themes and patterns of engagement generated in each case study and to explore in more depth how characteristics of the digital technologies or clinical cases could influence engagement with DHIs [337,459]. The explanation building process of developing a revised model based on the synthesis of the individual case studies and previous research highlighted questions around the value of conducting this mapping. The close relationship between all of the components, and especially affective and cognitive engagement, means that we must ask whether it is meaningful to try and pull apart and order the components of engagement. It could be argued that the components are so closely integrated that any attempt to capture the relationships in a model will be an oversimplification or that the representation of the relationship as bidirectional in the model avoids the question. While the definitions of each component might be somewhat artificial and there is a lot of overlap between them, I think there is value to be gained from the process of trying to tease them apart. The model provides a starting point for intervention developers to consider how users' will engage with their intervention to prompt consideration of these different aspects (which evidence has shown to be important in the overall process of digital health engagement) and how to incorporate strategies to support them in intervention designs, which will be explored further in the final

chapter. It also provides a framework for identifying how barriers might interfere with the engagement process - by examining barriers reported by users in co-production or early stage evaluations, developers could examine where in the process those barriers are occurring, what component(s) of engagement they are affecting, and how they could best be mitigated. It would be interesting for future research to examine the importance of order in engagement - whether it is meaningful to identify the order of affective, cognitive, and behavioural components to mitigate barriers or whether affective and cognitive integration are so closely linked that it doesn't matter which comes 'first'.

Chapter 7: Discussion, Strengths and Limitations, Implications, and Conclusions

7.1 Chapter Overview

The main aim of this thesis was to improve our understanding of how users engage with digital health interventions (DHIs) by examining how different components of engagement interact, how this is influenced by the characteristics of the DHI, and how we can use those insights to develop strategies to support engagement with DHIs. Chapter 6 synthesised the results of the three individual case studies described in Chapter 5 to compare the empirically-observed relationships between the different components of engagement in the three case studies and previous literature to generate a revised model and to determine what additional insights could be generated about how users engage with different types of DHI.

The research question outlined in Chapter 1 asked: *“How are the different components of engagement (affective, cognitive, behavioural, micro, and macro) related in digital behaviour change app interventions for weight management and mental health and in an autonomous telemedicine post-surgical follow-up assessment?”* The final chapter will bring together the insights from all of the previous chapters to summarise what has been learnt in exploring this research question and to discuss the implications of this thesis for theory, future research, intervention design, and clinical care. Through the process of developing and examining this research question, I identified a significant gap in our understanding of engagement with DHIs that has important implications for their impact. DHIs cannot have a positive impact on behavioural, health, or service-related outcomes without sufficient engagement to achieve their aims, but supporting this engagement is an ongoing challenge. Previous research broadly agrees that engagement is a multi-faceted concept including various components (e.g. affective, cognitive, behavioural), but evaluations of engagement with DHIs are often limited to simple behavioural intervention use metrics. We cannot reliably design interventions that support engagement without an understanding of why users struggle to engage.

To understand this, we need clear definitions of these different components and a good understanding of how they are related to and influence each other; this was lacking from the literature. This thesis aimed to address this gap by examining previous theory and literature about the components of engagement and synthesising those findings with three case studies that mapped the interactions between these components in three different DHIs. The analysis demonstrated that each component was important for the overall process of digital health engagement, despite substantial overlap and close integration. As a whole, the thesis

addresses the research question and the gap in knowledge by providing a framework that can be used for both intervention design and evaluation to guide thinking around how different components influence the overall process of engagement and how they can be supported to achieve target outcomes (Figure 7-1).

Chapter 7 will start by summarising the main findings, followed by an examination of the overall strengths and limitations of the thesis. The proposed framework and broader implications will be discussed, including highlighting unanswered questions and gaps for future research, proposing recommendations for how different components could be individually targeted to improve engagement with DHIs, and exploring how these findings could support health and care. The chapter concludes with reflections on what I have learnt about conducting evaluations of engagement as a multi-faceted concept in digital health and about research more broadly over the process of conducting my thesis and with a summary of the contributions of this thesis.

Publications: Chapter 7 refers to data and analyses that have previously been published, but none of the discussion presented in this chapter had been published when the thesis was submitted. Sections of Chapter 7 have since been prepared for publication; the reference and author contributions for this paper are listed below.

Milne-Ives M, Homer S, Andrade J, Meinert E. Mapping the process of engagement with digital health interventions: A cross-case synthesis. *Mayo Clinic Proceedings: Innovations, Quality & Outcomes*. 2025; Under Review.

MMI conceived the topic, conducted the analysis, and drafted and revised the manuscript. SH, JA, and EM supervised the research and contributed revisions.

7.2 Principal Findings

Given the previous research about strategies to improve engagement, the thesis began with a systematic review of existing evidence for how individual components of engagement could be supported by theory-based features, focusing specifically on BCTs. There was little research identified that specifically evaluated the associations between these BCTs and different components of engagement, although the review found evidence supporting the association of 9 common BCTs (feedback, self-monitoring of behaviour, goal-setting and action planning, prompts/cues, social support, rewards, information about health consequences, and instruction on how to perform the behaviour) with engagement more generally. It also highlighted the importance of mixed methods in digital health research, as standard measures of engagement such as system usage data are unable to capture affective

and cognitive elements of engagement. These were better captured using qualitative methods or questionnaires, but there was a lack of intentionality to examine these components and a plethora of measures used. This diversity in measures highlighted another gap in the literature - a lack of clear definitions of what the different components of engagement entail.

These are critical limitations of digital health engagement research; to grow our understanding of engagement, assess generalisability, and enable the efficacy of specific strategies to be tested, we need to be able to capture and compare the components of engagement across studies. To address this gap in the conceptualisation and measurement of engagement, Chapter 3 investigated two recently developed self-report measures that proposed to capture a multi-faceted conceptualisation of engagement. The comparison and analysis of these two scales supported the need for further clarity in the definitions and conceptualisation of the different components of engagement in scale development to improve our ability to capture those components. An examination of the scales' theoretical backgrounds, development, and items revealed several unanswered questions in our understanding of digital health engagement: whether engagement has distinct components, how those components should be defined, and - if distinct - how each can be measured. Comparing this analysis to broader literature around the associations between affective, cognitive, and behavioural elements underlined the substantial integration and overlap between the components of engagement. Despite this, I argued that there are practical benefits that can be obtained through investigating the components as distinct elements. Through the process of trying to tease these components apart, the ways in which they are connected can become more clear. This can provide the opportunity to identify barriers relating to these components at various points in the process of engagement, to enable intervention designers to incorporate engagement strategies targeted to address those specific issues, and to improve the ability of DHIs to support engagement within the context of their use.

The process of teasing apart the influence of the components was conducted in the analysis of the individual case studies and synthesised in the cross-case analysis. The results of each case study emphasised the importance of affective, cognitive, and behavioural engagement throughout the process of engagement with a DHI and identified factors that could influence these components. Synthesis of these individual findings with each other and previous literature highlighted the close integration of cognitive and affective engagement, to the extent that it was difficult to determine the order of influence (whether cognitive engagement influenced affective engagement or vice versa). There was a discrepancy between previous research and the current case studies on this 'order of influence' question; potential

explanations included their bidirectional influence at different times in the process of engagement, the methodology used to examine the associations in the case studies compared to the literature used to develop the hypothesised model (how participants are asked about their experience of engagement and how they order their recollections), and the context and characteristics of the particular influences on the pattern of engagement in question.

Within the broader picture of how the process of engagement progressed, the thematic synthesis of the three case studies generated four main themes relating to factors influencing these components of engagement: connection with the intervention, attention and interest, usability and accessibility for the target population, and confidence in the intervention. Within these themes, there were a variety of factors that influenced cognitive, affective, and behavioural components to shape user engagement. They present targets for intervention developers aiming to support engagement; design recommendations based on the theoretical and empirical evidence generated as part of this thesis will be explored in [Section 7.4.3](#).

7.3 Strengths and Limitations of the Thesis

7.3.1 Terminology and definitions issues

As a key focus of this thesis was the investigation of the associations between the different components of engagement, it was critical to establish clear definitions of what was meant by each of the components considered. These definitions were outlined at the start of the thesis ([Section 1.3.2](#)) and generated based on previous theory, definitions, and measures examined over the course of the first three chapters. The challenge identified during the systematic review (Chapter 2) was that there was a lack of agreed definitions for the individual components in the literature.

The definitions generated for this thesis were based on a variety of sources from the literature [42,43,45,46,49,62–70]; however, these sources were obtained through various readings around the theory and evaluations of engagement rather than through a systematic review process and they cannot be assumed to be final or comprehensive definitions of the components. This is a limitation because it introduces a source of bias into the development of the definitions and is worth exploring further because these definitions were used as the basis for much of the coding in the thematic analyses. While a number of papers were used to develop the definitions, it is possible that sources were missed that would have introduced additional detail, contradicting perspectives, or otherwise shaped the definitions that were generated. Unfortunately, it was not possible to conduct another systematic review to address this question given time constraints.

What is possible to assess is the potential bias, as an integrative review of definitions of engagement with mental health and health behaviour change interventions was recently published (April 2024), while I was revising the draft of my thesis [476]. While this review was not specific to digital health (it included in-person, digital, and blended interventions), it compiled and discussed the definitions of various components of engagement in its 40 included studies. They found that definitions of engagement predominantly focused on behaviour, which was variously defined in terms of uptake, adherence, usage, and effort (or active participation) with the intervention or beyond it [476]. Definitions of cognitive engagement captured a variety of elements, including understanding of the intervention, processes of attention, interest, motivation, and relation to the intervention, and assessments of the credibility, potential benefits, and potential difficulties of an intervention [476]. There were fewer definitions that captured the affective component of engagement, but these primarily focused on the “emotional experience with an intervention or the quality of the intervention experience” [476], although aesthetics, immersion, and care-provider relationships were also mentioned in some studies. In the model of the process of engagement that they developed, the authors synthesised their review into the following definitions:

Behavioural engagement: “inside-intervention effort - adherence (degree of actual intervention usage versus the prescribed usage by the intervention) plus effortful behaviour within the intervention - and outside-intervention effort (real-life skills-practice and behaviour change)” [476]

Cognitive engagement: “a subjective experience of attention and interest relating to the intervention in terms of credibility and expectancy” [476]

Affective engagement: “the subjective positive, negative, or neutral emotional experiences with the intervention” [476]

The definitions captured from the literature, and the final definitions Bijkerk *et al.* created for their model [476], are largely similar to the definitions I developed in [Section 3.6.3](#). This is particularly interesting as there was only marginal overlap in the papers cited in support of the definitions, but increases the confidence that the definitions are capturing their components well. It is possible that this similarity is due to many of the studies deriving their definitions from a few key papers (such as [42]). The two main differences between my definitions and those developed by Bijkerk *et al.* [476] were the inclusion of motivation and effort. Unlike Bijkerk *et al.*, I included the concept of motivation in my definitions of engagement, but like Bijkerk *et al.*, I also observed that it was unclear in the literature “whether motivation can be

considered a part of engagement or a related concept” [476]. In the face of this lack of clarity, Bijkerk *et al.* [476] and I have taken opposite perspectives. I included elements of motivation in my definitions because I wanted to ensure that these elements were captured in my analysis ([Sections 3.6.3](#) and [3.9.1](#)). Bijkerk *et al.* [476], while proposing a dynamic, bidirectional relationship between engagement and motivation, consider them as separate concepts. Given my argument about the benefit of considering the different components of engagement separately, despite their close integration, I see the value in Bijkerk *et al.*’s argument, particularly given the large body of research and theory established around the concept of motivation. This is an area where future research and further conceptual clarity is still required. The other main difference between our definitions is in the inclusion of effort in behavioural engagement. I think this is a difference in how effort is being conceptualised; Bijkerk *et al.*’s [476] definition appears to define effort as the performance of the behaviour, whereas I considered effort (associated with DHI specifically) as primarily cognitive, which is why I excluded it from my behavioural engagement definition.

The differences between the various multifaceted conceptualisations of engagement and the findings from this thesis about the close integration between the components (particularly affective and cognitive) mean that there is still work required to formalise our understanding and definitions of these concepts in digital health. The lack of clarity in definitions in the literature was a limitation for this evaluation, as it introduced potential ambiguity into the coding of components of engagement in the analyses. Having established definitions would have facilitated the process of mapping the components by increasing confidence in what factors were being deemed cognitive, affective, and behavioural. The benefit of the lack of established definitions was that it prompted careful consideration of how the components would be defined for this thesis and justification of key decisions in the coding and mapping of engagement components; this was a strength as it prompted potential alternative interpretations to be explored. Through the process of defining the engagement components, the thesis also highlighted a gap in the literature and provided an initial step for addressing it that can be built on in future work.

7.3.2 Value of case study research

A strength of the thesis was that the multiple case study approach enabled an in-depth investigation of how and why users engage with different types of DHIs [337]. This methodology was well-suited for exploring the research question of how the different components of engagement are related because it captured detail about users’ experiences with the intervention ([Section 4.2](#)). It also meant that the context of the cases could be

explored and compared to determine how certain characteristics of the intervention might influence engagement. As well as being a good fit for the exploratory and process-focused nature of the research question method, the case study methodology was well aligned with the topic of interest, as we need to understand how users engage with DHIs in the real world, rather than in experimentally-controlled conditions that may differ from actual use conditions. The limitation of using a case study design rather than an experimental design is that I could not test or draw conclusions about causal relationships between different components of engagement. Additional limitations relating to the case study design are discussed in the following sections.

7.3.3 Case selection and comparability

There were limitations associated with the selection of cases. The three cases were selected from a small pool of DHIs that we had won funding to evaluate, and were thus primarily a convenience sample. The research questions and primary aims of each study were different, and although all used mixed methods and assessed engagement, this meant that the execution of each case study varied. Ideally, this multiple case study would have been conducted with three dedicated case studies with the same aim and methodology, to improve the comparability of the results. The cross-case analysis would have been strengthened if one theoretically-based, pre-developed topic guide focusing on the components of engagement had been used in qualitative interviews for all of the case studies. As it was, some of the differences in findings could have been related to methodological differences rather than case differences; this has been discussed in the cross-case analysis where relevant ([Section 6.3](#)). Despite this, the sequential execution of the case studies was a strength, as it meant that learnings from the initial case studies could be incorporated into the design and execution of the subsequent ones. This helped shape the SSI guides particularly for cases 2 and 3 and resulted in more comprehensive and theoretically-based data for analysis. The analyses conducted may have been more limited had all three studies been designed and executed with the same methodology in parallel, as these learnings could not be incorporated.

The comparability of the cases did limit the analysis somewhat; the different DHIs, study aims, and topic guides generated meant that not all of the case studies captured all of the points of engagement. The findings around micro affective, cognitive, and behavioural engagement with the intervention could be synthesised across all three studies, but only two of the case studies (2 and 3) captured data about factors influencing initial micro engagement and the data about macro engagement with a health behaviour was limited to the two behaviourally-focused interventions (cases 1 and 2). The strength of the cases selected is that,

although comparability over the whole process of engagement was limited, the evaluation of different types of DHI enabled an exploration of which factors are likely to be generalisable across digital health as well as how various intervention characteristics might shape users' process of engagement.

7.3.4 Generalisability

The issue of generalisability or transferability is particularly important for the contribution and impact of the research. The use of a multiple case study design, rather than a single case study design, was a strength, as it enabled an examination of the extent to which the findings might be applicable to other DHIs or contexts [477]. Unlike quantitative or experimental studies, the aim of case study research is not to achieve statistical generalisability [337]; rather, the purpose of examining these cases was to generate insights to shape theoretical propositions about how users engage with DHIs. Including three distinct cases helped to enhance the generalisability of the findings beyond what would have been possible with a single case study because it incorporated contextual factors in the analysis and enabled the comparison of different cases to highlight similarities and differences between them. Considering how these similarities and differences might be shaped by the various characteristics and contexts of the DHIs examined allowed hypotheses to be generated about how the findings from these three cases might be applied to a broader model of engagement with digital health.

Although the use of multiple cases helped strengthen the transferability of the findings, generalisability is still limited. The cases represented only two types of DHIs (mobile health apps and AI conversational agents) and three clinical areas. Although there was general alignment of the mappings of engagement in the three cases, the inclusion of additional (or different) case studies may well have produced different results. Based on these strengths and limitations, the framework presented here is intended as a hypothesis for guiding further investigation of engagement with DHIs more broadly - based on in-depth, real-world investigation but requiring further evidence to extend the generalisability and confirm proposed associations and order of influence.

7.3.5 Validity and reliability

Although the thesis used mixed methods, the primary analyses of the case studies were qualitative. In qualitative research, the accurate reflection of the data (validity) and the trustworthiness or consistency of the analysis (reliability) [477] is complicated by the interpretation required and the unavoidable influence of a researchers' own perspectives and

biases. The qualitative analyses in this thesis have been conducted using a variety of strategies to improve the trustworthiness of the findings [477]. All of the initial case study coding and thematic analyses involved multiple researchers, which enabled individual interpretations to be challenged and to evolve. Additionally, I was extensively engaged in the data over prolonged periods, including conducting the majority of the interviews included in this analysis, and included in the appendices how the themes evolved over the course of the analyses. A diversity of participant quotations were included in the results to support the interpretation of the results with data.

Despite these efforts to strengthen the validity and reliability of the analyses, there were also limitations. The thesis-specific engagement mappings, cross-case analysis, and thematic synthesis were conducted by myself alone, and the lack of additional viewpoints to challenge my perspective is a limitation. This was mitigated through engagement with the data and review of the analyses and findings over a long, separated period of time, to enable it to be considered again through fresh eyes.

Where possible, qualitative data was triangulated with questionnaire and system use data to provide an additional source of evidence to contextualise the qualitative analysis, but the ability to triangulate was often limited based on the quantitative data available. Although each case study used mixed methods, the cross-case analysis had to rely primarily on qualitative data. The analysis would have been strengthened if that qualitative data could have been triangulated with quantitative data regarding engagement. App use data was collected in cases 1 and 2, but the available data was limited, different across the studies, and did not substantially improve the analysis. Although it was not possible to modify the system use data that was captured in these studies, future research would benefit from more detailed app use data - particularly information about which elements of the intervention a user is engaging with - to compare with user feedback about engagement.

If I was redoing the study, I would also include both the DBCI-ES and the TWEETS as measures in the case studies. This could not be done in this study as, by the time Chapter 3 had been conducted and identified the most appropriate self-report measures, the case studies were already underway. Including these measures could also have added to the investigation by enabling an evaluation of the strengths and limitations of each scale for addressing this type of research question. The limitations of each scale highlighted in [Section 3.9.1](#) would likely have limited the value of the questionnaire responses for mapping the components of engagement. Including these measures would still have strengthened the analysis by providing a source of quantitative data specific to engagement that could be captured from a larger

sample and triangulated with the qualitative data to either increase confidence in the findings or highlight discrepancies requiring further consideration. Responses to individual items could have been considered in the synthesis with the qualitative data and potentially generated additional insights regarding barriers to engagement across the different DHIs at different time points.

7.3.6 Other methodological issues

7.3.6.1 Engagement mapping analysis

There were several limitations with the engagement mapping analysis; these have been discussed in [Sections 5.2.6, 5.3.6, and 5.4.5](#), but are worth mentioning here as this was a key analysis for the thesis. Unlike the initial thematic analyses for each of the case studies, the thesis-specific engagement mapping analysis was conducted by only one researcher. This is a major limitation for qualitative analyses because there is a lack of external perspective to challenge assumptions made during the interpretation of the data. This was of particular concern for this analysis because of the lack of agreed conceptualisation and definitions of the different components of engagement. I attempted to mitigate these limitations in a couple of ways: first, I referred back regularly to the definitions of affective, cognitive, and behavioural engagement that I developed in [Section 3.6.3](#). These definitions were created from a variety of published theory and literature around engagement, and as demonstrated in [Section 7.3.1](#) are aligned with a subsequent, independently-generated set of definitions [476]; while this does not make them definitive, it helps to improve confidence in the association of particular components of engagement with case study themes. Additionally, while there may be alternative perspectives of how each component should be defined, describing the definitions clarifies the interpretation of the components in this analysis, increasing transparency.

The second strategy I used to mitigate the bias of being the only researcher conducting the analysis was to conduct the analysis twice, over three months apart. While this cannot achieve the same level of independence of having two researchers interpreting the data, it provided me with an opportunity to re-analyse the data with fresh eyes, compare my two analyses, and further interrogate any differences before generating my final interpretation. In this way, I aimed to mimic the process of independent then collaborative coding and theming of data in standard thematic analyses. The third strategy was substantiating the interpretation with data. By including participant quotes in the thematic frameworks developed in this analysis, I provided evidence to justify my interpretation of the data.

These strategies are not foolproof, and it is likely that with the inclusion of additional researchers and perspectives, the interpretation of the data and patterns of engagement generated might evolve further. Indeed, I hope that this will be conducted in future research, where the model I propose mapping relationships between the various components of engagement can be further explored and refined.

7.3.6.2 Presentation of thematic analyses

Presentations of thematic analyses typically follow a structure that is organised into sections and subsections by theme and subtheme [397]. Case studies 2 and 3 follow a version of this structure - they are organised by high level themes to present the most interesting and relevant results, use thematic maps to provide a visual overview of their associations, and incorporate participant quotes to strengthen the narrative of the data [397]. Compared to the presentation of typical thematic analyses, however, the results presented for these analyses may appear to lack detail; many thematic analyses present narratives for each subtheme with several, detailed quotes to enrich their descriptions (e.g. [111,478–480]). The reason that these analyses are somewhat abbreviated in Chapter 5 is because the main focus for the thesis was the mapping analyses and their synthesis in Chapter 6. Rather than exploring each theme and subtheme generated from the original case study thematic analyses in detail, the key findings and narrative summaries are presented to set the stage for the thesis-specific analyses. This is a limitation in that the individual thematic analyses are not fully explored in detail, but was necessary in the context of the thesis. To mitigate this limitation, full thematic frameworks and additional participant quotes are included in the Chapter 5 appendices to provide more detail.

Compared to case studies 2 and 3, case study 1 presents the results in a less traditional manner for a thematic analysis. This was because there was an additional analysis conducted within case study 1 after the thematic framework was finalised. As discussed in [Section 5.2.2.2](#), this was conducted to address the case study research question, which focused on barriers and facilitators (factors influencing) engagement and perceived impact. I determined that the presentation of results structured based on this subsequent analysis would be more relevant for readers looking to develop or evaluate DHIs and better answer the research question than the more traditional presentation of results structured by theme and subtheme. The limitation is that this is again a step away from the standard thematic analysis method, which increases potential confusion about how those insights were derived. I have aimed to address this by detailing the steps used in the analysis in Chapter 4 and including a table detailing how the initial thematic framework was mapped to the COM-B model and factors in [Appendix 5-B](#).

7.3.6.3 Reflexivity

Throughout the process of my thesis, I have reflected on and aimed to justify the methodological decisions I have made and how I have interpreted the data and included a positionality and reflexivity statement reflecting on how my personal characteristics and experiences have influenced my research ([Appendix 7-A](#)). Through the process of developing this statement, however, I have learned that a more structured approach to reflexivity throughout the process of conducting my thesis would have been more valuable than only formally reflecting at the end [481]. Keeping a more detailed research diary outlining not just my process of research decision-making but my reflections on experiences of interviews and how my perspectives might shape my methodological choices and interpretations of data would have better enabled me to challenge my assumptions and strengthen the trustworthiness of my findings.

7.4 Research, Practical, and Clinical implications

7.4.1 Research implications

7.4.1.1 Theoretical contribution to knowledge

This thesis makes a theoretical contribution to knowledge that builds on a substantial body of previous theory and research. The literature has already established that engagement is an essential aspect to consider when evaluating DHIs ([Section 1.6.1](#)); this is demonstrated by the inconsistent evidence of effectiveness of DHIs [2,116,171,172] and findings that a lack of engagement is a key barrier to the potential impact of such interventions [12,14,31,33]. The growing awareness of the importance of engagement for DHIs has been reflected in the increasing focus on engagement and has resulted in a body of literature examining a variety of factors associated with engagement (including user characteristics, design features, active ingredients, and context - eg. [122,166,185,204,226,482]). Despite these advances in knowledge, there are many gaps remaining in our understanding of digital health engagement and how to support it. While there is general agreement that engagement is a multi-faceted concept [14,42,45,49–51], there is a lack of research examining non-behavioural components of engagement (Chapter 2) [14,44,45], a lack of clarity in the conceptualisations of the specific components (Chapters 2 and 3) [47], and a lack of guidance on how to investigate the different components of engagement in a way that is comprehensive, consistent, and comparable (Chapter 3) [38]. To address the problem of insufficient engagement, we need to understand why people struggle with engagement - what factors influence engagement, why and in what circumstances, and how engagement with the intervention and the target health

behaviour(s) influence each other over time. To build up a strong, generalisable understanding of how to support engagement with DHIs, we need to look beyond behavioural measures of engagement and we need a framework for how to investigate engagement as a multi-faceted concept.

That is this thesis’ theoretical contribution to research: a theory- and evidence-based framework that represents a step forward in our understanding of how the different components of engagement relate to each other and shape the overall process of engagement with DHIs as well as how they can be influenced by context and other factors (Figures 6-4 and 7-1). This provides a starting point for future research to test and refine, but even in an early stage, this analysis adds value by providing guidance for how future studies can think about the different components of engagement and incorporate them into their data collection and by creating a structure for exploring the relationships between components of engagement and particular factors that can influence them. An improved understanding of how different components of engagement interact will enable us to design better DHIs that target particular and personalised barriers to engagement and to improve intervention’s impact on behavioural, health, and other outcomes. The thesis also highlights key gaps that still need to be addressed, which will be discussed in the following section.

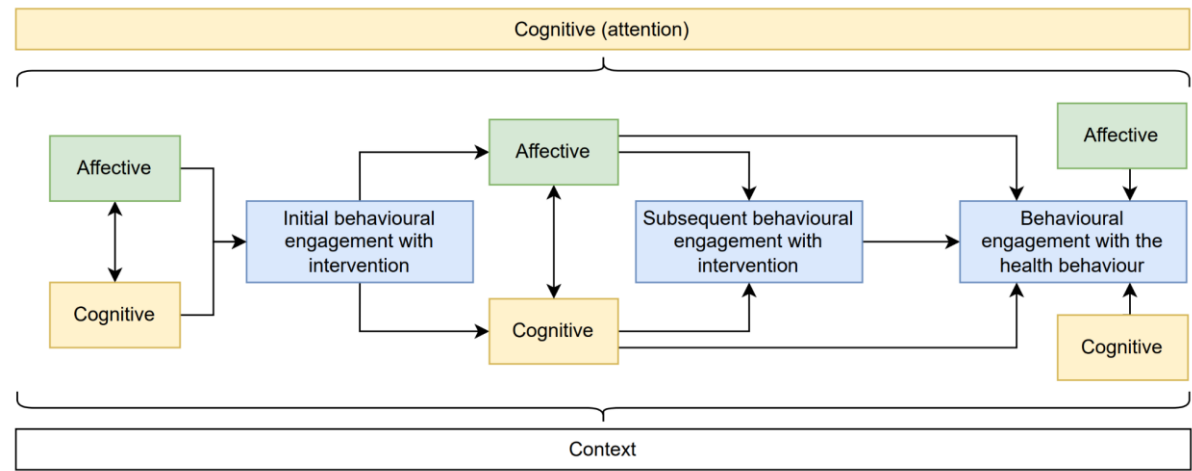


Figure 6-4 (duplicate). Proposed model of relationships over the process of affective, cognitive, and behavioural engagement with digital health interventions

7.4.1.2 Unanswered questions and future directions

A key contribution of this thesis is in the identification of areas where more research is needed to enable a better understanding and evaluation of digital health engagement. The first is the need for the development of a questionnaire measure that enables examination of distinct components of engagement. Chapter 3 highlighted the strengths and limitations of the

two available measures of digital health engagement as a multi-faceted concept; the DBCI-ES captured well the experiential aspect of engagement with DHIs but has questionable validity for the behavioural component, does not enable affective and cognitive engagement to be considered separately, and is focused solely on micro engagement. The TWEETS provides a good holistic measure of engagement as a multi-faceted concept, but despite including different items for affective, cognitive, and behavioural engagement, there was theoretical overlap between the items and micro and macro engagement were blurred together. The scales both represent improvements in the available measures and could nicely address particular research questions, but there is still a need for further development to enable engagement components to be examined separately.

A scale that attempts to comprehensively capture and distinguish between the components of engagement would come with both benefits and challenges. The ability to examine components separately is important for research that aims to understand how particular strategies influence engagement or what particular barriers users face to engagement. It would be a valuable source of quantitative data to triangulate with qualitative data or enable these questions to be examined if qualitative methods are not feasible. Such a scale could also help facilitate comparability between evaluations of engagement of different DHIs. This does not mean that such a scale should be the universal measure of engagement; the diversity of DHIs and the potential questions around the complex concept of engagement means that there will be value in having a variety of measures available. Developing a new measure would also face the challenges identified around the overlap between affective and cognitive components; the DBCI-ES did not try to separate them and EFA of the TWEETS found that, despite including them as different components, they all loaded onto one factor. Further research is needed to explore whether and how these components could be distinguished in quantitative measures.

Another key question is how important it is to understand the order of influence between affective and cognitive engagement. This was an area of investigation in this thesis that could not be satisfactorily resolved; the hypothesised model, based on previous theory and studies, suggested a pattern of influence of affective engagement on behavioural engagement via cognitive engagement, whereas the investigation of the thesis case studies observed a pattern of influence of cognitive engagement on behavioural engagement via affective engagement. [Section 6.4.3](#) discussed potential explanations for this difference between the hypothesised and observed mapping of relationships between the components, but did not have the data to reach evidenced conclusions. Given the close relationship between affective and cognitive

engagement, establishing a definitive order of influence may not be possible or may not add value - e.g. the relationship could go in either direction depending on circumstances or characteristics; strategies to address barriers to each might be necessary to support engagement - but this area is worth exploring further to determine if there are circumstances in which the order of delivery of strategies to support affective or cognitive engagement has a material difference in terms of impact on behavioural engagement with the intervention and target behaviour.

For example, the influence of affective and cognitive engagement could vary at particular points in the process of engagement - maybe affective engagement from initial impressions of a DHI drives initial cognitive and behavioural engagement while cognitive engagement with interesting intervention content is especially important for supporting continued positive affective engagement and sustaining longer-term behavioural engagement. This could be explored in the context of the broader behaviour change literature. There are a large number of theories of behaviour change: a 2016 systematic review specifically investigating theories of maintenance of behaviour change included 100 theories [483], including prominent theories such as Social Cognitive Theory [484], Self-Determination Theory [316], the Theory of Planned Behaviour [485], and the Transtheoretical Model [486], among others (see *The Handbook of Behaviour Change* for descriptions of these and other theories [487]). This might provide further theoretical context for exploring the contribution of different components at different stages in the process of engagement with the intervention and health behaviour change. Alternatively, the importance of affective or cognitive engagement could depend on what matters most to individual users or on the purpose of the intervention. Learning more about the factors influencing the relative importance of each component at each stage in the process of engagement with the intervention and target health behaviour will further help improve intervention design and implementation.

In terms of intervention design, another open area of investigation is the relationship between the components of engagement and particular features and BCTs incorporated in DHIs. There is existing evidence for the efficacy of certain features at supporting engagement and many themes generated in this analysis align with previous findings - e.g. usability, tone and credibility influencing trust, visual aesthetic [122,204] - but understanding the psychological basis for why these aspects support engagement will be useful to ensure that they are incorporated in a way that meets users' needs. Beyond existing strategies, there is the potential for the identification and development of new means of supporting engagement; the BCTs identified in Chapter 2 as having associations with engagement were also some of the

most commonly used BCTs. It is likely that there are other BCTs that could influence engagement but haven't been incorporated in sufficient DHIs to demonstrate evidence of their impact on engagement. A useful next step would be to conduct a theoretical analysis of which BCTs might be related to which components of engagement. This could build off existing research being conducted with the Human Behaviour Change Project [488]; its Theory and Techniques tool maps BCTs to mechanisms of action and would provide a good starting point for theorising about how particular BCTs might support particular components of engagement. The relationships with the strongest theoretical justification would then be good targets for empirical investigations to determine the impact of the BCTs on engagement. I have suggested BCTs as a starting point because they are well-structured and described and are being increasingly incorporated into DHI design; however, research on strategies to support engagement should not be limited to BCTs. Other features of app design and context have proven important factors for engagement, and psychologically-based strategies to support engagement around these aspects should also be considered. A comprehensive framework of how particular strategies could address particular barriers would be a valuable tool for DHI developers, especially if it captures how different components and strategies influence engagement at different points in the process. The next section provides a summary of recommendations that can be derived to support intervention design from this current work.

7.4.2 Practical implications: recommendations for digital health intervention development

Both the systematic review and the cross-case synthesis identified evidence for the ability of particular design features and BCTs to support engagement with DHIs. This section summarises these findings to provide recommendations for means of designing interventions that avoid key barriers to engagement and include evidence-based features to support it. An intervention needs to be sufficiently affectively and cognitively engaging to prompt initial behavioural engagement and to sustain sufficient behavioural engagement with the active ingredients of the intervention to achieve its aims. Factors that the systematic review and case studies identified as influencing cognitive and/or affective engagement included context, BCTs, and other design and content features (Table 7-1).

7.4.2.1 Context

Before addressing particular features incorporated into the interventions, which will be the main focus of these recommendations, we must first consider why users' initially pick up and try an intervention. This will depend on the context and type of the intervention - users may

come across a mobile health app on an app store or other online platform, it may be recommended to them by a friend or clinician, or it may be delivered as part of a clinical care pathway. Social and other contextual factors play a role here - in line with previous research [431,489–491], the Wysa and Dora case studies demonstrated that the providence of an intervention from an official healthcare setting can support engagement by increasing the credibility of, and trust in, an intervention.

Other contextual factors can influence subsequent behavioural engagement with the intervention and target behaviour. Accessibility was a key factor that was highlighted across the three case studies, but how it related to engagement varied depending on the type of intervention. Basic accessibility considerations such as language ability and vision or hearing difficulties can make an intervention completely inaccessible to some potential users, but other accessibility considerations are more subtle. Case study 3 highlighted how neurodivergence and specific clinical characteristics could influence cognitive and affective (and thus potentially behavioural) engagement with the intervention. Case study 1 demonstrated how socioeconomic factors (e.g. affordability, local availability of required resources) could be a barrier to behavioural engagement with the target health behaviour, even if a user is cognitively and affectively engaged.

7.4.2.2 Behaviour Change Techniques

Regarding the association of BCTs with engagement, the empirical findings from the case studies supported the exploratory analysis conducted in the systematic review. The five most common BCTs associated with engagement from Chapter 2 (feedback, self-monitoring, goal setting, prompts/cues, and social support) were all identified as important factors influencing engagement in case study 1. Interpretation of user feedback suggested that many of these BCTs helped support cognitive and affective engagement and motivation to behaviourally engage with the intervention and the target behaviour. Case studies 2 and 3 did not include all of these BCTs, but also identified an important role for the ‘prompts/cues’ BCT. The associations between the particular BCTs highlighted in the case studies and engagement are described in Table 7-1.

7.4.2.3 Other design and content features

Regardless of context and beyond specific BCTs, an intervention’s design and mode of delivery can have a significant influence on engagement. An intervention that is visually appealing and appears to address an unmet need of the user is likely to increase cognitive and affective engagement and encourage initial behavioural engagement. In terms of sustaining

engagement, all of the case studies highlighted the importance of a positive, supportive, non-judgmental tone and of an interface that is easy to use and navigate. Personalisation and customisation of the intervention could also support affective and cognitive engagement and influence the success of many other features. For example, case studies 1 and 2 demonstrated that the ability to schedule reminders that were tailored to a user's personal schedule could be the difference between notifications supporting cognitive engagement or reducing affective engagement. Another key factor influencing confidence in, and engagement with, an intervention that was highlighted in the thematic synthesis is credibility; this was already discussed in [Section 7.4.2.1](#) in terms of how the method of provision of the intervention can influence credibility, but case study 2 demonstrated that credibility can also be supported through the inclusion of reliable sources in the intervention and through the provision of clear privacy and security information. Beyond the design of the intervention, novelty and variety of content provided - whether this be health education and suggestions or AI conversations - was important for supporting cognitive and affective, and thus subsequent behavioural, engagement.

7.4.2.4 Summary of features and associations with engagement

Table 7-1 provides an overview of the features generated from the systematic review and cross-case analysis (highlighted in the previous subsections) and considerations for how these factors could influence the different components of engagement. This is not an exhaustive list, but includes recommendations and cautions based on the findings of this analysis. It is also worth noting that, despite these general recommendations, what will appeal to the target audience of an intervention is likely to vary and should be explored through user-centred design and co-production.

Table 7-1. List of features and considerations for incorporating them

Feature	Considerations for incorporating into intervention design	Evidence
<i>Context</i>		
Accessibility	The needs of users with different characteristics should be considered during the development of any intervention as this can influence all three components of engagement with both the intervention and the target health behaviour. Where possible, strategies for improving the accessibility of these interventions should be developed in collaboration with people with lived experience as accessibility barriers are likely to be related to the specific context of the target population and the intervention characteristics.	Cases 1-3

<i>Behaviour Change Techniques</i>		
Goal setting and self-monitoring of behaviour	Goal setting and self-monitoring of behaviour can support motivation to engage with the intervention (cognitive and affective engagement) and target behaviour by increasing awareness of a user's current behaviour and highlighting discrepancies with their desired behaviour. Designs incorporating goal setting should enable personalisation of goals to align with users' abilities to avoid triggering negative emotions and disengagement if the goal is too difficult [276].	Systematic review, Case 1
Feedback	Delivering encouragement to users can support positive affective engagement and providing advice on adapting goals or strategies to achieve them can support users' cognitive engagement with an intervention.	Systematic review, Case 1
Social support	Social support, depending on how it is delivered in an intervention, has the potential to support affective and cognitive engagement by providing a source of motivation to engage with the intervention. For example, social interaction with other users could make an intervention more enjoyable to use. However, social comparison, which might occur as a result of creating a digital community within an intervention, could trigger affective disengagement if users feel demotivated by the comparison or judged by others.	Systematic review, Case 1
Rewards	Previous research has highlighted the association of rewards with fun and motivation; however, to support affective and cognitive engagement, the rewards need to be meaningful to the user. One means of achieving this could be to enable users to choose their own rewards and track their progress towards them or to have a digital rewards system that unlocks new content or features (e.g. for children, new avatars or personalisable features).	Systematic review, Case 1
Prompts/cues (reminders)	Reminders are a key means of supporting cognitive engagement by drawing users' attention to the intervention. However, badly-timed or overly frequent reminders can cause frustration and irritation, which can lead to affective disengagement. To avoid this, users should be able to personalise reminders to the amount and time that best suits them.	Systematic review, Cases 1-3
<i>Design features</i>		
Aesthetic design	Intervention designs should be visually appealing to support affective engagement - although preferences will vary, this broadly includes simple, cheerful styles and colours that are enticing but not overwhelming to look at.	Cases 1 and 2
Tone	The intervention's tone is an important factor in users'	Cases 1-3

	connections with it; it should be positive and non-judgmental. In conversational agents, a brisk, impersonal, and mechanical voice or tone could lead to affective disengagement. While clarity is important, developers should also be careful not to be overly simplistic or appear condescending. Tone can be conveyed by the written or verbal content of the intervention and by visual elements (e.g. colour scheme, avatars, etc.).	
Personalisation	Personalisation is a key element that influences many of the other factors and can help support users' sense of connection with an intervention. Personalisation should be easy to set up and be as flexible as possible, but optional, so that there is not an excessive burden on users who are just wanting to try out the intervention or start using it quickly.	Cases 1 and 2
Usability and guidance	A simple, easy to use intervention that has (or does not require) clear guidance can help support cognitive engagement. A time consuming or overwhelming intervention that requires a high degree of attention and cognitive resources to use and navigate may lead to cognitive disengagement.	Systematic review, Cases 1-3
Credibility	Trust and confidence in the intervention can be supported through the use of credible sources (in terms of the intervention content and source).	Cases 2 and 3
Privacy and data security information	Potential concerns around data security and privacy can hinder behavioural engagement; these could be alleviated by providing easily comprehensible information about policies to improve the credibility of the intervention.	Cases 1 and 2
Gamification	Gamification can support affective and cognitive engagement by providing a fun, enjoyable, and rewarding experience. However, it has similar potential issues as goal setting and social support where excessive difficulty or comparison could decrease motivation and lead to disengagement.	Systematic review, Case 1
<i>Other content-related factors</i>		
Interest and novelty of content	Although the intervention should be easy to use and navigate, it should include a large amount and variety of content to sustain users' interest and support cognitive and behavioural engagement. A basic intervention without new content to consume can cause boredom and potentially lead to disengagement.	Cases 1 and 2
Suggestions	Suggestions (e.g. for goals or target behaviours) can support affective and cognitive engagement and macro behavioural engagement by reducing the cognitive effort required to engage or by helping frame the target behaviour as enjoyable.	Case 1

Conversational agent capability	For interventions that include a CA, its capability (as well as its tone) is a key factor that can influence affective, cognitive, and behavioural engagement. Repetition, inconsistencies, inaccuracies, and shallow responses can cause frustration and decrease users' confidence in the ability of the intervention, which can lead to disengagement.	Cases 2 and 3

7.4.2.5 Comparison and contribution to the literature

Comparison of the factors generated from this thesis' analyses with recent literature demonstrates good alignment. An evaluation of design features that users' perceive to support initial and subsequent engagement with alcohol and smoking reduction apps highlighted the importance of an app's aesthetic appeal, usability, and perceived credibility ("social proof" as indicated by ratings, reviews, and recognition) on users' decision to try it [204]. In terms of supporting continued engagement, that evaluation also found that users felt goal-setting, monitoring, feedback, rewards, reminders, a positive and non-judgmental tone, personalisation, credible content, and clear privacy information were important factors [204]. Similar features were identified as important in a systematic review and thematic analysis of a broader range of mobile technologies (app, website, and text message interventions for a variety of health issues): an appealing and simple aesthetic, easy navigation with guidance, reinforcement through rewards and reminders, the ability to connect with healthcare professionals and other users, simple and positive delivery of content, credible sources and privacy information, and a variety of features to personalise the experience (e.g. monitoring and incorporation of user preferences and behaviours to provide personalised content and feedback and the ability to tailor the app to their needs) [122]. This review also provides a handy checklist for intervention development specifically relating to design features that can support user engagement [122]. Similar conclusions about key design features (including BCTs) were also drawn in a systematic review of design features that enhance engagement specifically with physical activity interventions for young people [492] and in the framework developed as part of Perski *et al.*'s systematic review where they defined their conceptualisation of engagement [42].

Given the various previous analyses of associations between design features and engagement and the high degree of similarity between those analyses and this thesis, one might ask what the value is in re-examining this topic. There are two main ways in which the inclusion of this analysis of features in this thesis represents a contribution to the literature. Firstly, the existing evidence for similar features in the literature improves confidence in the generalisability of the importance of these features for supporting engagement across different types and clinical

applications of DHIs. Further findings supporting these associations in different cases only helps to build the body of evidence. Secondly, this thesis has advanced the current state of the literature by explicitly exploring how these features relate to the different components of the engagement process. Conducting this analysis enabled the factors identified in the various case studies to be mapped to the proposed model mapping the relationships between different components of engagement (Figure 7-1). While this is still preliminary and could be further developed with future investigations and testing, it provides a guide for developers to identify where users are experiencing barriers to engagement and what particular factors might be most relevant at that stage to help support a particular pattern of engagement.

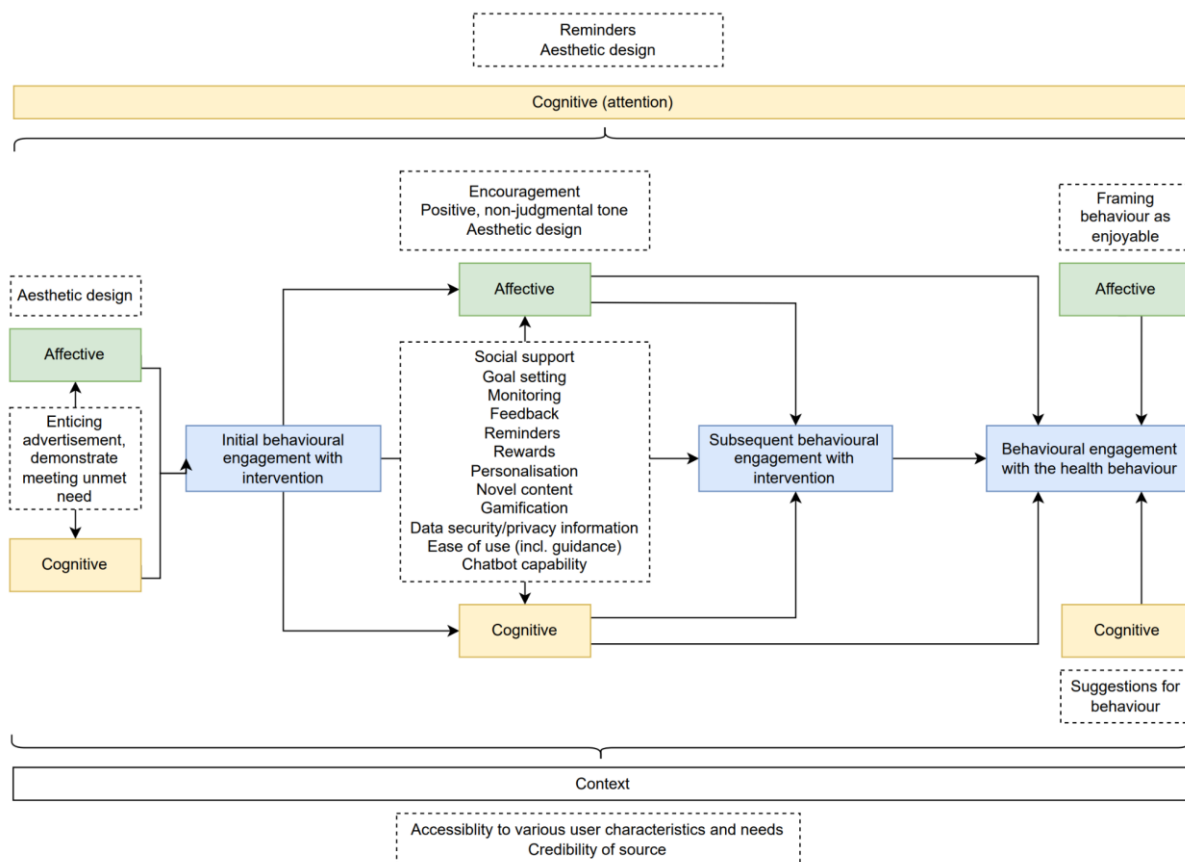


Figure 7-1. Framework of how various factors can influence the relationships between different components of engagement

7.4.3 Clinical implications

The clinical need for better supporting engagement with DHIs (detailed in [Sections 1.4](#) and [1.5](#)) relates to the need for solutions to address the gap between clinical resources and demand [15,16]. As the previous sections have highlighted, advancement of our theoretical understanding of digital health engagement can have practical implications for the design of

DHIs. If DHIs can better support ‘effective engagement’ [14], this could have valuable implications for public health and clinical care in several ways.

One potential clinical implication is improved patient self-management and empowerment. If patients engage more effectively with DHIs, we would expect to see improved behavioural and clinical outcomes. Empowering patients to take greater responsibility for their health has been associated with increased self-efficacy, which has in turn been associated with increased motivation and behavioural adherence [493]. Therefore, we can hypothesise that supporting engagement will increase the likelihood that an intervention will achieve its outcomes. If so, this would likely snowball into other positive implications for clinical care; if patients are better able to self-manage their health and engage in preventive health behaviours, there may be a decrease in clinical burden and a freeing of clinical resources to focus on patients with more complex conditions.

In the case of interventions such as Dora, successful adoption and engagement with AI-based clinical interactions could improve clinical efficiency. Automating routine tasks also has the potential to reduce clinical burden and enable clinicians to spend their time engaging in high-skill tasks that have greater patient benefits.

Though more indirectly linked, another potential clinical implication is the advance in knowledge for measuring digital health engagement. The thesis supports this in a couple ways: by comparing existing self-report measures, by providing guidance on the potential for the development of new scales, and by providing a framework mapping the different components over the process of digital health engagement that can guide specific qualitative or quantitative investigations. Better evaluations of digital health engagement in research will help build a body of knowledge about what types of DHIs are most effective at supporting engagement, which could facilitate healthcare professionals or systems in adopting or recommending high quality interventions to patients. The thesis findings suggest that this in turn could help support engagement with those interventions.

While the clinical implications are a step removed from the specific findings and conclusions of this thesis, they shape both the rationale and the key long-term impact of this research. The purpose of examining in detail the theoretical considerations of how we conceptualise, measure, and aim to support engagement is to enable improved clinical care and patient benefit on a population level.

7.5 Reflections and Lessons Learnt

Beyond the specific learnings that derived from the findings of the thesis, I would like to include a brief section reflecting on what I have learnt about research through the process of conducting this thesis. In developing this section, I considered my skills, knowledge, and perspectives as a researcher now compared to when I started my PhD, how and why the plan for my thesis evolved over the course of my PhD, and what lessons or advice I would share with a new PhD student just starting out.

One of the most important lessons that I learnt during the process of conducting my thesis was the importance of flexibility in research. I started my PhD with a research plan developed for my PhD proposal, which aimed to examine the impact of specific BCTs on effective engagement with DHIs for childhood obesity. While my overall thesis topic remained similar - exploring engagement with DHIs - the specific research question and method of execution evolved substantially over the course of my PhD. There were a couple key reasons for these changes; the first was the natural evolution of my idea as I learned more about the current state of the field. My initial plan was to evaluate the NoObesity app, re-design it incorporating key BCTs selected based on feedback from the initial study, and compare the direct and indirect effects of those BCTs on engagement in a subsequent study using a factorial randomised controlled trial design. As I conducted my systematic review and learnt more about the theory around engagement, I found that conducting this analysis might be skipping a few important steps. My systematic review aimed to examine what was already known about the associations between BCTs and the different components of engagement, but what I found was that this had rarely been intentionally explored. While my initial planned project would have contributed to this gap, my other finding from the systematic review - that further research was needed to address the lack of clarity in the conceptualisation and measurement of the different components of engagement - provided an alternative gap for exploration.

Another key reason for the evolution of the project was logistical. The redevelopment of the NoObesity app, which was a necessary prerequisite for the subsequent planned evaluation, was delayed. This could have been a significant barrier to my thesis, as it would have put my research on hold for a long period. In exploring how to continue my thesis in view of this, my primary supervisor suggested I consider using a multiple case study method to address my research question. The process of changing my original plan for my thesis - including my main research question and methodology - highlighted to me the importance and value of flexibility. Through the learnings I had gathered from my systematic review, I developed a new research question and thesis design that brought together three case studies to explore an

important gap in our theoretical understanding of engagement. Being flexible with my initial idea enabled me to adapt to what I was learning from the literature and turn a challenge into a new opportunity. In the end, I am grateful to have experienced these logistical difficulties, as it prompted me to reconsider my plan and explore how else I could make a contribution to literature in my area of research. This was a particularly valuable experience for me, as I have a tendency to want to stick to my plans, but doing so would have missed the opportunity to explore the relationships between the different components of engagement and bring together the findings from different cases, which I believe has strengthened my analysis and contribution to the literature.

Another major way in which I have developed as a researcher through the process of conducting my thesis, similar to the previous learning about flexibility, was in learning to work iteratively to balance productivity and perfection. For example, my initial way of writing would have been to sit down and try to draft the perfect paper or chapter. I found this way of working challenging and slow, because I would get stuck on a particular sentence or section and struggle to progress, leading to delays in the overall process of the research. Over the course of my thesis, however, I have learnt how to combat that tendency and work more effectively and efficiently. One strategy that I have found useful is starting by drafting outlines with a structure and notes about key points. The main learning for me, however, has been the process of accepting that a first draft will not be my best writing, but is a good starting place to get constructive feedback and to subsequently iterate and improve. I think this learning also applies to research more broadly. One study is unlikely to definitively answer a particular research question; there are so many contextual factors (eg. participant and intervention characteristics) and logistical constraints (e.g. time, budget, resource) that can influence a study that it will always have limitations. This does not mean that these studies are not valuable - they each contribute to the overall body of knowledge, which shapes our overall understanding of particular research areas and questions. Essentially, through the research I have conducted, I have learnt that research is a gradual process that depends on collaboration, cooperation, and synthesis of knowledge, and that it is okay to work within limitations to take a step towards the next stage, rather than achieving everything all at once.

This collaboration should not be limited to within academia. A key learning from conducting evaluations of engagement with DHIs is the importance of including the end users in both the process of intervention design and evaluation. Patient and public involvement and co-creation have become widely recognised as increasingly important, but the process of conducting this research has impressed on me why this is so important, particularly in digital health. DHIs are

often delivered without significant clinical support - many behaviour change interventions, for example, might be found by users on their own. My thesis has highlighted how affective and cognitive engagement are essential for supporting users' motivation to engage with the intervention and target behaviour and how various barriers can interfere with engagement. It is crucial that intervention design meets actual, rather than perceived, needs and achieving this requires collaboration with the people who will be using the intervention. While my thesis did not involve much PPI, as it was focused on the evaluation of three previously-developed interventions, this is a learning that I will take forward in my career.

Although my thesis did not involve co-creation, the emphasis on qualitative research methods enabled me to capture participants' experiences and perceptions of the interventions in their own voices. My initial thesis plan was more quantitatively focused and would have primarily been a statistical comparison of the impact of particular BCTs on engagement, with a qualitative aspect to support it. Through examining the concept of engagement, and particularly its affective and cognitive components, the value of qualitative research has become extremely clear. While it is not as clear-cut to analyse as quantitative research, it enables more in-depth explorations of 'how' and 'why' questions, which lead to better understanding of key concepts. Chapter 3 explored how the different components of engagement could be examined using quantitative self-report measures, but from conducting this research, I strongly believe that mixed-methods incorporating both qualitative and quantitative data provide the best opportunity for gaining a comprehensive understanding of a research question.

Conducting primarily qualitative analyses has impressed upon me the importance of justification and transparency in research. Qualitative analysis involves researcher interpretation that is impossible to avoid. While I have learnt various strategies for how to improve credibility and confidence in findings, the one that stands out most to me is transparency, particularly as many of the qualitative analyses in this thesis were conducted independently. In reviewing the initial analyses I conducted mapping and synthesising patterns of engagement and how I had written them up in Chapters 5 and 6, I realised that they were not sufficiently transparent or defensible. I had conducted the analysis and explained my findings but not demonstrated clearly how I had gotten from the analysis to the conclusions. In my re-analysis for each of these, I added additional thematic frameworks to the appendices, outlining how I had mapped factors to engagement components and generated patterns of engagement from those factors, complete with participant quotes. These analyses may still be biased by my own life experiences and perspectives, but including this evidence

of my analytic process means that while others might disagree with my interpretation, it will be clearer to them how I reached it. This provides an opportunity for further evaluation and the advancement of knowledge by incorporating new evidence or additional perspectives that build on my initial analysis.

I have found the process of conducting this thesis, though challenging at times, to have been very rewarding and to have helped me grow as a researcher. My perspectives on what research is and how it should be conducted have evolved over the course of my PhD to include a greater emphasis on the importance of flexibility, iteration, collaboration and the inclusion of different perspectives, and justification and transparency in knowledge sharing.

7.6 Conclusions

Engagement is an essential element for achieving positive behavioural and health impacts of DHIs. Previous research has established that engagement is a multifaceted construct including affective, cognitive, and behavioural elements, but there is a gap in our understanding of how these components interact with each other over the process of engagement with DHIs. The thesis aimed to address the research question “*How are the different components of engagement (affective, cognitive, behavioural, micro, and macro) related in digital behaviour change app interventions for weight management and mental health and in an autonomous telemedicine post-surgical follow-up assessment?*” by mapping how these components influence each other, comparing different types of DHI to assess generalisability, and developing recommendations for how these individual components could be supported. This thesis makes an original contribution to knowledge by providing a framework for understanding the relationships between the different components of engagement. This framework encompasses three key contributions: 1) definitions of affective, cognitive, and behavioural engagement, 2) a model mapping the associations between these components over the process of engagement with DHIs, and 3) recommendations for key factors to consider in the development and evaluation of DHIs and how intervention features can support the different components of engagement.

Digital technologies have the potential to make health interventions more accessible and easily available to the general population, but people often do not engage with digital interventions enough to achieve the intended behavioural or health outcomes. With the increasing demand on healthcare services and emphasis on digital delivery of care, it is essential that issues of low engagement and adherence are addressed. To address these issues, we need to understand how users engage with DHIs, which means addressing key gaps in our

knowledge. On the evaluation side, studies rarely examine engagement comprehensively; when it is investigated, it is primarily captured through limited usage or adherence measures [14,41,44,45]. These behavioural measures cannot capture affective and cognitive components of engagement, so there is still a limited understanding of how these different components of engagement are related to each other and how they can be supported to achieve intervention aims. On the theoretical side, research has moved beyond a focus on behavioural engagement with the technology to a multifaceted conceptualisation of engagement. What is still missing is a consistent conceptualisation of its components and an understanding of how they interact over the process of engagement [476]. This gap in knowledge is important to understand for three key reasons. Clarity on the conceptualisation, measurement, and relationships between different components of engagement could 1) improve the quality and comparability of evidence of digital health engagement and the effectiveness of various engagement strategies, 2) enable the identification of specific barriers to engagement, and 3) provide a guide for incorporating particular engagement strategies to address those barriers.

The first contribution of the framework is in providing definitions of the key components of engagement. This was an essential element of the framework because I found that these components were not well or consistently defined in the literature. The definitions were generated based on various previous conceptualisations and compared to a recent integrated review that also examined and synthesised definitions of engagement in research [476]. In contrast to the review, my definitions focused specifically on digital health engagement; however, the contribution to the literature is best viewed in synthesis with this review. The definitions generated in this thesis and in that review were largely similar, indicating that we are approaching greater consensus in the literature around how these components can be defined. This will be valuable for future research, as it will provide a structure for determining how these components can be assessed in individual studies while supporting their comparability with other evaluations.

Although the definitions were a necessary element, the main contribution of the framework is the model developed that maps the relationships between affective, cognitive, and behavioural components over the process of engagement with a DHI (and target behaviour, if applicable). This was a key gap in the literature [476] but has significant implications for DHI design. Engagement is a necessary prerequisite for a DHI to have a positive impact, so it is essential that interventions can support sufficient engagement to achieve their aims (“effective engagement”) [14]. To optimise the selection and impact of particular engagement strategies,

we need to know what the barrier to engagement is. By mapping the relationships between the different components, this model provides a framework for understanding how particular barriers might be interfering with engagement by identifying in which components engagement is poor. Once we have an understanding of where in the process the problem is and how it is influencing other components of engagement, the model can be used to help identify particular intervention strategies that could mitigate the barrier and support engagement.

Previous models of engagement have been generated in the literature - for example, by Perksi *et al.* [42] and more recently Bijkerk *et al.* [476] and Nahum-Shani and Yoon [281,476] - so what does my model contribute to this body of knowledge? Nahum-Shani and Yoon's model focuses on the process of engagement with the various digital and non-digital stimuli and tasks that comprise the overall engagement and behaviour change process; while they adopt a multifaceted conceptualisation of engagement, this is not explicit in the model [281]. While both Perksi *et al.* [42] and Bijkerk *et al.* [476] use multifaceted conceptualisations of engagement in their models, neither examine how these components are related to each other. The key contribution of my model is in the exploration of these associations. The model outlines the high degree of interaction between all three components across engagement with the intervention and target behaviour, with a particularly close relationship between affective and cognitive engagement. This relationship has previously been a point of debate in the field, with some conceptualisations of engagement combining these components into one overarching category (e.g. subjective experience [42]). While in practice there may be such a degree of overlap between the components that it is difficult to distinguish the influence of particular factors or engagement strategies on them, it is important that affective and cognitive aspects are considered when evaluating or designing for engagement with DHIs. Synthesising my model with other recent models of engagement such as these three [42,281,476] could further increase their value in providing a comprehensive overview of the process of digital health engagement. For example, while I have highlighted context as a key factor influencing engagement across DHIs, contextual factors are captured in greater detail in Perksi *et al.*'s conceptual framework [42]. Nahum-Shani and Yoon's model, meanwhile, provides a more detailed breakdown of the steps involved in the overall process of engagement [281]. It would be interesting for future research to explore whether the model of associations I have proposed equally applies in this finer-grained view of engagement with particular stimuli and tasks.

The final main contribution of the framework developed in this thesis relates to the implications of my model for intervention design and evaluation. The synthesis of case studies of three different DHIs helped to broaden the applicability of my framework. Examining the similarities and differences across these cases enabled me to generate a set of key factors that appear to be associated with engagement across different types of DHI. This provides a general guide for key factors to consider when designing or evaluating DHIs. More specifically, I also generated a list of intervention design and content features that were related to engagement in the literature and the three case studies examined in this thesis and posited, based on this evidence, how they could influence the relationships between the different components of engagement captured in my model.

In conclusion, this thesis addressed a key gap in knowledge around our understanding of the relationships between different components of engagement by conducting a theoretical and empirical investigation to generate a model mapping the relationships between these components. Overall, the thesis provides a framework for intervention developers and researchers looking to explore potential barriers to users' engagement with an intervention, understand why those barriers might be hindering engagement from a psychological perspective, and determine how they could be mitigated through intervention features and design. This thesis makes an original contribution to our understanding of how and why people engage with DHIs. It extends previous theoretical work that established engagement as a multi-faceted construct by delving into how those different components can be conceptualised and how they are influenced by each other and various factors. By applying psychological theory to digital health evaluations and mapping the relationships between the different components, the thesis provides an initial framework for further research and development to address barriers to digital health engagement and impact.

References

1. Milne-Ives M, Homer S, Andrade J, Meinert E. Associations Between Behavior Change Techniques and Engagement With Mobile Health Apps: Protocol for a Systematic Review. *JMIR Res Protoc* [Internet]. 2022 Mar [cited 2023 Mar 9];11(3). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9006128/>
2. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile Apps for Health Behavior Change in Physical Activity, Diet, Drug and Alcohol Use, and Mental Health: Systematic Review. *JMIR Mhealth Uhealth* [Internet]. 2020 Mar 18;8(3):e17046. Available from: <http://dx.doi.org/10.2196/17046>
3. Milne-Ives M, Homer S, Andrade J, Meinert E. The conceptualisation and measurement of engagement in digital health. *Internet Interventions* [Internet]. 2024 Jun [cited 2024 Apr 9];36. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10979253/>
4. Milne-Ives M, Ananthakrishnan A, Homer S, Andrade J, Meinert E. Engagement with and impact of a mobile health app for childhood obesity prevention and management: a mixed methods study protocol [Internet]. *Research Square* (Preprint). 2024. Available from: <https://www.researchsquare.com/article/rs-4510115/v1>
5. Huckvale K, Wang CJ, Majeed A, Car J. Digital health at fifteen: more human (more needed). *BMC Med* [Internet]. 2019 Mar 18 [cited 2023 Jul 24];17(1):1–4. Available from: <https://link.springer.com/articles/10.1186/s12916-019-1302-0>
6. Kostkova P. Grand Challenges in Digital Health. *Front Public Health* [Internet]. 2015 May 5 [cited 2023 Jul 24];3:147199. Available from: <http://dx.doi.org/10.3389/fpubh.2015.00134>
7. Senbekov M, Saliev T, Bukeyeva Z, Almabayeva A, Zhanaliyeva M, Aitenova N, et al. The Recent Progress and Applications of Digital Technologies in Healthcare: A Review. *Int J Telemed Appl* [Internet]. 2020 [cited 2023 Jun 27];2020. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7732404/>
8. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of Attrition and Dropout in App-Based Interventions for Chronic Disease: Systematic Review and Meta-Analysis. *J Med Internet Res* [Internet]. 2020 Sep 29 [cited 2021 Nov 5];22(9):e20283. Available from: <https://www.jmir.org/2020/9/e20283>
9. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis. *J Affect Disord* [Internet]. 2020 Feb 15;263:413–9. Available from: <http://dx.doi.org/10.1016/j.jad.2019.11.167>

10. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *npj Digital Medicine* [Internet]. 2020 Feb 17 [cited 2021 Nov 8];3(1):1–10. Available from: <https://www.nature.com/articles/s41746-020-0224-8>
11. Birnbaum F, Lewis DM, Rosen R, Ranney ML. Patient engagement and the design of digital health. *Acad Emerg Med* [Internet]. 2015 Jun [cited 2022 Aug 31];22(6):754. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4674428/>
12. Yeager CM, Benight CC. If we build it, will they come? Issues of engagement with digital health interventions for trauma recovery. *Mhealth* [Internet]. 2018 Sep 11;4:37. Available from: <http://dx.doi.org/10.21037/mhealth.2018.08.04>
13. Baumel A, Muench F, Edan S, Kane JM. Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis. *J Med Internet Res* [Internet]. 2019 Sep 25 [cited 2021 Oct 11];21(9):e14567. Available from: <https://www.jmir.org/2019/9/e14567>
14. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. *Am J Prev Med* [Internet]. 2016 Nov;51(5):833–42. Available from: <http://dx.doi.org/10.1016/j.amepre.2016.06.015>
15. Dawkins B, Renwick C, Ensor T, Shinkins B, Jayne D, Meads D. What factors affect patients' ability to access healthcare? An overview of systematic reviews. *Trop Med Int Health* [Internet]. 2021 Oct;26(10):1177–88. Available from: <http://dx.doi.org/10.1111/tmi.13651>
16. De Hert S. Burnout in Healthcare Workers: Prevalence, Impact and Preventative Strategies. *Local Reg Anesth* [Internet]. 2020 Oct 28;13:171–83. Available from: <http://dx.doi.org/10.2147/LRA.S240564>
17. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in Health Care Services Caused by the Adoption of Digital Health Technologies: Scoping Review. *J Med Internet Res* [Internet]. 2022 Mar 21 [cited 2023 Jun 27];24(3):e34144. Available from: <https://www.jmir.org/2022/3/e34144>
18. The NHS Long Term Plan [Internet]. NHS; 2019 Jan [cited 2021 Mar 2]. Available from: <https://www.longtermplan.nhs.uk>
19. CDC's Global Digital Health Strategy [Internet]. Centers for Disease Control and Prevention. 2023 [cited 2023 Jul 24]. Available from: <https://www.cdc.gov/globalhealth/topics/gdhs/index.html#:~:text=The%20Centers%20for%20Di,sease%20Control,%2C%20regional%2C%20and%20global%20levels.>

20. Fatehi F, Samadbeik M, Kazemi A. What is Digital Health? Review of Definitions. *Stud Health Technol Inform* [Internet]. 2020 Nov 23;275:67–71. Available from: <http://dx.doi.org/10.3233/SHTI200696>
21. Sharma A, Harrington RA, McClellan MB, Turakhia MP, Eapen ZJ, Steinhubl M Jr, et al. Using Digital Health Technology to Better Generate Evidence and Deliver Evidence-Based Care. *J Am Coll Cardiol* [Internet]. 2018 Jun 12 [cited 2023 Jun 27];71(23):2680–90. Available from: <http://dx.doi.org/10.1016/j.jacc.2018.03.523>
22. Meskó B, Drobni Z, Bényei É, Gergely B, Györfy Z. Digital health is a cultural transformation of traditional healthcare. *mHealth* [Internet]. 2017 [cited 2023 Jun 27];3. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5682364/>
23. Clancy CM. Patient Engagement in Health Care [Internet]. Vol. 46, *Health Services Research*. 2011. p. 389–93. Available from: <http://dx.doi.org/10.1111/j.1475-6773.2011.01254.x>
24. Sawesi S, Rashrash M, Phalakornkule K, Carpenter JS, Jones JF. The Impact of Information Technology on Patient Engagement and Health Behavior Change: A Systematic Review of the Literature [Internet]. Vol. 4, *JMIR Medical Informatics*. 2016. p. e1. Available from: <http://dx.doi.org/10.2196/medinform.4514>
25. Ferrara G, Kim J, Lin S, Hua J, Seto E. A Focused Review of Smartphone Diet-Tracking Apps: Usability, Functionality, Coherence With Behavior Change Theory, and Comparative Validity of Nutrient Intake and Energy Estimates. *JMIR Mhealth Uhealth* [Internet]. 2019 May 17;7(5):e9232. Available from: <http://dx.doi.org/10.2196/mhealth.9232>
26. Digital Health Trends 2021: Innovation, Evidence, Regulation, and Adoption [Internet]. IQVIA Institute for Human Data Science; 2021 Jul [cited 2023 Jul 24]. Available from: <https://www.iqvia.com/form-pages/institute-gated>
27. Pandit JA, Radin JM, Quer G, Topol EJ. Smartphone apps in the COVID-19 pandemic. *Nat Biotechnol* [Internet]. 2022 Jun 20 [cited 2023 Jul 24];40(7):1013–22. Available from: <https://www.nature.com/articles/s41587-022-01350-x>
28. Trifan A, Oliveira M, Oliveira JL. Passive Sensing of Health Outcomes Through Smartphones: Systematic Review of Current Solutions and Possible Limitations. *JMIR Mhealth Uhealth* [Internet]. 2019 Aug 23;7(8):e12649. Available from: <http://dx.doi.org/10.2196/12649>
29. Harari GM, Lane ND, Wang R, Crosier BS, Campbell AT, Gosling SD. Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges. *Perspect Psychol Sci* [Internet]. 2016 Nov;11(6):838–54. Available from: <http://dx.doi.org/10.1177/1745691616650285>

30. Torous J, Michalak EE, O'Brien HL. Digital Health and Engagement—Looking Behind the Measures and Methods. *JAMA Network Open* [Internet]. 2020;3(7):e2010918. Available from: <http://dx.doi.org/10.1001/jamanetworkopen.2020.10918>
31. Grady A, Yoong S, Sutherland R, Lee H, Nathan N, Wolfenden L. Improving the public health impact of eHealth and mHealth interventions. *Australian and New Zealand Journal of Public Health* [Internet]. 2018;42(2):118–9. Available from: <http://dx.doi.org/10.1111/1753-6405.12771>
32. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HWW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. *J Cancer Surviv* [Internet]. 2017 Dec;11(6):704–19. Available from: <http://dx.doi.org/10.1007/s11764-017-0632-1>
33. Etminani K, Tao Engström A, Göransson C, Sant'Anna A, Nowaczyk S. How Behavior Change Strategies are Used to Design Digital Interventions to Improve Medication Adherence and Blood Pressure Among Patients With Hypertension: Systematic Review. *J Med Internet Res* [Internet]. 2020 Apr 9;22(4):e17201. Available from: <http://dx.doi.org/10.2196/17201>
34. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the Trial: Systematic Review of Real-World Uptake and Engagement With Digital Self-Help Interventions for Depression, Low Mood, or Anxiety. *J Med Internet Res* [Internet]. 2018 Jun 6;20(6):e199. Available from: <http://dx.doi.org/10.2196/jmir.9275>
35. Graham AK, Kwasny MJ, Lattie EG, Greene CJ, Gupta NV, Reddy M, et al. Targeting subjective engagement in experimental therapeutics for digital mental health interventions. *Internet Interv* [Internet]. 2021 Sep;25:100403. Available from: <http://dx.doi.org/10.1016/j.invent.2021.100403>
36. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A Systematic Review of the Impact of Adherence on the Effectiveness of e-Therapies. *J Med Internet Res* [Internet]. 2011 [cited 2023 Jul 25];13(3). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3222162/>
37. McLaughlin M, Delaney T, Hall A, Byaruhanga J, Mackie P, Grady A, et al. Associations Between Digital Health Intervention Engagement, Physical Activity, and Sedentary Behavior: Systematic Review and Meta-analysis. *J Med Internet Res* [Internet]. 2021 Feb [cited 2022 Nov 19];23(2). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8011420/>
38. Delaney T, McLaughlin M, Hall A, Yoong SL, Brown A, O'Brien K, et al. Associations between Digital Health Intervention Engagement and Dietary Intake: A Systematic Review. *Nutrients* [Internet]. 2021 Sep 20 [cited 2023 Jul 25];13(9):3281. Available from: <https://www.mdpi.com/2072-6643/13/9/3281>

39. Auster-Gussman LA, Lockwood KG, Graham SA, Pitter V, Branch OH. Engagement in Digital Health App-Based Prevention Programs Is Associated With Weight Loss Among Adults Age 65. *Front Digit Health* [Internet]. 2022 May 19;4:886783. Available from: <http://dx.doi.org/10.3389/fdgth.2022.886783>
40. Gan DZQ, McGillivray L, Han J, Christensen H, Torok M. Effect of Engagement With Digital Interventions on Mental Health Outcomes: A Systematic Review and Meta-Analysis. *Front Digit Health* [Internet]. 2021 Nov 4;3:764079. Available from: <http://dx.doi.org/10.3389/fdgth.2021.764079>
41. Oakley-Girvan I, Yunis R, Longmire M, Ouillon JS. What Works Best to Engage Participants in Mobile App Interventions and e-Health: A Scoping Review. *Telemed J E Health* [Internet]. 2022 Jun [cited 2024 Feb 19];28(6):768. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9231655/>
42. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* [Internet]. 2017 Jun;7(2):254–67. Available from: <http://dx.doi.org/10.1007/s13142-016-0453-1>
43. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the Concept of Adherence to eHealth Technology: Systematic Review on When Usage Becomes Adherence. *J Med Internet Res* [Internet]. 2017 Dec 6;19(12):e402. Available from: <http://dx.doi.org/10.2196/jmir.8578>
44. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring Engagement in eHealth and mHealth Behavior Change Interventions: Viewpoint of Methodologies. *J Med Internet Res* [Internet]. 2018 Nov [cited 2022 Nov 14];20(11). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6269627/>
45. Kelders SM, van Zyl LE, Ludden GDS. The Concept and Components of Engagement in Different Domains Applied to eHealth: A Systematic Scoping Review. *Front Psychol* [Internet]. 2020 [cited 2021 Oct 11];0. Available from: <http://dx.doi.org/10.3389/fpsyg.2020.00926>
46. Bijkerk LE, Oenema A, Geschwind N, Spigt M. Measuring Engagement with Mental Health and Behavior Change Interventions: an Integrative Review of Methods and Instruments. *Int J Behav Med* [Internet]. 2023 Apr [cited 2024 Apr 9];30(2). Available from: <https://pubmed.ncbi.nlm.nih.gov/35578099/>
47. Milne-Ives M, Homer SR, Andrade J, Meinert E. Potential associations between behavior change techniques and engagement with mobile health apps: a systematic review. *Front Psychol* [Internet]. 2023 Sep 18 [cited 2023 Dec 15];14:1227443. Available from: <http://dx.doi.org/10.3389/fpsyg.2023.1227443>

48. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. *Transl Behav Med* [Internet]. 2021 May 8 [cited 2023 Jun 13];11(7):1473–80. Available from: <https://academic.oup.com/tbm/article-pdf/11/7/1473/39456753/ibab048.pdf>
49. Nahum-Shani I, Shaw SD, Carpenter SM, Murphy SA, Yoon C. Engagement in digital interventions. *Am Psychol* [Internet]. 2022 Oct;77(7):836–52. Available from: <http://dx.doi.org/10.1037/amp0000983>
50. Cole-Lewis H, Ezeanochie N, Turgiss J. Understanding Health Behavior Technology Engagement: Pathway to Measuring Digital Behavior Change Interventions. *JMIR Form Res* [Internet]. 2019 Oct 10;3(4):e14052. Available from: <http://dx.doi.org/10.2196/14052>
51. O'Brien H. Theoretical Perspectives on User Engagement. In: O'Brien H, Cairns P, editors. *Why Engagement Matters* [Internet]. Cham: Springer International Publishing; 2016. p. 1–26. Available from: http://link.springer.com/10.1007/978-3-319-27446-1_1
52. APA Dictionary of Psychology [Internet]. American Psychological Association. 2018 [cited 2024 Jan 29]. Available from: <https://dictionary.apa.org/behavior>
53. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implement Sci* [Internet]. 2011 [cited 2022 Nov 14];6:42. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3096582/>
54. Marques MM Dr, Wright AJ, Johnston M, West R PhD, Hastings J, Zhang L, et al. The Behaviour Change Technique Ontology: Transforming the Behaviour Change Technique Taxonomy v1 [Internet]. *PsyArXiv*. 2023. Available from: <https://osf.io/vxypn>
55. Schenk PM, Wright AJ, West R, Hastings J, Lorencatto F, Moore C, et al. An ontology of mechanisms of action in behaviour change interventions. *Wellcome Open Res* [Internet]. 2023 Aug 11;8:337. Available from: <https://wellcomeopenresearch.org/articles/8-337/v1>
56. Michie S, Johnston M. Theories and techniques of behaviour change: Developing a cumulative science of behaviour change. *Health Psychol Rev* [Internet]. 2012 Mar;6(1):1–6. Available from: <http://www.tandfonline.com/doi/abs/10.1080/17437199.2012.654964>
57. Martinengo L, Lin X, Jabir AI, Kowatsch T, Atun R, Car J, et al. Conversational Agents in Health Care: Expert Interviews to Inform the Definition, Classification, and Conceptual Framework. *J Med Internet Res* [Internet]. 2023 Nov 1 [cited 2024 Jan 29];25(1):e50767. Available from: <https://www.jmir.org/2023/1/e50767>

58. Hekler EB, Michie S, Pavel M, Rivera DE, Collins LM, Jimison HB, et al. Advancing Models and Theories for Digital Behavior Change Interventions. *Am J Prev Med* [Internet]. 2016 Nov;51(5):825–32. Available from: <http://dx.doi.org/10.1016/j.amepre.2016.06.013>
59. Yardley L, Choudhury T, Patrick K, Michie S. Current Issues and Future Directions for Research Into Digital Behavior Change Interventions. *Am J Prev Med* [Internet]. 2016 Nov;51(5):814–5. Available from: <http://dx.doi.org/10.1016/j.amepre.2016.07.019>
60. Wienert J, Jahnelt T, Maaß L. What are Digital Public Health Interventions? First Steps Toward a Definition and an Intervention Classification Framework. *J Med Internet Res* [Internet]. 2022 Jun [cited 2024 Jan 29];24(6). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9277526/>
61. Smith PG, Morrow RH, Ross DA. Types of intervention and their development. In: Smith PG, Morrow RH, Ross DA, editors. *Field Trials of Health Interventions: A Toolbox* 3rd edition [Internet]. OUP Oxford; 2015 [cited 2024 Jan 29]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK305514/>
62. Liu P, Xu Z, Zhao X. Road tests of self-driving vehicles: Affective and cognitive pathways in acceptance formation. *Transp Res Part A: Policy Pract* [Internet]. 2019 Jun;124:354–69. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0965856418307183>
63. Ben-Eliyahu A, Moore D, Dorph R, Schunn CD. Investigating the multidimensionality of engagement: Affective, behavioral, and cognitive engagement across science activities and contexts. *Contemp Educ Psychol* [Internet]. 2018 Apr;53:87–105. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0361476X17300334>
64. Makransky G, Lilleholt L. A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educ Technol Res Dev* [Internet]. 2018 Oct;66(5):1141–64. Available from: <http://link.springer.com/10.1007/s11423-018-9581-2>
65. Friedrich T, Schlauderer S, Overhage S. The impact of social commerce feature richness on website stickiness through cognitive and affective factors: An experimental study. *Electron Commer Res Appl* [Internet]. 2019 Jul;36(100861):100861. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1567422319300389>
66. Denning J, Zacharia K, Ball K, George ES, Islam SMS. Exploring engagement with a web-based dietary intervention for adults with type 2 diabetes: A mixed methods evaluation of the T2Diet study. *PLoS One* [Internet]. 2022 Dec 30;17(12):e0279466. Available from: <http://dx.doi.org/10.1371/journal.pone.0279466>

67. Kelders SM, Sommers-Spijkerman M, Goldberg J. Investigating the Direct Impact of a Gamified Versus Nongamified Well-Being Intervention: An Exploratory Experiment. *J Med Internet Res* [Internet]. 2018 Jul 26 [cited 2023 Jun 15];20(7):e9923. Available from: <https://www.jmir.org/2018/7/e247>
68. Salas-Pilco SZ, Yang Y, Zhang Z. Student engagement in online learning in Latin American higher education during the COVID-19 pandemic: A systematic review. *Br J Educ Technol* [Internet]. 2022 May;53(3):593–619. Available from: <http://dx.doi.org/10.1111/bjet.13190>
69. Westgate EC, Wilson TD. Boring thoughts and bored minds: The MAC model of boredom and cognitive engagement. *Psychol Rev* [Internet]. 2018 Oct;125(5):689–713. Available from: <http://dx.doi.org/10.1037/rev0000097>
70. Cook DA, Artino AR Jr. Motivation to learn: an overview of contemporary theories. *Med Educ* [Internet]. 2016 Oct;50(10):997–1014. Available from: <http://dx.doi.org/10.1111/medu.13074>
71. Short SE, Mollborn S. Social Determinants and Health Behaviors: Conceptual Frames and Empirical Advances. *Current opinion in psychology* [Internet]. 2015 Oct [cited 2024 Jan 29];5:78. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4511598/>
72. Bhavnani SP, Narula J, Sengupta PP. Mobile technology and the digitization of healthcare. *Eur Heart J* [Internet]. 2016 Feb 11 [cited 2021 Oct 11];37(18):1428–38. Available from: <https://academic.oup.com/eurheartj/article-pdf/37/18/1428/17356722/ehv770.pdf>
73. Facing the future: opportunities and challenges for 21st-century public health in implementing the Sustainable Development Goals and the Health 2020 policy framework [Internet]. World Health Organization, Regional Office for Europe; 2018 Jun [cited 2021 Feb 7]. Available from: https://www.euro.who.int/__data/assets/pdf_file/0003/374052/180278-public-health-future-eng.pdf
74. World mental health report: Transforming mental health for all [Internet]. World Health Organization; 2022 [cited 2024 Jan 29]. Available from: <https://www.who.int/publications/i/item/9789240049338>
75. Verduci E, Di Profio E, Fiore G, Zuccotti G. Integrated Approaches to Combatting Childhood Obesity. *Ann Nutr Metab* [Internet]. 2022 Jun 9;78 Suppl 2:8–19. Available from: <http://dx.doi.org/10.1159/000524962>
76. van Agteren J, Iasiello M, Lo L, Bartholomaeus J, Kopsaftis Z, Carey M, et al. A systematic review and meta-analysis of psychological interventions to improve mental wellbeing. *Nature Human Behaviour* [Internet]. 2021 Apr 19 [cited 2024 Apr 9];5(5):631–52. Available from: <https://www.nature.com/articles/s41562-021-01093-w>

77. Milne-Ives M, Rahman E, Bradwell H, Baines R, Boey T, Potter A, et al. Barriers and facilitators to parents' engagement with and perceived impact of a childhood obesity app: A mixed-methods study. *PLOS Digital Health* [Internet]. 2024 Mar [cited 2024 Apr 9];3(3). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10971669/>
78. Obesity [Internet]. Nuffield Trust. 2019 [cited 2021 Oct 28]. Available from: <https://www.nuffieldtrust.org.uk/resource/obesity>
79. Cuschieri S, Grech S. COVID-19: a one-way ticket to a global childhood obesity crisis? *J Diabetes Metab Disord* [Internet]. 2020 Nov 6;1–4. Available from: <http://dx.doi.org/10.1007/s40200-020-00682-2>
80. Obesity and overweight [Internet]. World Health Organization. 2021 [cited 2022 Apr 22]. Available from: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>
81. Baker C. Obesity Statistics [Internet]. House of Commons Library; 2021 Jan [cited 2021 Oct 28]. Report No.: BRIEFING PAPER Number 3336. Available from: <https://researchbriefings.files.parliament.uk/documents/SN03336/SN03336.pdf>
82. FastStats [Internet]. CDC National Center for Health Statistics. 2021 [cited 2021 Nov 5]. Available from: <https://www.cdc.gov/nchs/fastats/obesity-overweight.htm>
83. The Health Effects of Overweight and Obesity [Internet]. Centers for Disease Control and Prevention. 2020 [cited 2021 Oct 28]. Available from: <https://www.cdc.gov/healthyweight/effects/index.html>
84. Djalalinia S, Qorbani M, Peykari N, Kelishadi R. Health impacts of Obesity. *Pak J Med Sci Q* [Internet]. 2015 Jan;31(1):239–42. Available from: <http://dx.doi.org/10.12669/pjms.311.7033>
85. Smith JD, Fu E, Kobayashi M. Prevention and Management of Childhood Obesity and its Psychological and Health Comorbidities. *Annu Rev Clin Psychol* [Internet]. 2020 May 7 [cited 2021 Oct 28];16:351. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7259820/>
86. Bass R, Eneli I. Severe childhood obesity: an under-recognised and growing health problem. *Postgrad Med J* [Internet]. 2015 Nov;91(1081):639–45. Available from: <http://dx.doi.org/10.1136/postgradmedj-2014-133033>
87. Morrison KM, Shin S, Tarnopolsky M, Taylor VH. Association of depression & health related quality of life with body composition in children and youth with obesity. *J Affect Disord* [Internet]. 2015 Feb 1;172:18–23. Available from: <http://dx.doi.org/10.1016/j.jad.2014.09.014>
88. Beck AR. Psychosocial Aspects of Obesity. *NASN Sch Nurse* [Internet]. 2016 Jan;31(1):23–7. Available from: <http://dx.doi.org/10.1177/1942602X15619756>

89. Solmi F, Sharpe PhD H, Gage SH, Maddock J, Lewis G, Patalay P. Changes in the Prevalence and Correlates of Weight-Control Behaviors and Weight Perception in Adolescents in the UK, 1986-2015. *JAMA Pediatr* [Internet]. 2021 Mar 1;175(3):267–75. Available from: <http://dx.doi.org/10.1001/jamapediatrics.2020.4746>
90. Pischon T, Boeing H, Hoffmann K, Bergmann M, Schulze MB, Overvad K, et al. General and abdominal adiposity and risk of death in Europe. *N Engl J Med* [Internet]. 2008 Nov 13;359(20):2105–20. Available from: <http://dx.doi.org/10.1056/NEJMoa0801891>
91. Calculating the costs of the consequences of obesity [Internet]. World Obesity Federation. 2017 [cited 2022 Apr 12]. Available from: <https://www.worldobesity.org/resources/resource-library/calculating-the-costs-of-the-consequences-of-obesity>
92. Department of Health and Social Care. Childhood obesity: a plan for action, chapter 2 [Internet]. GOV.UK; 2018 [cited 2021 Oct 28]. Available from: <https://www.gov.uk/government/publications/childhood-obesity-a-plan-for-action-chapter-2>
93. Health matters: obesity and the food environment [Internet]. Public Health England. 2017 [cited 2021 Oct 28]. Available from: <https://www.gov.uk/government/publications/health-matters-obesity-and-the-food-environment/health-matters-obesity-and-the-food-environment--2>
94. Why It Matters [Internet]. Centers for Disease Control and Prevention. 2022 [cited 2023 May 13]. Available from: <https://www.cdc.gov/obesity/about-obesity/why-it-matters.html>
95. Robinson N, McKay JA, Pearce MS, Albani V, Wright CM, Adamson AJ, et al. The Biological and Social Determinants of Childhood Obesity: Comparison of 2 Cohorts 50 Years Apart. *J Pediatr* [Internet]. 2021 Jan 1 [cited 2023 Jan 17];228:138–46.e5. Available from: <http://dx.doi.org/10.1016/j.jpeds.2020.09.031>
96. Wu S, Ding Y, Wu F, Li R, Hu Y, Hou J, et al. Socio-economic position as an intervention against overweight and obesity in children: a systematic review and meta-analysis. *Sci Rep* [Internet]. 2015 Jun 26 [cited 2023 Mar 9];5(1):1–11. Available from: <https://www.nature.com/articles/srep11354>
97. Wang L, Morelen D, Alamian A. A prospective cohort study of the association between key family and individual factors and obesity status among youth. *Sci Rep* [Internet]. 2022 Sep 19 [cited 2023 Mar 9];12(1):1–10. Available from: <https://www.nature.com/articles/s41598-022-19585-8>
98. Preventing excess weight gain [Internet]. National Institute for Health and Care Excellence. NICE; 2015 [cited 2021 Oct 28]. Available from: <https://www.nice.org.uk/guidance/ng7>

99. Strategies to Prevent & Manage Obesity [Internet]. Centers for Disease Control and Prevention. 2021 [cited 2021 Oct 28]. Available from: <https://www.cdc.gov/obesity/strategies/index.html>
100. Ruth S. M Chan JW. Prevention of Overweight and Obesity: How Effective is the Current Public Health Approach. *Int J Environ Res Public Health* [Internet]. 2010 Mar [cited 2021 Oct 28];7(3):765. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2872299/>
101. Geirhos A, Stephan M, Wehrle M, Mack C, Messner EM, Schmitt A, et al. Standardized evaluation of the quality and persuasiveness of mobile health applications for diabetes management. *Sci Rep* [Internet]. 2022 Mar 7 [cited 2023 Mar 10];12(1):1–10. Available from: <https://www.nature.com/articles/s41598-022-07544-2>
102. Wang Y, Xue H, Huang Y, Huang L, Zhang D. A Systematic Review of Application and Effectiveness of mHealth Interventions for Obesity and Diabetes Treatment and Self-Management. *Adv Nutr* [Internet]. 2017 May 5 [cited 2022 Aug 21];8(3):449–62. Available from: <https://academic.oup.com/advances/article-pdf/8/3/449/23755031/an014100.pdf>
103. Ghelani DP, Moran LJ, Johnson C, Mousa A, Naderpoor N. Mobile Apps for Weight Management: A Review of the Latest Evidence to Inform Practice. *Front Endocrinol* [Internet]. 2020 [cited 2022 Apr 13];11. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7326765/>
104. Rice L, Sara R. Updating the determinants of health model in the Information Age. *Health Promot Int* [Internet]. 2019 Dec 1;34(6):1241–9. Available from: <http://dx.doi.org/10.1093/heapro/day064>
105. Fitness App Market Size, Share & Growth Report, 2021 - 2028 [Internet]. Grand View Research. [cited 2021 Oct 28]. Available from: <https://www.grandviewresearch.com/industry-analysis/fitness-app-market>
106. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. *Int J Behav Nutr Phys Act* [Internet]. 2016 Mar 10 [cited 2023 Mar 10];13(1):1–9. Available from: <https://ijbnpa.biomedcentral.com/articles/10.1186/s12966-016-0359-9>
107. Tully L, Burls A, Sorensen J, El-Moslemany R, O'Malley G. Mobile Health for Pediatric Weight Management: Systematic Scoping Review. *JMIR mHealth and uHealth* [Internet]. 2020 Jun 3 [cited 2022 Feb 11];8(6):e16214. Available from: <https://mhealth.jmir.org/2020/6/e16214>
108. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? *npj Digital Medicine* [Internet]. 2020 Jan 13 [cited 2022 Apr 4];3(1):1–6.

Available from: <https://www.nature.com/articles/s41746-019-0206-x>

109. Singh K, Drouin K, Newmark LP, Filkins M, Silvers E, Bain PA, et al. Patient-Facing Mobile Apps to Treat High-Need, High-Cost Populations: A Scoping Review. *JMIR Mhealth Uhealth* [Internet]. 2016 Dec 19;4(4):e136. Available from: <http://dx.doi.org/10.2196/mhealth.6445>
110. Chen J, Cade JE, Allman-Farinelli M. The Most Popular Smartphone Apps for Weight Loss: A Quality Assessment. *JMIR Mhealth Uhealth* [Internet]. 2015 Dec 16;3(4):e104. Available from: <http://dx.doi.org/10.2196/mhealth.4334>
111. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* [Internet]. 2013 Apr 18;15(4):e86. Available from: <http://dx.doi.org/10.2196/jmir.2583>
112. Tackling obesity: empowering adults and children to live healthier lives [Internet]. Department of Health & Social Care; 2020 Jul [cited 2021 Oct 28]. Available from: <https://www.gov.uk/government/publications/tackling-obesity-government-strategy/tackling-obesity-empowering-adults-and-children-to-live-healthier-lives>
113. Forman EM, Evans BC, Flack D, Goldstein SP. Could technology help us tackle the obesity crisis? *Future Science OA* [Internet]. 2016 Dec [cited 2021 Nov 5];2(4). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5242208/>
114. Digital Implementation Investment Guide (DIIG): Integrating Digital Interventions into Health Programmes [Internet]. World Health Organization; 2020 [cited 2021 Feb 7]. Available from: <https://www.who.int/publications/i/item/9789240010567>
115. Taj F, Klein MCA, van Halteren A. Digital Health Behavior Change Technology: Bibliometric and Scoping Review of Two Decades of Research. *JMIR mHealth and uHealth* [Internet]. 2019 [cited 2021 Jan 11];7(12):e13311. Available from: <https://mhealth.jmir.org/2019/12/e13311/>
116. Dounavi K, Tsoumani O. Mobile Health Applications in Weight Management: A Systematic Literature Review. *Am J Prev Med* [Internet]. 2019 Jun [cited 2022 Aug 21];56(6). Available from: <https://pubmed.ncbi.nlm.nih.gov/31003801/>
117. Oh C, Carducci B, Vaivada T, Bhutta ZA. Digital Interventions for Universal Health Promotion in Children and Adolescents: A Systematic Review. *Pediatrics* [Internet]. 2022 May 1;149(Suppl 5). Available from: <http://dx.doi.org/10.1542/peds.2021-053852H>

118. Yau KW, Tang TS, G6rges M, Pinkney S, Kim AD, Kalia A, et al. Effectiveness of Mobile Apps in Promoting Healthy Behavior Changes and Preventing Obesity in Children: Systematic Review. *JMIR Pediatrics and Parenting* [Internet]. 2022 [cited 2022 Nov 11];5(1). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9002598/>
119. Bonvicini L, Pingani I, Venturelli F, Patrignani N, Bassi MC, Broccoli S, et al. Effectiveness of mobile health interventions targeting parents to prevent and treat childhood Obesity: Systematic review. *Prev Med Rep* [Internet]. 2022 Oct;29:101940. Available from: <http://dx.doi.org/10.1016/j.pmedr.2022.101940>
120. Chai LK, Farletti R, Fathi L, Littlewood R. A Rapid Review of the Impact of Family-Based Digital Interventions for Obesity Prevention and Treatment on Obesity-Related Outcomes in Primary School-Aged Children. *Nutrients* [Internet]. 2022 Nov 15 [cited 2022 Nov 30];14(22):4837. Available from: <https://www.mdpi.com/2072-6643/14/22/4837>
121. Morrison LG. Theory-based strategies for enhancing the impact and usage of digital health behaviour change interventions: A review. *Digit Health* [Internet]. 2015 Jul 17;1. Available from: <http://dx.doi.org/10.1177/2055207615595335>
122. Wei Y, Zheng P, Deng H, Wang X, Li X, Fu H. Design Features for Improving Mobile Health Intervention User Engagement: Systematic Review and Thematic Analysis. *J Med Internet Res* [Internet]. 2020 Dec 9 [cited 2022 Nov 21];22(12):e21687. Available from: <https://www.jmir.org/2020/12/e21687>
123. Smith R, Michalopoulou M, Reid H, Payne Riches S, Wango YN, Kenworthy Y, et al. Applying the behaviour change wheel to develop a smartphone application “stay-active” to increase physical activity in women with gestational diabetes. *BMC Pregnancy Childbirth* [Internet]. 2022 [cited 2022 Nov 30];22. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8962081/>
124. Shoneye CL, Mullan B, Begley A, Pollard CM, Jancey J, Kerr DA. Design and Development of a Digital Weight Management Intervention (ToDAy): Qualitative Study. *JMIR mHealth and uHealth* [Internet]. 2020 Sep [cited 2022 Nov 30];8(9). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7511863/>
125. Adu MD, Malabu UH, Malau-Aduli AEO, Malau-Aduli BS. The development of My Care Hub Mobile-Phone App to Support Self-Management in Australians with Type 1 or Type 2 Diabetes. *Sci Rep* [Internet]. 2020 Jan 8 [cited 2023 Mar 9];10(1):1–10. Available from: <https://www.nature.com/articles/s41598-019-56411-0>

126. Partridge SR, Redfern J. Strategies to Engage Adolescents in Digital Health Interventions for Obesity Prevention and Management. *Healthcare (Basel)* [Internet]. 2018 Jun 21;6(3). Available from: <http://dx.doi.org/10.3390/healthcare6030070>
127. Arthurs N, Tully L, O'Malley G, Browne S. Usability and Engagement Testing of mHealth Apps in Paediatric Obesity: A Narrative Review of Current Literature. *Int J Environ Res Public Health* [Internet]. 2022 Jan 27;19(3). Available from: <http://dx.doi.org/10.3390/ijerph19031453>
128. Henriksson P, Migueles JH, Söderström E, Sandborg J, Maddison R, Löf M. User engagement in relation to effectiveness of a digital lifestyle intervention (the HealthyMoms app) in pregnancy. *Sci Rep* [Internet]. 2022 Aug 13 [cited 2023 Mar 9];12(1):1–9. Available from: <https://www.nature.com/articles/s41598-022-17554-9>
129. Meinert E, Milne-Ives M, Taylor E, Inkster B, Paik A, Ananthakrishnan A, et al. Real-world testing of an artificial intelligence conversational agent as an early intervention and support tool in the mental health referral care pathway: A randomised controlled trial. *International Journal of Social Psychiatry*. (*In Review*).
130. Mental health [Internet]. World Health Organization. [cited 2021 Mar 2]. Available from: https://www.who.int/health-topics/mental-health#tab=tab_1
131. The Improving Access to Psychological Therapies (IAPT) Pathway for People with Long-Term Conditions and Medically Unexplained Symptoms: Full implementation guidance [Internet]. London: National Collaborating Centre for Mental Health; 2018 [cited 2021 Mar 2]. Available from: https://www.rcpsych.ac.uk/docs/default-source/improving-care/nccmh/iapt/nccmh-iapt-ltc-full-implementation-guidance.pdf?sfvrsn=de824ea4_4
132. Cohen A. Addressing comorbidity between mental disorders and major noncommunicable diseases: Background technical report to support implementation of the WHO European Mental Health Action Plan 2013–2020 and the WHO European Action Plan for the Prevention and Control of Noncommunicable Diseases 2016–2025 [Internet]. World Health Organization; 2017 [cited 2021 Mar 2]. Available from: https://www.euro.who.int/__data/assets/pdf_file/0009/342297/Comorbidity-report_E-web.pdf?ua=1
133. Payne RA, Abel GA, Guthrie B, Mercer SW. The effect of physical multimorbidity, mental health conditions and socioeconomic deprivation on unplanned admissions to hospital: a retrospective cohort study. *Canadian Medical Association Journal* [Internet]. 2013;185(5):E221–8. Available from: <http://dx.doi.org/10.1503/cmaj.121349>

134. Naylor C, Parsonage M, McDaid D, Knapp M, Fossey M, Galea A. Long-term conditions and mental health: The cost of co-morbidities [Internet]. The King's Fund; 2012 Feb [cited 2021 Mar 2]. Available from: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/long-term-conditions-mental-health-cost-comorbidities-naylor-feb12.pdf
135. World Health Assembly. Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level: report by the Secretariat [Internet]. World Health Organization; 2012 [cited 2021 Mar 2]. Report No.: A65/10. Available from: <https://apps.who.int/iris/handle/10665/78898>
136. The Improving Access to Psychological Therapies Manual [Internet]. The National Collaborating Centre for Mental Health; 2019 Mar [cited 2021 Mar 2]. Available from: https://www.rcpsych.ac.uk/docs/default-source/improving-care/nccmh/iapt/the-iapt-manual--final--republished-7-3-18.pdf?sfvrsn=a192d1af_0
137. The Five Year Forward View for Mental Health [Internet]. The Mental Health Taskforce; 2016 Feb [cited 2021 Mar 2]. Available from: <https://www.england.nhs.uk/wp-content/uploads/2016/02/Mental-Health-Taskforce-FYFV-final.pdf>
138. McManus S, Bebbington P, Jenkins R, Brugha T, editors. Mental Health and Wellbeing in England: Adult Psychiatric Morbidity Survey 2014 [Internet]. Leeds: NHS Digital; 2016 [cited 2021 Mar 2]. Available from: https://files.digital.nhs.uk/pdf/q/3/mental_health_and_wellbeing_in_england_full_report.pdf
139. Alonso J, Liu Z, Evans-Lacko S, Sadikova E, Sampson N, Chatterji S, et al. Treatment gap for anxiety disorders is global: Results of the World Mental Health Surveys in 21 countries. *Depress Anxiety* [Internet]. 2018 Mar;35(3):195–208. Available from: <http://dx.doi.org/10.1002/da.22711>
140. Mental health in the workplace [Internet]. World Health Organization. [cited 2021 Mar 2]. Available from: <https://www.who.int/teams/mental-health-and-substance-use/mental-health-in-the-workplace>
141. Clark DM. A guide to Improving Access to Psychological Therapies services [Internet]. NHS England. 2018 [cited 2021 Mar 2]. Available from: <https://www.england.nhs.uk/blog/a-guide-to-improving-access-to-psychological-therapies-services/>
142. Psychological Therapies: Annual report on the use of IAPT services, England 2018-19 [Internet]. NHS Digital; 2019 Jul [cited 2021 Mar 2]. Available from: <https://files.digital.nhs.uk/1C/538E29/psych-ther-2018-19-ann-rep.pdf>

143. Community mental health survey 2020 [Internet]. Care Quality Commission. 2020 [cited 2021 Mar 2]. Available from: <https://www.cqc.org.uk/publications/surveys/community-mental-health-survey-2020>
144. Mental health and the NHS [Internet]. National Institute for Health and Care Excellence. [cited 2021 Mar 2]. Available from: <https://indepth.nice.org.uk/mental-health-and-the-nhs/index.html>
145. Two-fifths of patients waiting for mental health treatment forced to resort to emergency or crisis services [Internet]. Royal College of Psychiatrists. 2020 [cited 2021 Mar 2]. Available from: <https://www.rcpsych.ac.uk/news-and-features/latest-news/detail/2020/10/06/two-fifths-of-patients-waiting-for-mental-health-treatment-forced-to-resort-to-emergency-or-crisis-services>
146. Griffiths S, Steen S. Improving Access to Psychological Therapies (IAPT) Programme: Scrutinising IAPT Cost Estimates to Support Effective Commissioning. *The Journal of Psychological Therapies in Primary Care* [Internet]. 2013 Nov [cited 2021 Mar 2];2:142–56. Available from: <https://www.chester.ac.uk/sites/files/chester/IAPT%20-%20Scrutinising%20IAPT%20Cost%20Estimates.pdf>
147. Working days lost in Great Britain [Internet]. Health and Safety Executive. 2020 [cited 2021 Mar 2]. Available from: <https://www.hse.gov.uk/statistics/dayslost.htm>
148. Davies J. Mental health [Internet]. Nuffield Trust. 2020 [cited 2021 Mar 2]. Available from: <https://www.nuffieldtrust.org.uk/news-item/mental-health-2>
149. Parsonage M, Saini G. Mental health at work: The business costs ten years on [Internet]. Centre for Mental Health. 2017 [cited 2021 Mar 2]. Available from: <https://www.centreformentalhealth.org.uk/publications/mental-health-work-business-costs-ten-years>
150. Philippe TJ, Sikder N, Jackson A, Koblanski ME, Liow E, Pilarinos A, et al. Digital Health Interventions for Delivery of Mental Health Care: Systematic and Comprehensive Meta-Review. *JMIR Mental Health* [Internet]. 2022 May 12 [cited 2023 Jul 26];9(5):e35159. Available from: <https://mental.jmir.org/2022/5/e35159>
151. de Pennington N, Mole G, Lim E, Milne-Ives M, Normando E, Xue K, et al. Safety and Acceptability of a Natural Language Artificial Intelligence Assistant to Deliver Clinical Follow-up to Cataract Surgery Patients: Proposal. *JMIR Res Protoc* [Internet]. 2021 Jul [cited 2022 Aug 16];10(7). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8367096/>

152. Making our health and care systems fit for an ageing population [Internet]. The King's Fund. 2014 [cited 2023 Jun 27]. Available from: <https://www.kingsfund.org.uk/publications/making-our-health-and-care-systems-fit-ageing-population>
153. Illman J. Exclusive: NHS prepares to cancel elective ops in readiness for covid-19 surge [Internet]. Health Service Journal. 2020 [cited 2023 Jun 27]. Available from: <https://www.hsj.co.uk/free-for-non-subscribers/exclusive-nhs-prepares-to-cancel-elective-ops-in-readiness-for-covid-19-surge/7027110.article>
154. Griffin S. Covid-19: Waiting times in England reach record highs. BMJ [Internet]. 2020 Sep 11 [cited 2023 Jun 27];370. Available from: <https://www.bmj.com/content/370/bmj.m3557.abstract>
155. Department of Health and Social Care. Digital revolution to bust COVID backlogs and deliver more tailored care for patients [Internet]. GOV.UK. 2022 [cited 2022 Aug 16]. Available from: <https://www.gov.uk/government/news/digital-revolution-to-bust-covid-backlogs-and-deliver-more-tailored-care-for-patients>
156. 22/26 HSDR Evaluating the High Volume Low Complexity (HVLC) surgical hubs model Commissioning Brief [Internet]. National Institute for Health and Care Research. 2022 [cited 2022 Aug 16]. Available from: <https://www.nihr.ac.uk/documents/2226-hsdr-evaluating-the-high-volume-low-complexity-hvlc-surgical-hubs-model-commissioning-brief/29940>
157. Oliver D, Foot C, Humphries R. Making our health and care systems fit for an ageing population [Internet]. The King's Fund; 2014 [cited 2020 Nov 3]. Available from: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/making-health-care-systems-fit-ageing-population-oliver-foot-humphries-mar14.pdf
158. Myers LC, Liu VX. The COVID-19 Pandemic Strikes Again and Again and Again. JAMA Netw Open [Internet]. 2022 Mar 1 [cited 2022 Aug 16];5(3):e221760–e221760. Available from: https://jamanetwork.com/journals/jamanetworkopen/articlepdf/2789848/myers_2022_ic_220022_1646251687.29375.pdf
159. Alderwick H. Is the NHS overwhelmed? BMJ [Internet]. 2022 Jan 11 [cited 2022 Aug 16];376. Available from: <https://www.bmj.com/content/376/bmj.o51.abstract>
160. Dall TM, Gallo PD, Chakrabarti R, West T, Semilla AP, Storm MV. An Aging Population And Growing Disease Burden Will Require A Large And Specialized Health Care Workforce By 2025. Health Aff [Internet]. 2017 Aug 2 [cited 2022 Aug 16]; Available from: <http://dx.doi.org/10.1377/hlthaff.2013.0714>

161. Donachie PHJ, Sparrow JM. National Ophthalmology Database Audit: Year 5 Annual Report [Internet]. The Royal College of Ophthalmologists; 2020 Aug [cited 2020 May 20]. Available from:
<https://www.nodaudit.org.uk/u/docs/20/hqsrgrmurnv/NOD%20Audit%20Full%20Annual%20Report%202020.pdf>
162. Aaronson A, Viljanen A, Kanclerz P, Grzybowski A, Tuuminen R. Cataract complications study: an analysis of adverse effects among 14,520 eyes in relation to surgical experience. *Annals of Translational Medicine* [Internet]. 2020 Nov [cited 2022 Aug 26];8(22). Available from:
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7729371/>
163. Haripriya A, Chang DF, Reena M, Shekhar M. Complication rates of phacoemulsification and manual small-incision cataract surgery at Aravind Eye Hospital. *J Cataract Refract Surg* [Internet]. 2012 Aug [cited 2022 Aug 26];38(8). Available from:
<https://pubmed.ncbi.nlm.nih.gov/22814042/>
164. Hingorani M, El-Sherbiny S, Pradeep A, Stead R, Spiess K, Elgohary M. Managing an outbreak of postoperative endophthalmitis [Internet]. The Royal College of Ophthalmologists; 2022 Jan [cited 2024 May 16]. Available from: <https://www.rcophth.ac.uk/wp-content/uploads/2016/07/Managing-an-outbreak-of-postoperative-endophthalmitis-Final-2022.pdf>
165. MacEwen C, Davis A, Chang L. Ophthalmology: GIRFT Programme National Specialty Report [Internet]. Getting It Right First Time (GIRFT); 2019 Dec [cited 2020 May 20]. Available from: <https://gettingitrightfirsttime.co.uk/wp-content/uploads/2019/12/OphthalmologyReportGIRFT19P-FINAL.pdf>
166. Gan DZQ, McGillivray L, Larsen ME, Christensen H, Torok M. Technology-supported strategies for promoting user engagement with digital mental health interventions: A systematic review. *Digit Health* [Internet]. 2022 Jun 1;8:20552076221098268. Available from:
<http://dx.doi.org/10.1177/20552076221098268>
167. Jaworski BK, Webb Hooper M, Aklin WM, Jean-Francois B, Elwood WN, Belis D, et al. Advancing digital health equity: Directions for behavioral and social science research. *Transl Behav Med* [Internet]. 2023 Apr 3;13(3):132–9. Available from:
<http://dx.doi.org/10.1093/tbm/ibac088>
168. Edwards MK, Loprinzi PD. Comparative effects of meditation and exercise on physical and psychosocial health outcomes: a review of randomized controlled trials. *Postgrad Med* [Internet]. 2017 Nov 27 [cited 2023 Jun 19]; Available from:
<https://www.tandfonline.com/doi/abs/10.1080/00325481.2018.1409049>

169. Conner M, Norman P. Health behaviour: Current issues and challenges. *Psychol Health* [Internet]. 2017 Jun 14 [cited 2023 Jun 19]; Available from: <https://www.tandfonline.com/doi/abs/10.1080/08870446.2017.1336240>
170. Michie S, Atkins L, West R. *The Behaviour Change Wheel: A Guide to Designing Interventions* [Internet]. Silverback Publishing; 2014. 329 p. Available from: https://books.google.com/books/about/The_Behaviour_Change_Wheel.html?hl=&id=am6ytAEA CAAJ
171. Ng MM, Firth J, Minen M, Torous J. User Engagement in Mental Health Apps: A Review of Measurement, Reporting, and Validity. *Psychiatr Serv* [Internet]. 2019 Jul 1;70(7):538–44. Available from: <http://dx.doi.org/10.1176/appi.ps.201800519>
172. Moshe I, Terhorst Y, Philippi P, Domhardt M, Cuijpers P, Cristea I, et al. Digital interventions for the treatment of depression: A meta-analytic review. *Psychol Bull* [Internet]. 2021 Aug;147(8):749–86. Available from: <http://dx.doi.org/10.1037/bul0000334>
173. Samanez-Larkin GR, Knutson B. Decision making in the ageing brain: Changes in affective and motivational circuits. *Nat Rev Neurosci* [Internet]. 2015 May [cited 2023 Jun 17];16(5):278. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5645075/>
174. Uman LS. Systematic Reviews and Meta-Analyses. *J Can Acad Child Adolesc Psychiatry* [Internet]. 2011 Feb [cited 2023 May 13];20(1):57. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3024725/>
175. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* [Internet]. 2021 Mar 29;372:n160. Available from: <http://dx.doi.org/10.1136/bmj.n160>
176. Moller AC, Merchant G, Conroy DE, West R, Hekler E, Kugler KC, et al. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. *J Behav Med* [Internet]. 2017 Feb;40(1):85–98. Available from: <http://dx.doi.org/10.1007/s10865-016-9818-7>
177. Gourlan M, Bernard P, Bortolon C, Romain AJ, Lareyre O, Carayol M, et al. Efficacy of theory-based interventions to promote physical activity. A meta-analysis of randomised controlled trials. *Health Psychol Rev* [Internet]. 2016 [cited 2023 Jan 27];10(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/25402606/>
178. Klonoff DC. Behavioral Theory: The Missing Ingredient for Digital Health Tools to Change Behavior and Increase Adherence. *J Diabetes Sci Technol* [Internet]. 2019 Mar;13(2):276–81. Available from: <http://dx.doi.org/10.1177/1932296818820303>

179. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* [Internet]. 2013 Aug [cited 2023 Mar 10];46(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/23512568/>
180. Iribarren SJ, Akande TO, Kamp KJ, Barry D, Kader YG, Suelzer E. Effectiveness of Mobile Apps to Promote Health and Manage Disease: Systematic Review and Meta-analysis of Randomized Controlled Trials. *JMIR mHealth and uHealth* [Internet]. 2021 Jan 11 [cited 2022 Apr 13];9(1):e21563. Available from: <https://mhealth.jmir.org/2021/1/e21563>
181. Floryan M, Chow PI, Schueller SM, Ritterband LM. The Model of Gamification Principles for Digital Health Interventions: Evaluation of Validity and Potential Utility. *J Med Internet Res* [Internet]. 2020 Jun 10 [cited 2022 Aug 31];22(6):e16506. Available from: <https://www.jmir.org/2020/6/e16506>
182. Garnett C, Crane D, West R, Brown J, Michie S. Identification of Behavior Change Techniques and Engagement Strategies to Design a Smartphone App to Reduce Alcohol Consumption Using a Formal Consensus Method. *JMIR Mhealth Uhealth* [Internet]. 2015 Jun 29;3(2):e73. Available from: <http://dx.doi.org/10.2196/mhealth.3895>
183. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. *BMC Medical Informatics and Decision Making* [Internet]. 2016;16(1). Available from: <http://dx.doi.org/10.1186/s12911-016-0359-3>
184. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the Uptake of and Engagement With Health and Well-Being Smartphone Apps: Systematic Review. *J Med Internet Res* [Internet]. 2020 May 29;22(5):e17572. Available from: <http://dx.doi.org/10.2196/17572>
185. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and Facilitators of User Engagement With Digital Mental Health Interventions: Systematic Review. *J Med Internet Res* [Internet]. 2021 Mar 24;23(3):e24387. Available from: <http://dx.doi.org/10.2196/24387>
186. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* [Internet]. 2015 Jan 2;350:g7647. Available from: <http://dx.doi.org/10.1136/bmj.g7647>

187. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club* [Internet]. 1995 Nov;123(3):A12–3. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/7582737>
188. Counsell C. Formulating questions and locating primary studies for inclusion in systematic reviews. *Ann Intern Med* [Internet]. 1997 Sep 1;127(5):380–7. Available from: <http://dx.doi.org/10.7326/0003-4819-127-5-199709010-00008>
189. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. *Sci Transl Med* [Internet]. 2015 Apr 15;7(283):283rv3. Available from: <http://dx.doi.org/10.1126/scitranslmed.aaa3487>
190. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. *BMJ* [Internet]. 2011 Oct 18;343:d5928. Available from: <http://dx.doi.org/10.1136/bmj.d5928>
191. Risk of Bias 2 (RoB 2) tool [Internet]. [cited 2021 Nov 24]. Available from: <https://methods.cochrane.org/risk-bias-2>
192. Hong QN, Pluye P, Fàbregues S, Bartlett G, Boardman F, Cargo M, et al. Mixed Methods Appraisal Tool (MMAT) [Internet]. Registration of Copyright (#1148552), Canadian Intellectual Property Office, Industry Canada. ; 2018 [cited 2022 Jul 19]. Available from: http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/attach/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf
193. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* [Internet]. 2016 Oct 12;355:i4919. Available from: <http://dx.doi.org/10.1136/bmj.i4919>
194. Hendrie GA, Sazzad Hussain M, Brindal E, James-Martin G, Williams G, Crook A. Impact of a Mobile Phone App to Increase Vegetable Consumption and Variety in Adults: Large-Scale Community Cohort Study. *JMIR mHealth and uHealth* [Internet]. 2020 Apr 17 [cited 2024 Feb 22];8(4):e14726. Available from: <https://mhealth.jmir.org/2020/4/e14726>
195. Duff O, Walsh D, Malone S, McDermott L, Furlong B, O’Connor N, et al. MedFit App, a Behavior-Changing, Theoretically Informed Mobile App for Patient Self-Management of Cardiovascular Disease: User-Centered Development. *JMIR Formative Research* [Internet]. 2018;2(1):e8. Available from: <http://dx.doi.org/10.2196/formative.9550>
196. Delmas MA, Kohli A. Can Apps Make Air Pollution Visible? Learning About Health Impacts Through Engagement with Air Quality Information. *Journal of Business Ethics* [Internet]. 2020;161(2):279–302. Available from: <http://dx.doi.org/10.1007/s10551-019-04215-7>

197. Baumel A, Kane JM. Examining Predictors of Real-World User Engagement with Self-Guided eHealth Interventions: Analysis of Mobile Apps and Websites Using a Novel Dataset. *J Med Internet Res* [Internet]. 2018 Dec 14;20(12):e11491. Available from: <http://dx.doi.org/10.2196/11491>
198. Bell L, Garnett C, Qian T, Perski O, Williamson E, Potts HWW. Engagement With a Behavior Change App for Alcohol Reduction: Data Visualization for Longitudinal Observational Study. *J Med Internet Res* [Internet]. 2020 Dec [cited 2024 Feb 22];22(12). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7762688/>
199. Willms A, Rhodes RE, Liu S. The Development of a Hypertension Prevention and Financial-Incentive mHealth Program Using a “No-Code” Mobile App Builder: Development and Usability Study. *JMIR Formative Research* [Internet]. 2023 [cited 2024 Feb 22];7. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10131686/>
200. Attwood S, Parke H, Larsen J, Morton KL. Using a mobile health application to reduce alcohol consumption: a mixed-methods evaluation of the drinkaware track & calculate units application. *BMC Public Health* [Internet]. 2017 May 17 [cited 2024 Feb 22];17(1):1–21. Available from: <https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-017-4358-9>
201. Szinay D, Perski O, Jones A, Chadborn T, Brown J, Naughton F. Perceptions of Factors Influencing Engagement With Health and Well-being Apps in the United Kingdom: Qualitative Interview Study. *JMIR Mhealth Uhealth* [Internet]. 2021 Dec 16;9(12):e29098. Available from: <http://dx.doi.org/10.2196/29098>
202. Tang J, Abraham C, Stamp E, Greaves C. How can weight-loss app designers’ best engage and support users? A qualitative investigation. *Br J Health Psychol* [Internet]. 2015 Feb;20(1):151–71. Available from: <http://dx.doi.org/10.1111/bjhp.12114>
203. Direito A, Jiang Y, Whittaker R, Maddison R. Apps for IMproving FITness and Increasing Physical Activity Among Young People: The AIMFIT Pragmatic Randomized Controlled Trial. *J Med Internet Res* [Internet]. 2015 Aug [cited 2024 Feb 22];17(8). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4642788/>
204. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers’ and drinkers’ choice of smartphone applications and expectations of engagement: a think aloud and interview study. *BMC Med Inform Decis Mak* [Internet]. 2017 Feb 28 [cited 2023 Jun 30];17(1):1–14. Available from: <https://bmcmmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-017-0422-8>

205. Flaherty SJ, McCarthy M, Collins AM, McCafferty C, McAuliffe FM. Exploring engagement with health apps: the emerging importance of situational involvement and individual characteristics. *Eur J Mark* [Internet]. 2021 May 6 [cited 2024 Feb 22];55(13):122–47. Available from: <https://www.emerald.com/insight/content/doi/10.1108/EJM-06-2019-0531/full/html>
206. Roberts AL, Potts HWW, Koutoukidis DA, Smith L, Fisher A. Breast, Prostate, and Colorectal Cancer Survivors' Experiences of Using Publicly Available Physical Activity Mobile Apps: Qualitative Study. *JMIR mHealth and uHealth* [Internet]. 2019 Jan [cited 2024 Feb 22];7(1). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6329432/>
207. Thornton L, Gardner LA, Osman B, Green O, Champion KE, Bryant Z, et al. A Multiple Health Behavior Change, Self-Monitoring Mobile App for Adolescents: Development and Usability Study of the Health4Life App. *JMIR Formative Research* [Internet]. 2021 Apr [cited 2022 Nov 19];5(4). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8076990/>
208. D'Addario M, Baretta D, Zanatta F, Greco A, Steca P. Engagement Features in Physical Activity Smartphone Apps: Focus Group Study With Sedentary People. *JMIR Mhealth Uhealth* [Internet]. 2020 Nov 16;8(11):e20460. Available from: <http://dx.doi.org/10.2196/20460>
209. Davis A, Ellis R. A quasi-experimental investigation of college students' ratings of two physical activity mobile apps with varied behavior change technique quantity. *Digital Health* [Internet]. 2019;5. Available from: <http://dx.doi.org/10.1177/2055207619891347>
210. DeSmet A, De Bourdeaudhuij I, Chastin S, Crombez G, Maddison R, Cardon G. Adults' Preferences for Behavior Change Techniques and Engagement Features in a Mobile App to Promote 24-Hour Movement Behaviors: Cross-Sectional Survey Study. *JMIR mHealth and uHealth* [Internet]. 2019;7(12):e15707. Available from: <http://dx.doi.org/10.2196/15707>
211. Cueto V, Wang CJ, Sanders LM. Impact of a Mobile App-Based Health Coaching and Behavior Change Program on Participant Engagement and Weight Status of Overweight and Obese Children: Retrospective Cohort Study. *JMIR Mhealth Uhealth* [Internet]. 2019 Nov 15;7(11):e14458. Available from: <http://dx.doi.org/10.2196/14458>
212. Valle CG, Nezami BT, Tate DF. Designing in-app messages to nudge behavior change: Lessons learned from a weight management app for young adults. *Organizational Behavior and Human Decision Processes* [Internet]. 2020;161:95–101. Available from: <http://dx.doi.org/10.1016/j.obhdp.2020.10.004>
213. Domin A, Ouzzahra Y, Vögele C. Features and Components Preferred by Adolescents in Smartphone Apps for the Promotion of Physical Activity: Focus Group Study. *JMIR Human Factors* [Internet]. 2022 [cited 2024 Feb 22];9(2). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9227785/>

214. Mummah SA, King AC, Gardner CD, Sutton S. Iterative development of Vegethon: a theory-based mobile app intervention to increase vegetable consumption. *Int J Behav Nutr Phys Act* [Internet]. 2016 Aug 8 [cited 2024 Feb 22];13(1):1–12. Available from: <https://ijbnpa.biomedcentral.com/articles/10.1186/s12966-016-0400-z>
215. Mummah SA, Mathur M, King AC, Gardner CD, Sutton S. Mobile Technology for Vegetable Consumption: A Randomized Controlled Pilot Study in Overweight Adults. *JMIR mHealth and uHealth* [Internet]. 2016 [cited 2024 Feb 22];4(2). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4889871/>
216. Sakane N, Suganuma A, Domichi M, Sukino S, Abe K, Fujisaki A, et al. The Effect of a mHealth App (KENPO-app) for Specific Health Guidance on Weight Changes in Adults With Obesity and Hypertension: Pilot Randomized Controlled Trial. *JMIR mHealth and uHealth* [Internet]. 2023 [cited 2024 Feb 22];11. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10134028/>
217. Schulte MHJ, Boumparis N, Kleiboer A, Wind TR, Olff M, Huizink AC, et al. The effectiveness of a mobile intervention to reduce young adults' alcohol consumption to not exceed low-risk drinking guidelines. *Frontiers in Digital Health* [Internet]. 2022 [cited 2024 Feb 22];4. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9763894/>
218. Whale K, Beasant L, Wright AJ, Yardley L, Wallace LM, Moody L, et al. A Smartphone App for Supporting the Self-management of Daytime Urinary Incontinence in Adolescents: Development and Formative Evaluation Study of URApp. *JMIR Pediatrics and Parenting* [Internet]. 2021 [cited 2024 Feb 22];4(4). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8663506/>
219. Ahonkhai AA, Pierce LJ, Mbugua S, Wasula B, Owino S, Nmoh A, et al. PEERNaija: A Gamified mHealth Behavioral Intervention to Improve Adherence to Antiretroviral Treatment Among Adolescents and Young Adults in Nigeria. *Frontiers in Reproductive Health* [Internet]. 2021;3. Available from: <http://dx.doi.org/10.3389/frph.2021.656507>
220. Buman MP, Epstein DR, Gutierrez M, Herb C, Hollingshead K, Huberty JL, et al. BeWell24: development and process evaluation of a smartphone “app” to improve sleep, sedentary, and active behaviors in US Veterans with increased metabolic risk. *Transl Behav Med* [Internet]. 2016 Sep;6(3):438–48. Available from: <http://dx.doi.org/10.1007/s13142-015-0359-3>
221. Caon M, Prinelli F, Angelini L, Carrino S, Mugellini E, Orte S, et al. PEGASO e-Diary: User Engagement and Dietary Behavior Change of a Mobile Food Record for Adolescents. *Front Nutr* [Internet]. 2022 Mar 17;9:727480. Available from: <http://dx.doi.org/10.3389/fnut.2022.727480>

222. Curtis K, Lebedev A, Aguirre E, Lobitz S. A Medication Adherence App for Children With Sickle Cell Disease: Qualitative Study. *JMIR Mhealth Uhealth* [Internet]. 2019 Jun 18;7(6):e8130. Available from: <http://dx.doi.org/10.2196/mhealth.8130>
223. Du H, Venkatakrishnan A, Youngblood GM, Ram A, Pirolli P. A Group-Based Mobile Application to Increase Adherence in Exercise and Nutrition Programs: A Factorial Design Feasibility Study. *JMIR Mhealth Uhealth* [Internet]. 2016 Jan 15;4(1):e4. Available from: <http://dx.doi.org/10.2196/mhealth.4900>
224. Höchsmann C, Infanger D, Klenk C, Königstein K, Walz SP, Schmidt-Trucksäss A. Effectiveness of a Behavior Change Technique-Based Smartphone Game to Improve Intrinsic Motivation and Physical Activity Adherence in Patients With Type 2 Diabetes: Randomized Controlled Trial. *JMIR Serious Games* [Internet]. 2019 Feb 13;7(1):e11444. Available from: <http://dx.doi.org/10.2196/11444>
225. Hales S, Turner-McGrievy G, Fahim A, Freix A, Wilcox S, Davis RE, et al. A Mixed-Methods Approach to the Development, Refinement, and Pilot Testing of Social Networks for Improving Healthy Behaviors. *JMIR Human Factors* [Internet]. 2016;3(1):e8. Available from: <http://dx.doi.org/10.2196/humanfactors.4512>
226. Lin Y, Mâsse LC. A look at engagement profiles and behavior change: A profile analysis examining engagement with the Aim2Be lifestyle behavior modification app for teens and their families. *Preventive Medicine Reports* [Internet]. 2021;24:101565. Available from: <http://dx.doi.org/10.1016/j.pmedr.2021.101565>
227. Nuijten RCY, Van Gorp P, Borghouts T, Le Blanc P, Van den Berg P, Kemperman A, et al. Preadolescent Students' Engagement With an mHealth Intervention Fostering Social Comparison for Health Behavior Change: Crossover Experimental Study. *J Med Internet Res* [Internet]. 2021 Jul 29;23(7):e21202. Available from: <http://dx.doi.org/10.2196/21202>
228. Nuijten R, Van Gorp P, Khanshan A, Le Blanc P, van den Berg P, Kemperman A, et al. Evaluating the Impact of Adaptive Personalized Goal Setting on Engagement Levels of Government Staff With a Gamified mHealth Tool: Results From a 2-Month Randomized Controlled Trial. *JMIR Mhealth Uhealth* [Internet]. 2022 Mar 31;10(3):e28801. Available from: <http://dx.doi.org/10.2196/28801>
229. Nurmi J, Knittle K, Ginchev T, Khattak F, Helf C, Zwickl P, et al. Engaging Users in the Behavior Change Process With Digitalized Motivational Interviewing and Gamification: Development and Feasibility Testing of the Precious App. *JMIR Mhealth Uhealth* [Internet]. 2020 Jan 30;8(1):e12884. Available from: <http://dx.doi.org/10.2196/12884>

230. Songtaweessin WN, LeGrand S, Bandara S, Piccone C, Wongharn P, Moonwong J, et al. Adaptation of a Theory-Based Social Networking and Gamified App-Based Intervention to Improve Pre-Exposure Prophylaxis Adherence Among Young Men Who Have Sex With Men in Bangkok, Thailand: Qualitative Study. *J Med Internet Res* [Internet]. 2021 Nov 4;23(11):e23852. Available from: <http://dx.doi.org/10.2196/23852>
231. Sporrel K, De Boer RDD, Wang S, Nibbeling N, Simons M, Deutekom M, et al. The Design and Development of a Personalized Leisure Time Physical Activity Application Based on Behavior Change Theories, End-User Perceptions, and Principles From Empirical Data Mining. *Front Public Health* [Internet]. 2020;8:528472. Available from: <http://dx.doi.org/10.3389/fpubh.2020.528472>
232. Whiteley L, Brown L, Lally M, Heck N, van den Berg JJ. A Mobile Gaming Intervention to Increase Adherence to Antiretroviral Treatment for Youth Living With HIV: Development Guided by the Information, Motivation, and Behavioral Skills Model. *JMIR Mhealth Uhealth* [Internet]. 2018 Apr 23;6(4):e96. Available from: <http://dx.doi.org/10.2196/mhealth.8155>
233. Whiteley L, Mena L, Craker LK, Healy MG, Brown LK. Creating a Theoretically Grounded Gaming App to Increase Adherence to Pre-Exposure Prophylaxis: Lessons From the Development of the Viral Combat Mobile Phone Game. *JMIR Serious Games* [Internet]. 2019 Mar 27;7(1):e11861. Available from: <http://dx.doi.org/10.2196/11861>
234. Abdul Khalil NM, Mydin FHM, Moy FM. Healthy adults' views and experiences on behavior change strategies in mobile applications for diet monitoring: A single centre qualitative study. *PLoS One* [Internet]. 2023 [cited 2024 Feb 22];18(11). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10653402/>
235. Beck AK, Kelly PJ, Deane FP, Baker AL, Hides L, Manning V, et al. Developing a mHealth Routine Outcome Monitoring and Feedback App ("SMART Track") to Support Self-Management of Addictive Behaviours. *Front Psychiatry* [Internet]. 2021 [cited 2024 Feb 22];12. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8249767/>
236. Bell L, Garnett C, Bao Y, Cheng Z, Qian T, Perski O, et al. How Notifications Affect Engagement With a Behavior Change App: Results From a Micro-Randomized Trial. *JMIR mHealth and uHealth* [Internet]. 2023 [cited 2024 Feb 22];11. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10337295/>
237. Brindal E, Hendrie GA, Freyne J, Noakes M. A Mobile Phone App Designed to Support Weight Loss Maintenance and Well-Being (MotiMate): Randomized Controlled Trial. *JMIR mHealth and uHealth* [Internet]. 2019 Sep [cited 2024 Feb 22];7(9). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6834303/>

238. Burgermaster M, Desai PM, Heitkemper EM, Juul F, Mitchell EG, Turchioe M, et al. Who needs what (features) when? Personalizing engagement with data-driven self-management to improve health equity. *J Biomed Inform* [Internet]. 2023 Aug;144:104419. Available from: <http://dx.doi.org/10.1016/j.jbi.2023.104419>
239. Chang C, Palermo E, Deswert S, Brown A, Nuske HJ. Money can't buy happiness: A randomized controlled trial of a digital mental health app with versus without financial incentives. *Digital Health* [Internet]. 2023 [cited 2024 Feb 22];9. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10286537/>
240. Curtis KE, Lahiri S, Brown KE. Targeting Parents for Childhood Weight Management: Development of a Theory-Driven and User-Centered Healthy Eating App. *JMIR mHealth and uHealth* [Internet]. 2015 [cited 2022 Aug 15];3(2). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4526951/>
241. Daniels K, Lemmens R, Knippenberg E, Marinus N, Vonck S, Baerts J, et al. Promoting physical activity and a healthy active lifestyle in community-dwelling older adults: a design thinking approach for the development of a mobile health application. *Frontiers in Public Health* [Internet]. 2023 [cited 2024 Feb 22];11. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10724027/>
242. Edney SM, Olds TS, Ryan JC, Vandelanotte C, Plotnikoff RC, Curtis RG, et al. A Social Networking and Gamified App to Increase Physical Activity: Cluster RCT. *Am J Prev Med* [Internet]. 2020 Feb;58(2):e51–62. Available from: <http://dx.doi.org/10.1016/j.amepre.2019.09.009>
243. Edwards EA, Caton H, Lumsden J, Rivas C, Steed L, Pirunsarn Y, et al. Creating a Theoretically Grounded, Gamified Health App: Lessons From Developing the Cigbreak Smoking Cessation Mobile Phone Game. *JMIR Serious Games* [Internet]. 2018 [cited 2024 Feb 22];6(4). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6293248/>
244. Ehlers DK, Huberty JL. Middle-aged women's preferred theory-based features in mobile physical activity applications. *J Phys Act Health* [Internet]. 2014 Sep [cited 2024 Feb 22];11(7). Available from: <https://pubmed.ncbi.nlm.nih.gov/24368818/>
245. Ekezie W, Dallosso H, Saravanan P, Khunti K, Hadjiconstantinou M. Experiences of using a digital type 2 diabetes prevention application designed to support women with previous gestational diabetes. *BMC Health Serv Res* [Internet]. 2021 [cited 2024 Feb 22];21. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8337145/>

246. Fatima S, Augusto JC, Moseley R, Urbonas P, Elliott A, Payne N. Applying motivational techniques for user adherence to adopt a healthy lifestyle in a gamified application. *Entertain Comput* [Internet]. 2023 May;46(100571):100571. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1875952123000265>
247. Garnett C, Crane D, West R, Brown J, Michie S. The development of Drink Less: an alcohol reduction smartphone app for excessive drinkers. *Transl Behav Med* [Internet]. 2019 Apr [cited 2024 Feb 22];9(2):296. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6417151/>
248. Hendrie GA, James-Martin G, Williams G, Brindal E, Whyte B, Crook A. The Development of VegEze: Smartphone App to Increase Vegetable Consumption in Australian Adults. *JMIR Formative Research* [Internet]. 2019 [cited 2024 Feb 22];3(1). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6456819/>
249. Kekkonen M, Korkiakangas E, Laitinen J, Oinas-Kukkonen H. Factors Reducing the Use of a Persuasive mHealth App and How to Mitigate Them: Thematic Analysis. *JMIR Human Factors* [Internet]. 2023 [cited 2024 Feb 22];10. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10337416/>
250. Lawrason SVC, Martin Ginis KA. Evaluating the Feasibility, Acceptability, and Engagement of an mHealth Physical Activity Intervention for Adults With Spinal Cord Injury Who Walk: A Randomized Controlled Trial. *J Sport Exerc Psychol* [Internet]. 2023 Apr 1;45(2):61–76. Available from: <http://dx.doi.org/10.1123/jsep.2022-0087>
251. Lim SL, Tay MHJ, Ong KW, Johal J, Yap QV, Chan YH, et al. Association Between Mobile Health App Engagement and Weight Loss and Glycemic Control in Adults With Type 2 Diabetes and Prediabetes (D’LITE Study): Prospective Cohort Study. *JMIR Diabetes* [Internet]. 2022 [cited 2024 Feb 22];7(3). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9568822/>
252. Litterbach EK, Russell CG, Taki S, Denney-Wilson E, Campbell KJ, Laws RA. Factors Influencing Engagement and Behavioral Determinants of Infant Feeding in an mHealth Program: Qualitative Evaluation of the Growing Healthy Program. *JMIR mHealth and uHealth* [Internet]. 2017 Dec 18 [cited 2022 Dec 1];5(12):e8515. Available from: <https://mhealth.jmir.org/2017/12/e196>
253. Lubans DR, Smith JJ, Skinner G, Morgan PJ. Development and Implementation of a Smartphone Application to Promote Physical Activity and Reduce Screen-Time in Adolescent Boys. *Frontiers in Public Health* [Internet]. 2014 [cited 2024 Feb 22];2. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4032995/>

254. McClure JB, Heffner JL, Krakauer C, Mun S, Klasnja P, Catz SL. Feasibility, Acceptability, and Potential Impact of a Novel mHealth App for Smokers Ambivalent About Quitting: Randomized Pilot Study. *JMIR mHealth and uHealth* [Internet]. 2023 [cited 2024 Feb 22];11. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10365568/>
255. Montilva-Monsalve J, Dimantas B, Perski O, Gutman LM. Barriers and Enablers to the Adoption of a Healthier Diet Using an App: Qualitative Interview Study With Patients With Type 2 Diabetes Mellitus. *JMIR Diabetes* [Internet]. 2023 [cited 2024 Feb 22];8. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10762608/>
256. Nahum-Shani I, Rabbi M, Yap J, Philyaw-Kotov ML, Klasnja P, Bonar EE, et al. Translating Strategies for Promoting Engagement in Mobile Health: A Proof-of-Concept Micro-Randomized Trial. *Health Psychol* [Internet]. 2021 Dec [cited 2023 May 28];40(12):974. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8738098/>
257. Nicholas JC, Ntoumanis N, Smith BJ, Quested E, Stamatakis E, Thøgersen-Ntoumani C. Development and feasibility of a mobile phone application designed to support physically inactive employees to increase walking. *BMC Med Inform Decis Mak* [Internet]. 2021 Jan 21 [cited 2024 Feb 22];21(1):1–16. Available from: <https://bmcmmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-021-01391-3>
258. Patterson K, Keegan R, Davey R, Freene N. Implementing a Sedentary Behavior Change Smartphone App in Cardiac Rehabilitation: A Qualitative Analysis Guided by the Theoretical Domains Framework and Capability, Opportunity, and Motivation-Behavior Model. *J Cardiovasc Nurs* [Internet]. 2024 [cited 2024 Feb 22];39(1):E12. Available from: https://journals.lww.com/jcnjournal/abstract/2024/01000/implementing_a_sedentary_behavior_change.13.aspx
259. Dale LP, White L, Mitchell M, Faulkner G. Smartphone app uses loyalty point incentives and push notifications to encourage influenza vaccine uptake. *Vaccine* [Internet]. 2019 Jul 26;37(32):4594–600. Available from: <http://dx.doi.org/10.1016/j.vaccine.2018.04.018>
260. Resnick D, Schapira MM, Smith JM, Bautista A, Xu C, Jones L, et al. Promoting Collaborative Goal Setting for Cancer Prevention Among Primary Care Patients Through mHealth: Mixed Methods Evaluation of a New App. *JMIR Formative Research* [Internet]. 2021 Jul [cited 2024 Feb 22];5(7). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8319782/>

261. Rohde A, Duensing A, Dawczynski C, Godemann J, Lorkowski S, Brombach C. An App to Improve Eating Habits of Adolescents and Young Adults (Challenge to Go): Systematic Development of a Theory-Based and Target Group-Adapted Mobile App Intervention. *JMIR mHealth and uHealth* [Internet]. 2019 Aug [cited 2024 Feb 22];7(8). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6709564/>
262. Sanders JP, Gokal K, Thomas JJC, Rawstorn JC, Sherar LB, Maddison R, et al. Development of a Mobile Health Snacktivity App to Promote Physical Activity in Inactive Adults (SnackApp): Intervention Mapping and User Testing Study. *JMIR Formative Research* [Internet]. 2023 May 22 [cited 2024 Feb 22];7(1):e41114. Available from: <https://formative.jmir.org/2023/1/e41114>
263. Simons D, De Bourdeaudhuij I, Clarys P, De Cocker K, Vandelanotte C, Deforche B. A Smartphone App to Promote an Active Lifestyle in Lower-Educated Working Young Adults: Development, Usability, Acceptability, and Feasibility Study. *JMIR mHealth and uHealth* [Internet]. 2018 Feb [cited 2024 Feb 22];6(2). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5840478/>
264. Stancu V, Frank DA, Lähteenmäki L, Grunert KG. Motivating consumers for health and fitness: The role of app features. *Journal of Consumer Behaviour* [Internet]. 2022 Nov 1 [cited 2024 Feb 22];21(6):1506–21. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/cb.2108>
265. Tombor I, Beard E, Brown J, Shahab L, Michie S, West R. Randomized factorial experiment of components of the SmokeFree Baby smartphone application to aid smoking cessation in pregnancy. *Transl Behav Med* [Internet]. 2019 Aug [cited 2024 Feb 22];9(4):583. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6629841/>
266. Zimmermann M, Greenberg L, Breland JY. Engagement and Use of a Blended mHealth Intervention for Health Behavior Change. *Int J Behav Med* [Internet]. 2023 May 22 [cited 2024 Feb 22];1–8. Available from: <https://link.springer.com/article/10.1007/s12529-023-10182-1>
267. Zhou L, Bao J, Made Agus Setiawan I, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. *JMIR mHealth and uHealth* [Internet]. 2019 Apr 11 [cited 2024 Mar 12];7(4):e11500. Available from: <https://mhealth.jmir.org/2019/4/e11500>
268. Tong HL, Laranjo L. The use of social features in mobile health interventions to promote physical activity: a systematic review. *npj Digital Medicine* [Internet]. 2018 Sep 4 [cited 2022 Nov 19];1(1):1–10. Available from: <https://www.nature.com/articles/s41746-018-0051-3>

269. Aguiar M, Trujillo M, Chaves D, Álvarez R, Epelde G. mHealth Apps Using Behavior Change Techniques to Self-report Data: Systematic Review. *JMIR mHealth and uHealth* [Internet]. 2022 Sep [cited 2024 Mar 20];10(9). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9508675/>
270. Antezana G, Venning A, Blake V, Smith D, Winsall M, Orlowski S, et al. An evaluation of behaviour change techniques in health and lifestyle mobile applications. *Health Informatics J* [Internet]. 2018 Nov 30 [cited 2024 Mar 20]; Available from: <https://journals.sagepub.com/doi/10.1177/1460458218813726>
271. Saleem M, Kühne L, De Santis KK, Christianson L, Brand T, Busse H. Understanding Engagement Strategies in Digital Interventions for Mental Health Promotion: Scoping Review. *JMIR Ment Health* [Internet]. 2021 Dec 20;8(12):e30000. Available from: <http://dx.doi.org/10.2196/30000>
272. Fulton EA, Kwah KL, Wild S, Brown KE. Lost in Translation: Transforming Behaviour Change Techniques into Engaging Digital Content and Design for the StopApp. *Healthc Pap* [Internet]. 2018 Sep [cited 2022 Jul 24];6(3). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6164572/>
273. Lemke M, de Vries RAJ. Operationalizing Behavior Change Theory as Part of Persuasive Technology: A Scoping Review on Social Comparison. *Front Comput Sci* [Internet]. 2021 [cited 2022 Jul 24];0. Available from: <http://dx.doi.org/10.3389/fcomp.2021.656873>
274. Abuhamdeh S, Csikszentmihalyi M, Jalal B. Enjoying the possibility of defeat: Outcome uncertainty, suspense, and intrinsic motivation. *Motiv Emot* [Internet]. 2015 Feb;39(1):1–10. Available from: <http://link.springer.com/10.1007/s11031-014-9425-2>
275. Druce KL, Dixon WG, McBeth J. Maximizing Engagement in Mobile Health Studies: Lessons Learned and Future Directions. *Rheum Dis Clin North Am* [Internet]. 2019 May;45(2):159–72. Available from: <http://dx.doi.org/10.1016/j.rdc.2019.01.004>
276. Kanstrup AM, Bertelsen P, Jensen MB. Contradictions in digital health engagement: An activity tracker's ambiguous influence on vulnerable young adults' engagement in own health. *Digit Health* [Internet]. 2018 May 9;4:2055207618775192. Available from: <http://dx.doi.org/10.1177/2055207618775192>
277. Perski O, Lumsden J, Garnett C, Blandford A, West R, Michie S. Assessing the Psychometric Properties of the Digital Behavior Change Intervention Engagement Scale in Users of an App for Reducing Alcohol Consumption: Evaluation Study. *J Med Internet Res* [Internet]. 2019 Nov 20 [cited 2023 Jun 29];21(11):e16197. Available from: <https://www.jmir.org/2019/11/e16197>

278. Kelders SM, Kip H, Greeff J. Psychometric Evaluation of the TWente Engagement with Ehealth Technologies Scale (TWEETS): Evaluation Study. *J Med Internet Res* [Internet]. 2020 Oct 6 [cited 2022 Oct 3];22(10):e17757. Available from: <https://www.jmir.org/2020/10/e17757>
279. Perski O, Blandford A, Garnett C, Crane D, West R, Michie S. A self-report measure of engagement with digital behavior change interventions (DBCIs): development and psychometric evaluation of the “DBCI Engagement Scale.” *Transl Behav Med* [Internet]. 2019 Mar 30 [cited 2023 Mar 27];10(1):267–77. Available from: <https://academic.oup.com/tbm/article-pdf/10/1/267/32315584/ibz039.pdf>
280. Kelders SM, Kip H. Development and initial validation of a scale to measure engagement with eHealth technologies. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* [Internet]. New York, NY, USA: ACM; 2019. Available from: <https://dl.acm.org/doi/10.1145/3290607.3312917>
281. Nahum-Shani I, Yoon C. Toward the science of engagement with digital interventions. *Curr Dir Psychol Sci* [Internet]. 2024 Jun 3; Available from: <https://journals.sagepub.com/doi/10.1177/09637214241254328>
282. Kaveladze BT, Wasil AR, Bunyi JB, Ramirez V, Schueller SM. User Experience, Engagement, and Popularity in Mental Health Apps: Secondary Analysis of App Analytics and Expert App Reviews. *JMIR Human Factors* [Internet]. 2022 [cited 2024 Mar 29];9(1). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8844980/>
283. Morrison L, Moss-Morris R, Michie S, Yardley L. Optimizing engagement with Internet-based health behaviour change interventions: Comparison of self-assessment with and without tailored feedback using a mixed methods approach. *Br J Health Psychol* [Internet]. 2014 Nov 1 [cited 2024 Feb 1];19(4):839–55. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/bjhp.12083>
284. Nguyen TD, Attkisson CC, Stegner BL. Assessment of patient satisfaction: Development and refinement of a Service Evaluation Questionnaire. *Eval Program Plann* [Internet]. 1983 Jan 1 [cited 2024 Feb 1];6(3-4):299–313. Available from: [http://dx.doi.org/10.1016/0149-7189\(83\)90010-1](http://dx.doi.org/10.1016/0149-7189(83)90010-1)
285. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. *Eval Program Plann* [Internet]. 1979;2(3):197–207. Available from: [http://dx.doi.org/10.1016/0149-7189\(79\)90094-6](http://dx.doi.org/10.1016/0149-7189(79)90094-6)
286. Agarwal R, Karahanna E. Time flies when you’re having fun: Cognitive absorption and beliefs about information technology usage. *Miss Q* [Internet]. 2000 Dec;24(4):665. Available from: <https://www.jstor.org/stable/3250951?origin=crossref>

287. Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction* [Internet]. 1995 Jan 1 [cited 2024 Feb 1]; Available from: <https://www.tandfonline.com/doi/abs/10.1080/10447319509526110>
288. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry* [Internet]. 2000 Jun;31(2):73–86. Available from: [http://dx.doi.org/10.1016/s0005-7916\(00\)00012-4](http://dx.doi.org/10.1016/s0005-7916(00)00012-4)
289. Lefebvre RC, Tada Y, Hilfiker SW, Baur C. The Assessment of User Engagement with eHealth Content: The eHealth Engagement Scale. *J Comput Mediat Commun* [Internet]. 2010 Jul 1 [cited 2024 Feb 1];15(4):666–81. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1083-6101.2009.01514.x>
290. Lin A, Gregor S, Ewing M. Developing a scale to measure the enjoyment of Web experiences. *J Interact Mark* [Internet]. 2008 Nov;22(4):40–57. Available from: <http://journals.sagepub.com/doi/10.1002/dir.20120>
291. Jackson SA, Marsh HW. Development and validation of a scale to measure optimal experience: The Flow State Scale. *J Sport Exerc Psychol* [Internet]. 1996 Mar;18(1):17–35. Available from: <https://journals.humankinetics.com/view/journals/jsep/18/1/article-p17.xml>
292. Brockmyer JH, Fox CM, Curtiss KA, McBroom E, Burkhart KM, Pidruzny JN. The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *J Exp Soc Psychol* [Internet]. 2009 Jul;45(4):624–34. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0022103109000444>
293. Jennett C, Cox AL, Cairns P, Dhoparee S, Epps A, Tijs T, et al. Measuring and defining the experience of immersion in games. *Int J Hum Comput Stud* [Internet]. 2008 Sep;66(9):641–61. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1071581908000499>
294. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. *JMIR mHealth and uHealth* [Internet]. 2015 Mar 11 [cited 2024 Feb 1];3(1):e3422. Available from: <https://mhealth.jmir.org/2015/1/e27>
295. Zaichkowsky JL. The personal involvement inventory: Reduction, revision, and application to advertising. *J Advert* [Internet]. 1994 Dec;23(4):59–70. Available from: <http://www.tandfonline.com/doi/abs/10.1080/00913367.1943.10673459>

296. Lewis JR. Psychometric Evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. Proceedings of the Human Factors Society Annual Meeting [Internet]. 1992 Oct 1 [cited 2024 Feb 1]; Available from: <https://journals.sagepub.com/doi/10.1177/154193129203601617>
297. Tait L, Birchwood M, Trower P. A new scale (SES) to measure engagement with community mental health services. J Ment Health [Internet]. 2002;11(2):191–8. Available from: <http://dx.doi.org/10.1080/09638230020023570-2>
298. Simões de Almeida R, Marques A. User engagement in mobile apps for people with schizophrenia: A scoping review. Front Digit Health [Internet]. 2022;4:1023592. Available from: <http://dx.doi.org/10.3389/fdgth.2022.1023592>
299. Brooke J. SUS - A quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry [Internet]. 1996. Available from: https://www.researchgate.net/publication/319394819_SUS_--_a_quick_and_dirty_usability_scale
300. Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. Int J Man Mach Stud [Internet]. 1993 Mar 1 [cited 2024 Feb 1];38(3):475–87. Available from: <http://dx.doi.org/10.1006/imms.1993.1022>
301. Venkatesh V, Davis FD. A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. Manage Sci [Internet]. 2000 Feb;46(2):186–204. Available from: <https://pubsonline.informs.org/doi/10.1287/mnsc.46.2.186.11926>
302. Shirk SR, Gudmundsen G, Kaplinski HC, McMakin DL. Alliance and outcome in cognitive-behavioral therapy for adolescent depression. J Clin Child Adolesc Psychol [Internet]. 2008 Jul;37(3):631–9. Available from: <http://dx.doi.org/10.1080/15374410802148061>
303. Lund AM. Measuring usability with the USE questionnaire. Usability interface [Internet]. 2001;8(2):3–6. Available from: https://www.researchgate.net/publication/230786746_Measuring_Usability_with_the_USE_Questionnaire?enrichId=rgreq-ff5924c9c769fc2ca8098a360f9083c0-XXX&enrichSource=Y292ZXJQYWdlOzIzMDc4Njc0NjBUzozMzkyMzYxMDU1Mzk1ODRAMTQ1Nzg5MTU5ODIyMA%3D%3D&el=1_x_2&_esc=publicationCoverPdf
304. O'Brien HL, Toms EG. The development and evaluation of a survey to measure user engagement. J Am Soc Inf Sci Technol [Internet]. 2010 Jan;61(1):50–69. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/asi.21229>

305. O'Brien HL, Cairns P, Hall M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *Int J Hum Comput Stud* [Internet]. 2018 Apr;112:28–39. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1071581918300041>
306. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. In: Holzinger A, editor. *CI and Usability for Education and Work USAB 2008* [Internet]. Berlin, Heidelberg: Springer; 2008. p. 63–76. (Lecture notes in computer science; vol. 5298). Available from: http://link.springer.com/10.1007/978-3-540-89350-9_6
307. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young SL. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Health* [Internet]. 2018 Jun 11;6:149. Available from: <http://dx.doi.org/10.3389/fpubh.2018.00149>
308. Fredricks JA, Blumenfeld PC, Paris AH. School engagement: Potential of the concept, state of the evidence. *Rev Educ Res* [Internet]. 2004 Mar;74(1):59–109. Available from: <http://journals.sagepub.com/doi/10.3102/00346543074001059>
309. Hollingshead A, Williamson P, Carnahan C. Cognitive and emotional engagement for students with severe intellectual disability defined by the scholars with expertise in the field. *Res Pract Persons Severe Disabl* [Internet]. 2018 Dec;43(4):269–84. Available from: <http://journals.sagepub.com/doi/10.1177/1540796918812803>
310. Makransky G, Petersen GB. The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educ Psychol Rev* [Internet]. 2021 Jan 6 [cited 2023 Nov 17];33(3):937–58. Available from: <https://link.springer.com/article/10.1007/s10648-020-09586-2>
311. Silvia PJ. *Exploring the Psychology of Interest* [Internet]. Oxford University Press, USA; 2006. 263 p. Available from: <https://play.google.com/store/books/details?id=YxhnDAAAQBAJ>
312. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. *J Am Soc Inf Sci Technol* [Internet]. 2008 Apr;59(6):938–55. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/asi.20801>
313. Triberti S, Kelders SM, Gaggioli A. User engagement. In: Kip H, Beerlage-de Jong N, van Gemert-Pijnen L (jewc), Sanderman R, Kelders SM, editors. *eHealth Research, Theory and Development* [Internet]. Routledge; 2018. p. 271–89. Available from: <https://www.taylorfrancis.com/books/9781315385891/chapters/10.4324/9781315385907-13>

314. Schwarz AF, Huertas-Delgado FJ, Cardon G, DeSmet A. Design Features Associated with User Engagement in Digital Games for Healthy Lifestyle Promotion in Youth: A Systematic Review of Qualitative and Quantitative Studies. *Games Health J* [Internet]. 2020 Jun;9(3):150–63. Available from: <http://dx.doi.org/10.1089/g4h.2019.0058>
315. Schunk DH, Usher EL. Social Cognitive Theory and Motivation. In: Ryan RM, editor. *The Oxford Handbook of Human Motivation* [Internet]. Oxford University Press; 2019. p. 9–26. Available from: <https://academic.oup.com/edited-volume/28143/chapter/212903358>
316. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* [Internet]. 2000 Jan;55(1):68–78. Available from: <http://dx.doi.org/10.1037//0003-066x.55.1.68>
317. Wigfield A, Eccles JS. Expectancy-Value Theory of Achievement Motivation. *Contemp Educ Psychol* [Internet]. 2000 Jan [cited 2023 Jul 1];25(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/10620382/>
318. Rodgers WM, Markland D, Selzler AM, Murray TC, Wilson PM. Distinguishing Perceived Competence and Self-Efficacy: An Example From Exercise. *Res Q Exerc Sport* [Internet]. 2014 Oct 2 [cited 2024 Feb 3]; Available from: <https://www.tandfonline.com/doi/abs/10.1080/02701367.2014.961050>
319. Kavanagh DJ, Teixeira H, Connolly J, Andrade J, May J, Godfrey S, et al. The Motivational Thought Frequency Scales for increased physical activity and reduced high-energy snacking. *Br J Health Psychol* [Internet]. 2020 Sep 1 [cited 2024 Feb 12];25(3):558–75. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/bjhp.12422>
320. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC Health Serv Res* [Internet]. 2017 Jan 26 [cited 2021 Jan 20];17(1):1–13. Available from: <https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-017-2031-8>
321. Tullis T, Albert B. Performance Metrics. In: Tullis T, Albert B, editors. *Measuring the User Experience* [Internet]. Elsevier; 2013. p. 63–97. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9780124157811000042>
322. Maqbool B, Herold S. Potential effectiveness and efficiency issues in usability evaluation within digital health: A systematic literature review. *J Syst Softw* [Internet]. 2024 Feb;208(111881):111881. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0164121223002765>

323. Wu R, Ferguson AM, Inzlicht M. Do humans prefer cognitive effort over doing nothing? *J Exp Psychol Gen* [Internet]. 2023 Apr;152(4):1069–79. Available from: <http://dx.doi.org/10.1037/xge0001320>
324. Lai VT, Hagoort P, Casasanto D. Affective Primacy vs. Cognitive Primacy: Dissolving the Debate. *Front Psychol* [Internet]. 2012 Jul 17 [cited 2023 Jul 4];3:26172. Available from: <http://dx.doi.org/10.3389/fpsyg.2012.00243>
325. Whissell C. Emotion and Cognition. In: *Engaging with Emotion* [Internet]. Springer International Publishing; 2023 [cited 2023 Oct 23]. p. 141–56. Available from: https://link.springer.com/chapter/10.1007/978-3-031-21398-4_9
326. Fox E. Perspectives from affective science on understanding the nature of emotion. *Brain and Neuroscience Advances* [Internet]. 2018 [cited 2023 Jul 4];2. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7058241/>
327. Barrett LF. The theory of constructed emotion: an active inference account of interoception and categorization. *Soc Cogn Affect Neurosci* [Internet]. 2017 Jan 1;12(1):1–23. Available from: <http://dx.doi.org/10.1093/scan/nsw154>
328. Adolphs R. How should neuroscience study emotions? by distinguishing emotion states, concepts, and experiences. *Soc Cogn Affect Neurosci* [Internet]. 2017 Jan 1;12(1):24–31. Available from: <http://dx.doi.org/10.1093/scan/nsw153>
329. Pessoa L. Embracing integration and complexity: placing emotion within a science of brain and behaviour. *Cognition and Emotion* [Internet]. 2018 Sep 11 [cited 2023 Jul 4];33(1):55–60. Available from: <https://www.tandfonline.com/doi/abs/10.1080/02699931.2018.1520079>
330. Smith R, Moutoussis M, Bilek E. Simulating the computational mechanisms of cognitive and behavioral psychotherapeutic interventions: insights from active inference. *Sci Rep* [Internet]. 2021 May 12 [cited 2023 Oct 20];11(1):1–16. Available from: <https://www.nature.com/articles/s41598-021-89047-0>
331. Fisher LB, Curreri AJ, Tan EK, Sprich SE. Cognitive Techniques. In: Sprich SE, Petersen T, Wilhelm S, editors. *The Massachusetts General Hospital Handbook of Cognitive Behavioral Therapy* [Internet]. Springer International Publishing; 2023 [cited 2023 Oct 20]. p. 19–38. Available from: https://link.springer.com/chapter/10.1007/978-3-031-29368-9_3
332. Bertollo M, Filho E, Terry PC. *Advancements in Mental Skills Training* [Internet]. Routledge; 2020. 302 p. Available from: <https://play.google.com/store/books/details?id=qeruDwAAQBAJ>

333. Early BP, Grady MD. Embracing the Contribution of Both Behavioral and Cognitive Theories to Cognitive Behavioral Therapy: Maximizing the Richness. *Clin Soc Work J* [Internet]. 2016 Jun 14 [cited 2023 Oct 20];45(1):39–48. Available from: <https://link.springer.com/article/10.1007/s10615-016-0590-5>
334. Skinner EA, Raine KE. Unlocking the Positive Synergy Between Engagement and Motivation. In: Reschly AL, Christenson SL, editors. *Handbook of Research on Student Engagement* [Internet]. Springer International Publishing; 2022 [cited 2023 Oct 23]. p. 25–56. Available from: https://link.springer.com/chapter/10.1007/978-3-031-07853-8_2
335. Nadal C, Sas C, Doherty G. Technology Acceptance in Mobile Health: Scoping Review of Definitions, Models, and Measurement. *J Med Internet Res* [Internet]. 2020 Jul [cited 2023 Jul 7];22(7). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7381045/>
336. Crane D, Garnett C, Michie S, West R, Brown J. A smartphone app to reduce excessive alcohol consumption: Identifying the effectiveness of intervention components in a factorial randomised control trial. *Sci Rep* [Internet]. 2018 [cited 2022 Mar 9];8. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5847600/>
337. Yin RK. *Case Study Research and Applications: Design and Methods 6e* [Internet]. SAGE Publications; 2018. 352 p. Available from: <https://play.google.com/store/books/details?id=uX1ZDwAAQBAJ>
338. Hull L, Goulding L, Khadjesari Z, Davis R, Healey A, Bakolis I, et al. Designing high-quality implementation research: development, application, feasibility and preliminary evaluation of the implementation science research development (ImpRes) tool and guide. *Implement Sci* [Internet]. 2019 [cited 2022 Aug 22];14. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6693182/>
339. Hamilton AB, Mittman BS. Implementation science in health care. In: Brownson RC, Colditz GA, Proctor EK, editors. *Dissemination and Implementation Research in Health: Translating Science to Practice (2nd edn)* [Internet]. Oxford University Press; 2017. p. 385–400. Available from: <https://academic.oup.com/book/26456/chapter/194888661>
340. Landes SJ, McBain SA, Curran GM. Reprint of: An introduction to effectiveness-implementation hybrid designs. *Psychiatry Res* [Internet]. 2020 Jan 1 [cited 2023 May 10];283:112630. Available from: <http://dx.doi.org/10.1016/j.psychres.2019.112630>

341. Greenhalgh T, Wherton J, Papoutsis C, Lynch J, Hughes G, A'Court C, et al. Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. *J Med Internet Res* [Internet]. 2017 Nov 1;19(11):e367. Available from: <http://dx.doi.org/10.2196/jmir.8775>
342. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health* [Internet]. 1999 Sep;89(9):1322–7. Available from: <http://dx.doi.org/10.2105/ajph.89.9.1322>
343. What is RE-AIM? [Internet]. RE-AIM. [cited 2021 Nov 5]. Available from: <https://re-aim.org/learn/what-is-re-aim/>
344. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM Planning and Evaluation Framework: Adapting to New Science and Practice With a 20-Year Review. *Front Public Health* [Internet]. 2019 Mar 29;7:64. Available from: <http://dx.doi.org/10.3389/fpubh.2019.00064>
345. Meinert E, Milne-Ives M, Lim E, Higham A, Boege S, de Pennington N, et al. Accuracy and safety of an autonomous artificial intelligence clinical assistant conducting telemedicine follow-up assessment for cataract surgery. *EClinicalMedicine* [Internet]. 2024 Jul;73(102692):102692. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2589537024002712>
346. McPherson AC, Hamilton J, Kingsnorth S, Knibbe TJ, Peters M, Swift JA, et al. Communicating with children and families about obesity and weight-related topics: a scoping review of best practices. *Obes Rev* [Internet]. 2017 Feb;18(2):164–82. Available from: <http://dx.doi.org/10.1111/obr.12485>
347. Sjunnestrand M, Nordin K, Eli K, Nowicka P, Ek A. Planting a seed - child health care nurses' perceptions of speaking to parents about overweight and obesity: a qualitative study within the STOP project. *BMC Public Health* [Internet]. 2019 [cited 2022 Mar 31];19. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6842180/>
348. Serban CL, Putnok S, Ek A, Eli K, Nowicka P, Chirita-Emandi A. Making Childhood Obesity a Priority: A Qualitative Study of Healthcare Professionals' Perspectives on Facilitating Communication and Improving Treatment. *Frontiers in Public Health* [Internet]. 2021 [cited 2022 Mar 31];9. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8321411/>
349. Bradbury D, Chisholm A, Watson PM, Bundy C, Bradbury N, Birtwistle S. Barriers and facilitators to health care professionals discussing child weight with parents: A meta-synthesis of qualitative studies. *Br J Health Psychol* [Internet]. 2018 Sep;23(3):701–22. Available from: <http://dx.doi.org/10.1111/bjhp.12312>

350. Lampe EW, Abber SR, Forman EM, Manasse SM. Guidelines for caregivers and healthcare professionals on speaking to children about overweight and obesity: a systematic review of the gray literature. *Transl Behav Med* [Internet]. 2020 Oct 12;10(5):1144–54. Available from: <http://dx.doi.org/10.1093/tbm/ibaa012>
351. King D, Rahman E, Potter A. NoObesity Apps – From Approach to Finished App. In: Arai K, Bhatia R, Kapoor S, editors. *Proceedings of the Future Technologies Conference (FTC) 2018* [Internet]. Springer International Publishing; 2019 [cited 2022 Aug 21]. p. 1145–57. (Advances in Intelligent Systems & Computing; vol. 881). Available from: https://link.springer.com/chapter/10.1007/978-3-030-02683-7_84
352. NoObesity Apps [Internet]. NHS Health Education England. [cited 2021 Feb 6]. Available from: <https://wessex.hee.nhs.uk/wider-workforce/population-health/noobesity-digital-app/>
353. Meinert E, Rahman E, Potter A, Lawrence W, Van Velthoven M. Acceptability and Usability of the Mobile Digital Health App NoObesity for Families and Health Care Professionals: Protocol for a Feasibility Study. *JMIR Res Protoc* [Internet]. 2020 Jul 22 [cited 2021 Mar 22];9(7):e18068. Available from: <https://www.researchprotocols.org/2020/7/e18068>
354. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* [Internet]. 2001 Sep;16(9):606–13. Available from: <http://dx.doi.org/10.1046/j.1525-1497.2001.016009606.x>
355. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* [Internet]. 2006 May 22;166(10):1092–7. Available from: <http://dx.doi.org/10.1001/archinte.166.10.1092>
356. Mozilla Foundation. Top Mental Health and Prayer Apps Fail Spectacularly at Privacy, Security [Internet]. *Privacy Not Included. 2022 [cited 2024 Feb 16]. Available from: <https://foundation.mozilla.org/en/privacynotincluded/articles/top-mental-health-and-prayer-apps-fail-spectacularly-at-privacy-security/>
357. Wysa [Internet]. Mozilla Foundation. [cited 2024 Feb 16]. Available from: <https://foundation.mozilla.org/en/privacynotincluded/wysa/>
358. Inkster B, Sarda S, Subramanian V. An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR mHealth and uHealth* [Internet]. 2018 Nov 23 [cited 2021 Mar 2];6(11):e12106. Available from: <https://mhealth.jmir.org/2018/11/e12106>
359. Selby E, Wade C. Pharmacy2U Report: Provision of AI enabled digital support to regular medication users. 2020.

360. Malik T, Ambrose AJ, Sinha C. Evaluating User Feedback for an Artificial Intelligence-Enabled, Cognitive Behavioral Therapy-Based Mental Health App (Wysa): Qualitative Thematic Analysis. *JMIR Hum Factors* [Internet]. 2022 Apr 12 [cited 2024 May 16];9(2):e35668–e35668. Available from: <https://europepmc.org/article/med/35249886?client=bot>
361. Moustafa GA, Borkar DS, Borboli-Gerogiannis S, Greenstein SH, Lorch AC, Vasan RA, et al. Optimization of cataract surgery follow-up: A standard set of questions can predict unexpected management changes at postoperative week one. *PLoS One* [Internet]. 2019 [cited 2022 Aug 25];14(9). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6752806/>
362. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med* [Internet]. 2014 Sep;89(9):1245–51. Available from: <http://dx.doi.org/10.1097/ACM.0000000000000388>
363. Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* [Internet]. 2004 Mar [cited 2022 Dec 7];94(3). Available from: <https://pubmed.ncbi.nlm.nih.gov/14998794/>
364. NoObesity App Evaluation [Internet]. NHS Health Education England South East - Working across Wessex. 2020 [cited 2021 Oct 25]. Available from: <https://wessex.hee.nhs.uk/wider-workforce/population-health/noobesity-digital-app/evaluation/>
365. Blank G. OxIS 2019 Questionnaire. All Parts. 2019 Jan 30 [cited 2021 Oct 25]; Available from: <https://papers.ssrn.com/abstract=3522118>
366. Bandura A. Guide for constructing self-efficacy scales. In: Urdan T, Pajares F, editors. *Self-Efficacy Beliefs of Adolescents* [Internet]. Greenwich, CT: Information Age Publishing; 2006. p. 307–37. Available from: https://play.google.com/store/books/details?id=P_onDwAAQBAJ
367. Brooks J, McCluskey S, Turley E, King N. The Utility of Template Analysis in Qualitative Psychology Research. *Qual Res Psychol* [Internet]. 2015 Apr 3;12(2):202–22. Available from: <http://dx.doi.org/10.1080/14780887.2014.955224>
368. Braun V, Clarke V. Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. *International Journal of Transgender Health* [Internet]. 2022 Oct 25 [cited 2023 May 24]; Available from: <https://www.tandfonline.com/doi/abs/10.1080/26895269.2022.2129597>
369. SocioCultural Research Consultants, LLC. Dedoose Version 9.0.17 web application for managing, analyzing, and presenting qualitative and mixed method research data [Internet]. Los Angeles, CA; 2021 [cited 2021 Oct 28]. Available from: www.dedoose.com

370. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: Striving to Meet the Trustworthiness Criteria. *Int J Qual Methods* [Internet]. 2017 Dec;16(1):160940691773384. Available from: <http://journals.sagepub.com/doi/10.1177/1609406917733847>
371. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* [Internet]. 2006;3(2):77–101. Available from: <http://dx.doi.org/10.1191/1478088706qp063oa>
372. Attride-Stirling J. Thematic networks: an analytic tool for qualitative research. *Qual Res* [Internet]. 2001 Dec;1(3):385–405. Available from: <http://journals.sagepub.com/doi/10.1177/146879410100100307>
373. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. *Implement Sci* [Internet]. 2017 Jun 21 [cited 2023 Feb 7];12(1):1–18. Available from: <https://implementationscience.biomedcentral.com/articles/10.1186/s13012-017-0605-9>
374. Taylor-Powell E, Henert E. Developing a Logic Model: Teaching and Training Guide [Internet]. University of Wisconsin System; 2008. Available from: https://books.google.com/books/about/Developing_a_Logic_Model.html?hl=&id=eKN5PgAACAAJ
375. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med* [Internet]. 2020 Sep;26(9):1364–74. Available from: <http://dx.doi.org/10.1038/s41591-020-1034-x>
376. Sealed Envelope [Internet]. [cited 2024 Jan 23]. Available from: <https://www.sealedenvelope.com/>
377. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* [Internet]. 2011 Dec;20(10):1727–36. Available from: <http://dx.doi.org/10.1007/s11136-011-9903-x>
378. Brazier J. Is the EQ-5D fit for purpose in mental health? *Br J Psychiatry* [Internet]. 2010 Nov;197(5):348–9. Available from: <http://dx.doi.org/10.1192/bjp.bp.110.082453>
379. Saldana J. The Coding Manual for Qualitative Researchers [Internet]. SAGE; 2021. 441 p. Available from: https://books.google.com/books/about/The_Coding_Manual_for_Qualitative_Resear.html?hl=&id=RwcVEAAAQBAJ

380. Khavandi S, Lim E, Mole G. 110 Patient acceptability of telephone follow up after cataract surgery. *BMJ Leader* [Internet]. 2020 Nov 1 [cited 2020 Dec 14];4(Suppl 1). Available from: https://bmjleader.bmj.com/content/4/Suppl_1/A41.3.abstract
381. Khavandi S, Lim E, Higham A, de Pennington N, Bindra M, Maling S, et al. User-acceptability of an automated telephone call for post-operative follow-up after uncomplicated cataract surgery. *Eye* [Internet]. 2022 Oct 23 [cited 2023 Apr 26]; Available from: <https://pubmed.ncbi.nlm.nih.gov/36274084/>
382. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med* [Internet]. 2022 May 18 [cited 2024 Feb 15];28(5):924–33. Available from: <https://www.nature.com/articles/s41591-022-01772-9>
383. Bowling A. *Research Methods in Health: Investigating Health and Health Services* [Internet]. McGraw-Hill Education (UK); 2014. 538 p. Available from: <https://play.google.com/store/books/details?id=woNFBgAAQBAJ>
384. Lewis JR. The System Usability Scale: Past, Present, and Future. *International Journal of Human–Computer Interaction* [Internet]. 2018;34(7):577–90. Available from: <http://dx.doi.org/10.1080/10447318.2018.1455307>
385. Parmanto B, Lewis AN Jr, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). *Int J Telerehabil* [Internet]. 2016 Jul 1;8(1):3–10. Available from: <http://dx.doi.org/10.5195/ijt.2016.6196>
386. What Is Net Promoter? [Internet]. Net Promoter Network. [cited 2022 Aug 16]. Available from: <https://www.netpromoter.com/know/>
387. StataCorp. *Stata Statistical Software: Release 17* [Internet]. College Station, TX: StataCorp LLC.; 2021 [cited 2022 Aug 22]. Available from: <https://www.stata.com/support/faqs/resources/citing-software-documentation-faqs/>
388. QSR International Pty Ltd. NVivo (released in March 2020) [Internet]. 2020. Available from: <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home>
389. Jones KC, Burns A. Unit costs of health and social care 2021 [Internet]. Personal Social Services Research Unit; 2021. Available from: <https://kar.kent.ac.uk/id/eprint/92342>
390. Ridder HG. The theory contribution of case study research designs. *Business Research* [Internet]. 2017 Feb 16 [cited 2023 Jun 22];10(2):281–305. Available from: <https://link.springer.com/article/10.1007/s40685-017-0045-z>

391. Cruzes DS, Dybå T, Runeson P, Höst M. Case studies synthesis: a thematic, cross-case, and narrative synthesis worked example. *Empir Softw Eng* [Internet]. 2015 Dec;20(6):1634–65. Available from: <http://link.springer.com/10.1007/s10664-014-9326-8>
392. Killingback C, Tsofliou F, Clark C. Older people’s adherence to community-based group exercise programmes: a multiple-case study. *BMC Public Health* [Internet]. 2017 [cited 2023 Jul 11];17. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5264485/>
393. Fischl C, Blusi M, Lindgren H, Nilsson I. Tailoring to support digital technology-mediated occupational engagement for older adults - a multiple case study. *Scand J Occup Ther* [Internet]. 2020 Nov;27(8):577–90. Available from: <http://dx.doi.org/10.1080/11038128.2020.1760347>
394. Thematic Synthesis [Internet]. Cochrane Training. [cited 2023 Nov 13]. Available from: <https://training.cochrane.org/resource/thematic-synthesis>
395. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* [Internet]. 2008 Jul 10 [cited 2023 Nov 13];8(1):1–10. Available from: <https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-8-45>
396. McCormack GR, Rock M, Swanson K, Burton L, Massolo A. Physical activity patterns in urban neighbourhood parks: insights from a multiple case study. *BMC Public Health* [Internet]. 2014 Sep 17;14:962. Available from: <http://dx.doi.org/10.1186/1471-2458-14-962>
397. Guest GS, MacQueen KM, Namey EE. Writing up thematic analyses. In: *Applied Thematic Analysis* [Internet]. Thousand Oaks, CA: SAGE Publications; 2012. Available from: <https://methods.sagepub.com/book/applied-thematic-analysis>
398. Byrne D. A worked example of Braun and Clarke’s approach to reflexive thematic analysis. *Qual Quant* [Internet]. 2022 Jun;56(3):1391–412. Available from: <https://link.springer.com/10.1007/s11135-021-01182-y>
399. Maguire M, Delahunt B. Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *AISHE-J* [Internet]. 2017 Oct 31 [cited 2024 May 9];9(3). Available from: <https://ojs.aishe.org/index.php/aishe-j/article/view/335>
400. UKHSA operational launch – rebranding digital channels [Internet]. UK Health Security Agency. 2021 [cited 2021 Nov 4]. Available from: <https://ukhsa.blog.gov.uk/2021/09/23/ukhsa-operational-launch-rebranding-digital-channels/>

401. Kouvari M, Karipidou M, Tsiampalis T, Mamalaki E, Poulimeneas D, Bathrellou E, et al. Digital Health Interventions for Weight Management in Children and Adolescents: Systematic Review and Meta-analysis. *J Med Internet Res* [Internet]. 2022 Feb 14 [cited 2023 Dec 6];24(2):e30675. Available from: <https://www.jmir.org/2022/2/e30675>
402. González ODJ, Tugault-Lafleur CN, Jean Buckler E, Hamilton J, Ho J, Buchholz A, et al. The Aim2Be mHealth Intervention for Children With Overweight or Obesity and Their Parents: Person-Centered Analyses to Uncover Digital Phenotypes. *J Med Internet Res* [Internet]. 2022 Jun [cited 2023 Dec 6];24(6). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9221987/>
403. Tripicchio GL, Kay M, Herring S, Cos T, Bresnahan C, Gartner D, et al. Development and Preliminary Feasibility of iByte4Health: A Mobile Health (mHealth) Pediatric Obesity Prevention Intervention to Engage Parents with Low-Income of Children 2–9 Years. *Nutrients* [Internet]. 2021 Dec [cited 2023 Dec 6];13(12). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8708273/>
404. König LM, Attig C, Franke T, Renner B. Barriers to and Facilitators for Using Nutrition Apps: Systematic Review and Conceptual Framework. *JMIR mHealth and uHealth* [Internet]. 2021 Jun 24 [cited 2022 Nov 19];9(6):e20037. Available from: <https://mhealth.jmir.org/2021/6/e20037>
405. Shabir H, D’Costa M, Mohiaddin Z, Moti Z, Rashid H, Sadowska D, et al. The Barriers and Facilitators to the Use of Lifestyle Apps: A Systematic Review of Qualitative Studies. *European Journal of Investigation in Health, Psychology and Education* [Internet]. 2022;12(2):144–65. Available from: <http://dx.doi.org/10.3390/ejihpe12020012>
406. Mauch CE, Laws RA, Prichard I, Maeder AJ, Wycherley TP, Golley RK. Commercially Available Apps to Support Healthy Family Meals: User Testing of App Utility, Acceptability, and Engagement. *JMIR Mhealth Uhealth* [Internet]. 2021 May 7;9(5):e22990. Available from: <http://dx.doi.org/10.2196/22990>
407. Taki S, Russell CG, Lymer S, Laws R, Campbell K, Appleton J, et al. A Mixed Methods Study to Explore the Effects of Program Design Elements and Participant Characteristics on Parents’ Engagement With an mHealth Program to Promote Healthy Infant Feeding: The Growing Healthy Program. *Front Endocrinol* [Internet]. 2019 Jun 25;10:397. Available from: <http://dx.doi.org/10.3389/fendo.2019.00397>

408. Thomas K, Neher M, Alexandrou C, Müssener U, Henriksson H, Löf M. Mobile phone-based lifestyle support for families with young children in primary health care (MINISTOP 2.0): Exploring behavioral change determinants for implementation using the COM-B model. *FrontHealth Serv* [Internet]. 2022 [cited 2022 Dec 1];0. Available from: <http://dx.doi.org/10.3389/frhs.2022.951879>
409. Willmott TJ, Pang B, Rundle-Thiele S. Capability, opportunity, and motivation: an across contexts empirical examination of the COM-B model. *BMC Public Health* [Internet]. 2021 [cited 2022 Nov 19];21. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8164288/>
410. Taghizadeh S, Farhangi MA, Khodayari-Zarnaq R. Stakeholders perspectives of barriers and facilitators of childhood obesity prevention policies in Iran: A Delphi method study. *BMC Public Health* [Internet]. 2021 Dec 11;21(1):2260. Available from: <http://dx.doi.org/10.1186/s12889-021-12282-7>
411. Tran S, Smith L, El-Den S, Carter S. The Use of Gamification and Incentives in Mobile Health Apps to Improve Medication Adherence: Scoping Review. *JMIR mHealth and uHealth* [Internet]. 2022 Feb [cited 2022 Nov 19];10(2). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8902658/>
412. Lewis ZH, Swartz MC, Lyons EJ. What's the Point?: A Review of Reward Systems Implemented in Gamification Interventions. *Games Health J* [Internet]. 2016 Apr;5(2):93–9. Available from: <http://dx.doi.org/10.1089/g4h.2015.0078>
413. Chow CY, Riantiningtyas RR, Kanstrup MB, Papavasileiou M, Liem GD, Olsen A. Can games change children's eating behaviour? A review of gamification and serious games. *Food Qual Prefer* [Internet]. 2020 Mar 1 [cited 2022 Oct 7];80:103823. Available from: <http://dx.doi.org/10.1016/j.foodqual.2019.103823>
414. Melcher J, Camacho E, Lagan S, Torous J. College student engagement with mental health apps: analysis of barriers to sustained use. *J Am Coll Health* [Internet]. 2022 Aug;70(6):1819–25. Available from: <http://dx.doi.org/10.1080/07448481.2020.1825225>
415. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To Prompt or Not to Prompt? A Microrandomized Trial of Time-Varying Push Notifications to Increase Proximal Engagement With a Mobile Health App. *JMIR mHealth and uHealth* [Internet]. 2018 Nov 29 [cited 2022 Nov 19];6(11):e10123. Available from: <https://mhealth.jmir.org/2018/11/e10123>
416. Freyne J, Yin J, Brindal E, Hendrie GA, Berkovsky S, Noakes M. Push Notifications in Diet Apps: Influencing Engagement Times and Tasks. *International Journal of Human–Computer Interaction* [Internet]. 2017; Available from: <http://dx.doi.org/10.1080/10447318.2017.1278896>

417. Nguyen MH, Bol N, King AJ. Customization versus personalisation of digital health information. *European Journal of Health Communication* [Internet]. 2020 Oct 1;1(1):30–54. Available from: <https://ejhc.org/article/view/ejhc.2020.003>
418. Samdal GB, Eide GE, Barth T, Williams G, Meland E. Effective behaviour change techniques for physical activity and healthy eating in overweight and obese adults; systematic review and meta-regression analyses. *Int J Behav Nutr Phys Act* [Internet]. 2017 Mar 28 [cited 2023 Mar 10];14(1):1–14. Available from: <https://ijbnpa.biomedcentral.com/articles/10.1186/s12966-017-0494-y>
419. Bentley MR, Mitchell N, Sutton L, Backhouse SH. Sports nutritionists’ perspectives on enablers and barriers to nutritional adherence in high performance sport: A qualitative analysis informed by the COM-B model and theoretical domains framework. *J Sports Sci* [Internet]. 2019 Sep;37(18):2075–85. Available from: <http://dx.doi.org/10.1080/02640414.2019.1620989>
420. Liddy C, Johnston S, Irving H, Nash K, Ward N. Improving awareness, accountability, and access through health coaching: qualitative study of patients’ perspectives. *Can Fam Physician* [Internet]. 2015 Mar;61(3):e158–64. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/25932483>
421. Oussedik E, Foy CG, Masicampo EJ, Kammrath LK, Anderson RE, Feldman SR. Accountability: a missing construct in models of adherence behavior and in clinical practice. *Patient Prefer Adherence* [Internet]. 2017 [cited 2022 Nov 21];11:1285. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5536091/>
422. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol Bull* [Internet]. 1999 Nov;125(6):627–68; discussion 692–700. Available from: <http://dx.doi.org/10.1037/0033-2909.125.6.627>
423. Venkatesh A, Chang A, Green EA, Randall T, Gallagher R, Wildes JE, et al. Perceived Facilitators and Barriers to Engaging with a Digital Intervention among Those with Food Insecurity, Binge Eating, and Obesity. *Nutrients* [Internet]. 2021 Jul 19;13(7). Available from: <http://dx.doi.org/10.3390/nu13072458>
424. Ray D, Sniehotta F, McColl E, Ells L. Barriers and facilitators to implementing practices for prevention of childhood obesity in primary care: A mixed methods systematic review. *Obes Rev* [Internet]. 2022 Apr;23(4):e13417. Available from: <http://dx.doi.org/10.1111/obr.13417>
425. Brenton-Peters JM, Vallis M, Grant S, Consedine NS, Kirk SFL, Roy R, et al. Rethinking weight: Finding self-compassion for “weight management.” *Clin Obes* [Internet]. 2023 Feb;13(1):e12562. Available from: <http://dx.doi.org/10.1111/cob.12562>

426. Kebbe M, Perez A, Buchholz A, McHugh TLF, Scott SS, Richard C, et al. Barriers and enablers for adopting lifestyle behavior changes in adolescents with obesity: A multi-centre, qualitative study. *PLoS One* [Internet]. 2018 Dec 18 [cited 2023 Oct 11];13(12):e0209219. Available from: <https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0209219&type=printable>
427. Van Kessel G, Kavanagh M, Maher C. A Qualitative Study to Examine Feasibility and Design of an Online Social Networking Intervention to Increase Physical Activity in Teenage Girls. *PLoS One* [Internet]. 2016 Mar 2 [cited 2023 Oct 11];11(3):e0150817. Available from: <https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0150817&type=printable>
428. Cornet VP, Daley C, Bolchini D, Toscos T, Mirro MJ, Holden RJ. Patient-centered Design Grounded in User and Clinical Realities: Towards Valid Digital Health. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care* [Internet]. 2019 Sep 15 [cited 2023 Dec 6]; Available from: <https://journals.sagepub.com/doi/10.1177/2327857919081023>
429. Analytical Impact Team. Overview of the UK population - Office for National Statistics [Internet]. Office for National Statistics; 2021 [cited 2021 Nov 5]. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/january2021>
430. Population of England and Wales [Internet]. GOV.UK Ethnicity Facts and Figures. 2018 [cited 2021 Nov 5]. Available from: <https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/population-of-england-and-wales/latest>
431. Lipschitz JM, Pike CK, Hogan TP, Murphy SA, Burdick KE. The Engagement Problem: a Review of Engagement with Digital Mental Health Interventions and Recommendations for a Path Forward. *Current Treatment Options in Psychiatry* [Internet]. 2023 Aug 25 [cited 2023 Dec 6];10(3):119–35. Available from: <https://link.springer.com/article/10.1007/s40501-023-00297-3>
432. Wild CEK, Egli V, Rawiri NT, Willing EJ, Hofman PL, Anderson YC. “It’s more personal if you can have that contact with a person”: Qualitative study of health information preferences of parents and caregivers of children with obesity in New Zealand. *Health Soc Care Community* [Internet]. 2022 Sep [cited 2023 Dec 6];30(5). Available from: <https://pubmed.ncbi.nlm.nih.gov/35170827/>
433. Kidd C, Hayden BY. The Psychology and Neuroscience of Curiosity. *Neuron* [Internet]. 2015 Nov 4;88(3):449–60. Available from: <http://dx.doi.org/10.1016/j.neuron.2015.09.010>

434. Murayama K, FitzGibbon L, Sakaki M. Process account of curiosity and interest: A reward-learning perspective. *Educ Psychol Rev* [Internet]. 2019 Dec;31(4):875–95. Available from: <http://link.springer.com/10.1007/s10648-019-09499-9>
435. Pekrun R. The murky distinction between curiosity and interest: State of the art and future prospects. *Educ Psychol Rev* [Internet]. 2019 Dec;31(4):905–14. Available from: <http://link.springer.com/10.1007/s10648-019-09512-1>
436. Lopez A, Schwenk S, Schneck CD, Griffin RJ, Mishkind MC. Technology-Based Mental Health Treatment and the Impact on the Therapeutic Alliance. *Curr Psychiatry Rep* [Internet]. 2019 Jul 8;21(8):76. Available from: <http://dx.doi.org/10.1007/s11920-019-1055-7>
437. Baier AL, Kline AC, Feeny NC. Therapeutic alliance as a mediator of change: A systematic review and evaluation of research. *Clin Psychol Rev* [Internet]. 2020 Dec;82(101921):101921. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0272735820301094>
438. Beatty C, Malik T, Meheli S, Sinha C. Evaluating the Therapeutic Alliance With a Free-Text CBT Conversational Agent (Wysa): A Mixed-Methods Study. *Front Digit Health* [Internet]. 2022 Apr 11;4:847991. Available from: <http://dx.doi.org/10.3389/fdgth.2022.847991>
439. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review. *J Consult Clin Psychol* [Internet]. 2020 Jan;88(1):1–13. Available from: <http://dx.doi.org/10.1037/ccp0000459>
440. Young AT, Amara D, Bhattacharya A, Wei ML. Patient and general public attitudes towards clinical artificial intelligence: a mixed methods systematic review. *The Lancet Digital Health* [Internet]. 2021 Sep 1 [cited 2024 Mar 7];3(9):e599–611. Available from: <http://www.thelancet.com/article/S2589750021001321/abstract>
441. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial Intelligence-Based Chatbots for Promoting Health Behavioral Changes: Systematic Review. *J Med Internet Res* [Internet]. 2023 Feb 24;25:e40789. Available from: <http://dx.doi.org/10.2196/40789>
442. Abd-Alrazaq AA, Alajlani M, Ali N, Denecke K, Bewick BM, Househ M. Perceptions and Opinions of Patients About Mental Health Chatbots: Scoping Review. *J Med Internet Res* [Internet]. 2021 Jan 13;23(1):e17828. Available from: <http://dx.doi.org/10.2196/17828>
443. He Y, Yang L, Qian C, Li T, Su Z, Zhang Q, et al. Conversational Agent Interventions for Mental Health Problems: Systematic Review and Meta-analysis of Randomized Controlled Trials. *J Med Internet Res* [Internet]. 2023 Apr 28 [cited 2024 Mar 29];25(1):e43862. Available from: <https://www.jmir.org/2023/1/e43862>

444. Hamdoun S, Monteleone R, Bookman T, Michael K. AI-Based and Digital Mental Health Apps: Balancing Need and Risk. *IEEE Technol Soc Mag* [Internet]. 2023 Mar;42(1):25–36. Available from: <http://dx.doi.org/10.1109/MTS.2023.3241309>
445. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evid Based Ment Health* [Internet]. 2018 Aug;21(3):116–9. Available from: <http://dx.doi.org/10.1136/eb-2018-102891>
446. Treweek S, Pitkethly M, Cook J, Fraser C, Mitchell E, Sullivan F, et al. Strategies to improve recruitment to randomised trials. *Cochrane Database Syst Rev* [Internet]. 2018 Feb 22;2(2):MR000013. Available from: <http://dx.doi.org/10.1002/14651858.MR000013.pub6>
447. Borschmann R, Patterson S, Poovendran D, Wilson D, Weaver T. Influences on recruitment to randomised controlled trials in mental health settings in England: a national cross-sectional survey of researchers working for the Mental Health Research Network. *BMC Med Res Methodol* [Internet]. 2014 Feb 17 [cited 2023 Aug 29];14(1):1–11. Available from: <https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-14-23>
448. Liu Y, Pencheon E, Hunter RM, Moncrieff J, Freemantle N. Recruitment and retention strategies in mental health trials - A systematic review. *PLoS One* [Internet]. 2018 Aug 29 [cited 2023 Aug 29];13(8). Available from: <https://pubmed.ncbi.nlm.nih.gov/30157250/>
449. Resnik DB. Bioethical Issues in Providing Financial Incentives to Research Participants. *Medicoleg Bioeth* [Internet]. 2015 Jun 24;5:35–41. Available from: <http://dx.doi.org/10.2147/MB.S70416>
450. Hall CL, Sanderson C, Brown BJ, Andr  n P, Bennett S, Chamberlain LR, et al. Opportunities and challenges of delivering digital clinical trials: lessons learned from a randomised controlled trial of an online behavioural intervention for children and young people. *Trials* [Internet]. 2020 Dec 9;21(1):1011. Available from: <http://dx.doi.org/10.1186/s13063-020-04902-1>
451. Pham Q, Wiljer D, Cafazzo JA. Beyond the Randomized Controlled Trial: A Review of Alternatives in mHealth Clinical Trial Methods. *JMIR mHealth and uHealth* [Internet]. 2016 [cited 2023 Dec 14];4(3). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5035379/>
452. Hrynyschyn R, Prediger C, Stock C, Helmer SM. Evaluation Methods Applied to Digital Health Interventions: What Is Being Used beyond Randomised Controlled Trials?—A Scoping Review. *Int J Environ Res Public Health* [Internet]. 2022 May [cited 2023 Dec 14];19(9). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9102232/>

453. Guo C, Ashrafian H, Ghafur S, Fontana G, Gardner C, Prime M. Challenges for the evaluation of digital health solutions—A call for innovative evidence generation approaches. *npj Digital Medicine* [Internet]. 2020 Aug 27 [cited 2023 Dec 14];3(1):1–14. Available from: <https://www.nature.com/articles/s41746-020-00314-2>
454. What is a Good Net Promoter Score? [Internet]. Qualtrics. 2021 [cited 2022 Aug 19]. Available from: <https://www.qualtrics.com/uk/experience-management/customer/good-net-promoter-score/>
455. Wutz M, Hermes M, Winter V, Köberlein-Neu J. Factors Influencing the Acceptability, Acceptance, and Adoption of Conversational Agents in Health Care: Integrative Review. *J Med Internet Res* [Internet]. 2023 [cited 2024 Mar 30];25. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10565637/>
456. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. *npj Digital Medicine* [Internet]. 2018 Mar 14 [cited 2022 Sep 23];1(1):1–4. Available from: <https://www.nature.com/articles/s41746-017-0012-2>
457. D’Amario D, Canonico F, Rodolico D, Borovac JA, Vergallo R, Montone RA, et al. Telemedicine, Artificial Intelligence and Humanisation of Clinical Pathways in Heart Failure Management: Back to the Future and Beyond. *Cardiac Failure Review* [Internet]. 2020 Mar [cited 2022 Sep 23];6. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7312813/>
458. Morris MR. AI and Accessibility [Internet]. *Communications of the ACM*. 2020 [cited 2022 Sep 23]. Available from: <https://cacm.acm.org/magazines/2020/6/245157-ai-and-accessibility/fulltext>
459. Sadang JM, Palompon DR, Agoncillo DJE, Alawiya JTP. A Multiple Case Study of Older Adults’ Internal Resiliency and Quality of Life during the COVID-19 Health Emergency. *Annals of Geriatric Medicine and Research* [Internet]. 2023 Jun [cited 2023 Nov 9];27(2):158. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10326400/>
460. Mergelsberg ELP, Mullan BA, Allom V, Scott A. An intervention designed to investigate habit formation in a novel health behaviour. *Psychol Health* [Internet]. 2021 Apr 3 [cited 2023 Nov 8]; Available from: <https://www.tandfonline.com/doi/abs/10.1080/08870446.2020.1779272>
461. Phillips LA, Mullan BA. Ramifications of behavioural complexity for habit conceptualisation, promotion, and measurement. *Health Psychol Rev* [Internet]. 2023 Jul 3 [cited 2023 Nov 8];17(3):402–15. Available from: <https://www.tandfonline.com/doi/abs/10.1080/17437199.2022.2060849>

462. Keller J, Kwasnicka D, Klaiber P, Sichert L, Lally P, Fleig L. Habit formation following routine-based versus time-based cue planning: A randomized controlled trial. *Br J Health Psychol* [Internet]. 2021 Sep;26(3):807–24. Available from: <http://dx.doi.org/10.1111/bjhp.12504>
463. Lally P, van Jaarsveld CHM, Potts HWW, Wardle J. How are habits formed: Modelling habit formation in the real world. *Eur J Soc Psychol* [Internet]. 2010 Oct 1 [cited 2023 Nov 8];40(6):998–1009. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/ejsp.674>
464. Timulak L. Meta-analysis of qualitative studies: a tool for reviewing qualitative research findings in psychotherapy. *Psychother Res* [Internet]. 2009 Jul;19(4-5):591–600. Available from: <http://dx.doi.org/10.1080/10503300802477989>
465. Egan SJ, Wade TD, Fitzallen G, O'Brien A, Shafran R. A meta-synthesis of qualitative studies of the link between anxiety, depression and perfectionism: implications for treatment. *Behav Cogn Psychother* [Internet]. 2022 Jan;50(1):89–105. Available from: <http://dx.doi.org/10.1017/S1352465821000357>
466. Lachal J, Orri M, Sibeoni J, Moro MR, Revah-Levy A. Metasynthesis of youth suicidal behaviours: perspectives of youth, parents, and health care professionals. *PLoS One* [Internet]. 2015 May 22;10(5):e0127359. Available from: <http://dx.doi.org/10.1371/journal.pone.0127359>
467. Sarwer DB, Polonsky HM. The Psychosocial Burden of Obesity. *Endocrinol Metab Clin North Am* [Internet]. 2016 Sep [cited 2023 Nov 20];45(3):677. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6052856/>
468. Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy* [Internet]. 1979;16(3):252–60. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/h0085885>
469. Eubanks CF, Muran JC, Safran JD. Alliance rupture repair: A meta-analysis. *Psychotherapy* [Internet]. 2018 Dec;55(4):508–19. Available from: <http://dx.doi.org/10.1037/pst0000185>
470. Tong F, Lederman R, D'Alfonso S, Berry K, Bucci S. Digital Therapeutic Alliance With Fully Automated Mental Health Smartphone Apps: A Narrative Review. *Front Psychiatry* [Internet]. 2022 Jun 22 [cited 2023 Nov 20];13:819623. Available from: <http://dx.doi.org/10.3389/fpsy.2022.819623>
471. D'Alfonso S, Lederman R, Bucci S, Berry K. The Digital Therapeutic Alliance and Human-Computer Interaction. *JMIR Mental Health* [Internet]. 2020 Dec 29 [cited 2023 Nov 20];7(12):e21895. Available from: <https://mental.jmir.org/2020/12/e21895>

472. Henson P, Wisniewski H, Hollis C, Keshavan M, Torous J. Digital mental health apps and the therapeutic alliance: initial review. *BJPsych Open* [Internet]. 2019 Jan;5(1):e15. Available from: <http://dx.doi.org/10.1192/bjo.2018.86>
473. Jeronimus BF, Laceulle OM. Frustration. *Encyclopedia of Personality and Individual Differences* [Internet]. 2017 [cited 2024 Jun 3];1–5. Available from: https://link.springer.com/referenceworkentry/10.1007/978-3-319-28099-8_815-1
474. Hertzum M, Hornbæk K. Frustration: Still a common user experience. *ACM Trans Comput Hum Interact* [Internet]. 2023 Jun 30;30(3):1–26. Available from: <https://dl.acm.org/doi/10.1145/3582432>
475. Larche CJ, Musielak N, Dixon MJ. The Candy Crush Sweet Tooth: How “Near-misses” in Candy Crush Increase Frustration, and the Urge to Continue Gameplay. *J Gambl Stud* [Internet]. 2017 Jun;33(2):599–615. Available from: <http://dx.doi.org/10.1007/s10899-016-9633-7>
476. Bijkerk LE, Spigt M, Oenema A, Geschwind N. Engagement with mental health and health behavior change interventions: An integrative review of key concepts. *J Contextual Behav Sci* [Internet]. 2024 Apr;32(100748):100748. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2212144724000280>
477. Noble H, Smith J. Issues of validity and reliability in qualitative research. *Evid Based Nurs* [Internet]. 2015 Apr 1 [cited 2024 Mar 25];18(2):34–5. Available from: <https://ebn.bmj.com/content/18/2/34.abstract>
478. Karlsen MMW, Gabrielsen AK, Falch AL, Stubberud DG. Intensive care nursing students’ perceptions of simulation for learning confirming communication skills: A descriptive qualitative study. *Intensive Crit Care Nurs* [Internet]. 2017 Oct;42:97–104. Available from: <http://dx.doi.org/10.1016/j.iccn.2017.04.005>
479. Vaportzis E, Giatsi Clausen M, Gow AJ. Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study. *Front Psychol* [Internet]. 2017 Oct 4 [cited 2024 May 10];8:294305. Available from: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.01687/pdf>
480. Shaw T, McGregor D, Brunner M, Keep M, Janssen A, Barnet S. What is eHealth (6)? Development of a Conceptual Model for eHealth: Qualitative Study with Key Informants. *J Med Internet Res* [Internet]. 2017 Oct 24 [cited 2024 May 10];19(10):e8106. Available from: <https://www.jmir.org/2017/10/e324>

481. Olmos-Vega FM, Stalmeijer RE, Varpio L, Kahlke R. A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. *Med Teach* [Internet]. 2023 Mar 4 [cited 2024 Jun 10]; Available from: <https://www.tandfonline.com/doi/abs/10.1080/0142159X.2022.2057287>
482. Garnett C, Perski O, Tombor I, West R, Michie S, Brown J. Predictors of Engagement, Response to Follow Up, and Extent of Alcohol Reduction in Users of a Smartphone App (Drink Less): Secondary Analysis of a Factorial Randomized Controlled Trial. *JMIR mHealth and uHealth* [Internet]. 2018 Dec 14 [cited 2023 Dec 11];6(12):e11175. Available from: <https://mhealth.jmir.org/2018/12/e11175>
483. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health Psychol Rev* [Internet]. 2016 Sep;10(3):277–96. Available from: <http://dx.doi.org/10.1080/17437199.2016.1151372>
484. Bandura A. Social Cognitive Theory. In: Van Lange PAM, Kruglanski AW, Higgins ET, editors. *Handbook of Theories of Social Psychology: Volume 1* [Internet]. 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom: SAGE Publications Ltd; 2014. p. 349–74. Available from: https://sk.sagepub.com/reference/hdbk_socialpsychtheories1/n18.xml
485. Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process* [Internet]. 1991 Dec;50(2):179–211. Available from: <https://linkinghub.elsevier.com/retrieve/pii/074959789190020T>
486. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot* [Internet]. 1997 Sep-Oct;12(1):38–48. Available from: <http://dx.doi.org/10.4278/0890-1171-12.1.38>
487. Hagger MS, Cameron LD, Hamilton K, Hankonen N, Lintunen T, Ajzen I, et al. *The Handbook of Behavior Change* [Internet]. Cambridge University Press; 2020 [cited 2024 May 13]. Available from: <https://www.cambridge.org/core/books/handbook-of-behavior-change/BD76D0E53DB1BC2EE8C53C1603D45E6A>
488. Welcome to the Human Behaviour Change Project (HBCP) [Internet]. Human Behaviour Change Project. [cited 2024 Apr 4]. Available from: <https://www.humanbehaviourchange.org/>
489. Lipschitz J, Miller CJ, Hogan TP, Burdick KE, Lippin-Foster R, Simon SR, et al. Adoption of Mobile Apps for Depression and Anxiety: Cross-Sectional Survey Study on Patient Interest and Barriers to Engagement. *JMIR Ment Health* [Internet]. 2019 Jan 25;6(1):e11334. Available from: <http://dx.doi.org/10.2196/11334>

490. Hogan TP, Etingen B, Lipschitz JM, Shimada SL, McMahon N, Bolivar D, et al. Factors Associated With Self-reported Use of Web and Mobile Health Apps Among US Military Veterans: Cross-sectional Survey. *JMIR Mhealth Uhealth* [Internet]. 2022 Dec 30;10(12):e41767. Available from: <http://dx.doi.org/10.2196/41767>
491. Young CL, Mohebbi M, Staudacher HM, Kay-Lambkin F, Berk M, Jacka FN, et al. Optimizing Engagement in an Online Dietary Intervention for Depression (My Food & Mood Version 3.0): Cohort Study. *JMIR Ment Health* [Internet]. 2021 Mar 31;8(3):e24871. Available from: <http://dx.doi.org/10.2196/24871>
492. Schwarz A, Winkens LHH, de Vet E, Ossendrijver D, Bouwsema K, Simons M. Design Features Associated With Engagement in Mobile Health Physical Activity Interventions Among Youth: Systematic Review of Qualitative and Quantitative Studies. *JMIR mHealth and uHealth* [Internet]. 2023 Mar 6 [cited 2024 Apr 5];11(1):e40898. Available from: <https://mhealth.jmir.org/2023/1/e40898>
493. Ahmed RE, Bdair IA, AL-Mugheed K, Alshahrani SH, Alalyani MM, Ramaiah R, et al. Empowering Self-Efficacy by Using Patient Empowerment among Chronic Obstructive Pulmonary Disease: Pre–Post-Test Study. *Healthc Pap* [Internet]. 2023 Feb [cited 2024 Apr 5];11(3). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9914704/>
494. Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychol Rev* [Internet]. 2015 Aug 8 [cited 2022 Jun 6];9(3):323. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4566873/>

Appendices

Appendix 2-A. PRISMA checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Title page
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Title page
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	1
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	1-2
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	2, Table 1
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	2
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	2, Table 2
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	2
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	2-3
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Textbox 1
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	N/A
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	3
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	N/A
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	3
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	N/A
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	N/A
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	3
	13e	Describe any methods used to explore possible causes of heterogeneity among	N/A

Section and Topic	Item #	Checklist item	Location where item is reported
		study results (e.g. subgroup analysis, meta-regression).	
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	N/A
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	N/A
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	N/A
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	3, Figure 1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Figure 1
Study characteristics	17	Cite each included study and present its characteristics.	4, Appendix 5
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	5-6, Tables 6 and 8, Appendix 6
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	4-6, Tables 3-7
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	4-6, Table 8
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	N/A
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	N/A
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	N/A
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N/A
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	N/A
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	6-7
	23b	Discuss any limitations of the evidence included in the review.	6-7
	23c	Discuss any limitations of the review processes used.	6-7
	23d	Discuss implications of the results for practice, policy, and future research.	7-8
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Title page
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	2
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	2
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	26
Competing interests	26	Declare any competing interests of review authors.	26

Section and Topic	Item #	Checklist item	Location where item is reported
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	27

Appendix 2-B. Search results by database (12 July 2022)

Database	Search string	References
PubMed	(((Behavior Control OR Psychological Theory[MeSH Terms]) OR ("behavior change techniques"[Title/Abstract] or "behavior change techniques"[Title/Abstract] or "BCT"[Title/Abstract] or "behavior change technique"[Title/Abstract] or "behavior change technique"[Title/Abstract] or "behavioral change strategies"[Title/Abstract] or "behavioral change strategies"[Title/Abstract] or "behavior change wheel"[Title/Abstract] or "behavior change wheel"[Title/Abstract] or "behavioral theory"[Title/Abstract] or "behavioral theory"[Title/Abstract] or "behavior change theory"[Title/Abstract] or "behavior change theory"[Title/Abstract] or "health behavior change"[Title/Abstract] or "behavior change"[Title/Abstract] or "behavior change"[Title/Abstract] or "digital behavior change intervention"[Title/Abstract] or "digital behavior change intervention"[Title/Abstract] or "DBCI"[Title/Abstract] or "behavior change intervention"[Title/Abstract])) AND ((Telemedicine OR Mobile Applications OR Internet-Based Intervention[MeSH Terms]) OR ("mHealth"[Title/Abstract] OR "eHealth"[Title/Abstract] OR "mobile health"[Title/Abstract] OR "eHealth"[Title/Abstract] OR telehealth[Title/Abstract] OR mobile[Title/Abstract] OR phone[Title/Abstract] OR smartphone[Title/Abstract] OR cell[Title/Abstract] OR digital[Title/Abstract] OR "app"[Title/Abstract] OR "apps"[Title/Abstract] OR application*[Title/Abstract] OR digital[Title/Abstract] OR web[Title/Abstract] OR internet[Title/Abstract] OR online[Title/Abstract] OR technology[Title/Abstract]))) AND ((Treatment Adherence and Compliance OR Patient Participation OR Patient Compliance[MeSH Terms]) OR (Engag*[Title/Abstract] OR "user engagement"[Title/Abstract] OR immersion[Title/Abstract] OR flow[Title/Abstract] OR involvement[Title/Abstract] OR presence[Title/Abstract] OR adherence[Title/Abstract] OR attrition compliance[Title/Abstract] OR maintenance[Title/Abstract] OR acceptability[Title/Abstract] OR satisfaction[Title/Abstract]))	12,397
Embase (Ovid)	((patient compliance/ or patient participation/) OR ((Engag* or "user engagement" or immersion or flow or involvement or presence or adherence or attrition compliance or maintenance or acceptability or satisfaction).ti,ab)) AND ((telemedicine/ or mobile application/ or internet-based intervention/) OR (("mHealth" or "eHealth" or "mobile health" or "eHealth" or telehealth or mobile or phone or smartphone or cell or digital or "app" or "apps" or application* or digital or web or internet).ti,ab)) AND ((behaviour control/ or psychological theory/) OR (("behaviour change techniques" or "behaviour change techniques" or "BCT" or "behaviour change technique" or "behaviour change technique" or "behavioural change strategies" or "behavioural change strategies" or "behaviour change wheel" or "behaviour change wheel" or "behavioural theory" or "behavioural theory" or "behaviour change theory" or "behaviour change theory" or "health behaviour change" or "behaviour change" or "behaviour change" or "digital behaviour	2,565

	change intervention" or "digital behaviour change intervention" or "DBCI" or "behaviour change intervention").ti,ab))	
CINAHL ^a	(MH (Treatment Adherence and Compliance OR Patient Participation OR Patient Compliance) OR AB (Engag* OR “user engagement” OR immersion OR flow OR involvement OR presence OR adherence OR attrition compliance OR maintenance OR acceptability OR satisfaction)) AND (MH (Telemedicine OR Mobile Applications OR Internet-Based Intervention) OR AB (“mHealth” OR “eHealth” OR “mobile health” OR “eHealth” OR telehealth OR mobile OR phone OR smartphone OR cell OR digital OR “app” OR “apps” OR application* OR digital OR web OR internet)) AND (MH (Behavior Control OR Psychological Theory) OR AB (“behaviour change techniques” or “behaviour change techniques” or “BCT” or “behaviour change technique” or “behaviour change technique” or “behavioural change strategies” or “behavioural change strategies” or “behaviour change wheel” or “behaviour change wheel” or “behavioural theory” or “behavioural theory” or “behaviour change theory” or “health behaviour change” or “behaviour change” or “behaviour change” or “digital behaviour change intervention” or “digital behaviour change intervention” or “DBCI” or “behaviour change intervention”))	672
APA PsycArticles ^a	(mainsubject(Treatment Adherence and Compliance OR Patient Participation OR Patient Compliance) OR ab(Engag* OR “user engagement” OR immersion OR flow OR involvement OR presence OR adherence OR attrition compliance OR maintenance OR acceptability OR satisfaction)) AND (mainsubject(Telemedicine OR Mobile Applications OR Internet-Based Intervention) OR ab(“mHealth” OR “eHealth” OR “mobile health” OR “eHealth” OR telehealth OR mobile OR phone OR smartphone OR cell OR digital OR “app” OR “apps” OR application* OR digital OR web OR internet)) AND (mainsubject(Behavior Control OR Psychological Theory) OR ab(“behaviour change techniques” or “behaviour change techniques” or “BCT” or “behaviour change technique” or “behaviour change technique” or “behavioural change strategies” or “behavioural change strategies” or “behaviour change wheel” or “behaviour change wheel” or “behavioural theory” or “behavioural theory” or “behaviour change theory” or “behaviour change theory” or “health behaviour change” or “behaviour change” or “behaviour change” or “digital behaviour change intervention” or “digital behaviour change intervention” or “DBCI” or “behaviour change intervention”))	66
Science Direct ^b	("engagement") AND ("mobile" OR "digital" OR "mHealth") AND ("BCT" OR "behaviour theory" OR "behavioural theory" OR "behaviour change intervention")	1,291
Cochrane Library ^c (Trials)	(Engag* OR “user engagement” OR immersion OR flow OR involvement OR presence OR adherence OR attrition compliance OR maintenance OR acceptability OR satisfaction) AND (mHealth OR eHealth OR “mobile health” OR telehealth OR mobile OR phone OR smartphone OR cell OR digital OR “app” OR “apps” OR application* OR digital OR web OR internet) AND (“behaviour change techniques” or “behaviour change techniques” or “BCT” or “behaviour change	1,779

	technique” or “behaviour change technique” or “behavioural change strategies” or “behavioural change strategies” or “behaviour change wheel” or “behaviour change wheel” or “behavioural theory” or “behavioural theory” or “behaviour change theory” or “behaviour change theory” or “behaviour change theory” or “health behaviour change” or “behaviour change” or “behaviour change” or “digital behaviour change intervention” or “digital behaviour change intervention” or “DBCI” or “behaviour change intervention”)	
Web of Science ^{c,d}	(Engag* OR “user engagement” OR immersion OR flow OR involvement OR presence OR adherence OR attrition compliance OR maintenance OR acceptability OR satisfaction) AND ("mHealth" OR “eHealth” OR “mobile health” OR “eHealth” OR telehealth OR mobile OR phone OR smartphone OR cell OR digital OR “app” OR “apps” OR application* OR digital OR web OR internet) AND (“behaviour change techniques” or “behaviour change techniques” or “BCT” or “behaviour change technique” or “behaviour change technique” or “behavioural change strategies” or “behavioural change strategies” or “behaviour change wheel” or “behaviour change wheel” or “behavioural theory” or “behavioural theory” or “behaviour change theory” or “behaviour change theory” or “health behaviour change” or “behaviour change” or “behaviour change” or “digital behaviour change intervention” or “digital behaviour change intervention” or “DBCI” or “behaviour change intervention”)	2,415
Total		21,185
Total	<i>(with 3245 duplicates removed by EndNote)</i>	17,940

^aCINAHL and APA PsycArticles do not have a Title/Abstract option, so the search was limited to abstracts

^bScience Direct only allows 8 Boolean Operators to be used in a search, so key terms from each theme were selected and used in the search

^cCochrane Library and Web of Science do not use MeSH terms, so only keywords were searched

^dWeb of Science does not have a Title/Abstract search limiter, so ‘Topic’ was used (searches Title, Abstract, author keywords, and Keywords Plus)

Appendix 2-C. EndNote screening

Pass ^a	Search string	# of references remaining	Test articles present?
1	Year = GREATER OR EQUAL TO 2012	13,710	Yes
2	Title = NOT (review OR protocol OR chapter OR congress OR conference OR poster OR abstracts)	9,144	Yes
3 ^b	Any Field = (behaviour change OR behaviour change OR behaviour theory OR behavioural theory OR behaviour theory OR behavioural theory OR social cognitive theory OR planned behavio* OR transtheoretical OR IMB)	4,061	Yes
4	Any Field = (digital OR mobile OR “app” OR “apps” OR “application” OR internet OR web OR phone OR telehealth OR mHealth)	3,114	Yes
5	Any Field = (engag* OR adher* OR immers* OR involve* OR adopt OR maint* OR sustain* OR app us* OR acceptab*)	2,608	Yes
6	Title OR Abstract = (behavio*)	2,413	Yes
7	Abstract = (engage OR adhere OR app us* OR acceptability)	2,018	Yes
8	Abstract = (smartphone OR phone OR mobile OR cell OR mHealth OR iphone OR android OR tablet OR “app” OR “apps”)	1,171	Yes
9	Title = (behav* OR engag* OR adher*)	420	Yes
10	Any Field = NOT (text message OR SMS)	322	Yes

^aEach pass was conducted on the subset of studies retrieved in the previous pass

^bSix articles previously identified as relevant were used throughout the screening process to ensure that search terms were not unintentionally eliminating relevant articles [123,201,202,210,226,232]

^cFour specific theories that were not captured with the general ‘behaviour theory’ terms were included in this pass based on a scoping review which identified theories of behaviour change. These were the four that were cited at least 10 times in the included articles. [494]

Appendix 2-D. Engagement components codebook

- **Affective:** motivation, positive and negative feelings (e.g. enjoyment), sensory pleasure of UX [42,45,51]
- **Cognitive:** challenge, interest, thought, attention [42,45,51]
- **Behavioural:** interactivity, usage (amount, duration, frequency, etc.) [42,45,51]
- **Micro:** engagement with the mobile health app itself [14]
 - **UX:** “User interactions with the DBCI features and the context in which those interactions happen” [50]
 - **BCTs:** “Interactions with behaviour change intervention components/active ingredients specifically designed to influence the behavioural determinants which in turn influence the health behaviours” [50]
 - **Both**
 - **Unspecified**
- **Macro:** “engagement and identification with the wider intervention goals” [14]
 - **Behaviour:** engagement with the targeted health behaviour [50]

Appendix 2-E. Updated search results by database (15 January 2024)

Database	Search string	References
PubMed	(((Behavior Control OR Psychological Theory[MeSH Terms]) OR ("behavior change techniques"[Title/Abstract] or "behavior change techniques"[Title/Abstract] or "BCT"[Title/Abstract] or "behavior change technique"[Title/Abstract] or "behavior change technique"[Title/Abstract] or "behavioral change strategies"[Title/Abstract] or "behavioral change strategies"[Title/Abstract] or "behavior change wheel"[Title/Abstract] or "behavior change wheel"[Title/Abstract] or "behavioral theory"[Title/Abstract] or "behavioral theory"[Title/Abstract] or "behavior change theory"[Title/Abstract] or "behavior change theory"[Title/Abstract] or "health behavior change"[Title/Abstract] or "behavior change"[Title/Abstract] or "behavior change"[Title/Abstract] or "digital behavior change intervention"[Title/Abstract] or "digital behavior change intervention"[Title/Abstract] or "DBCI"[Title/Abstract] or "behavior change intervention"[Title/Abstract])) AND ((Telemedicine OR Mobile Applications OR Internet-Based Intervention[MeSH Terms]) OR ("mHealth"[Title/Abstract] OR "eHealth"[Title/Abstract] OR "mobile health"[Title/Abstract] OR "eHealth"[Title/Abstract] OR telehealth[Title/Abstract] OR mobile[Title/Abstract] OR phone[Title/Abstract] OR smartphone[Title/Abstract] OR cell[Title/Abstract] OR digital[Title/Abstract] OR "app"[Title/Abstract] OR "apps"[Title/Abstract] OR application*[Title/Abstract] OR digital[Title/Abstract] OR web[Title/Abstract] OR internet[Title/Abstract] OR online[Title/Abstract] OR technology[Title/Abstract]))) AND ((Treatment Adherence and Compliance OR Patient Participation OR Patient Compliance[MeSH Terms]) OR (Engag*[Title/Abstract] OR "user engagement"[Title/Abstract] OR immersion[Title/Abstract] OR flow[Title/Abstract] OR involvement[Title/Abstract] OR presence[Title/Abstract] OR adherence[Title/Abstract] OR attrition compliance[Title/Abstract] OR maintenance[Title/Abstract] OR acceptability[Title/Abstract] OR satisfaction[Title/Abstract]))	14,322 (+1,925 since last search)
Embase (Ovid)	((patient compliance/ or patient participation/) OR ((Engag* or "user engagement" or immersion or flow or involvement or presence or adherence or attrition or compliance or maintenance or acceptability or satisfaction).ti,ab)) AND ((telemedicine/ or mobile application/ or internet-based intervention/) OR (("mHealth" or "eHealth" or "mobile health" or telehealth or mobile or phone or smartphone or cell or digital or "app" or "apps" or application* or digital or web or internet).ti,ab)) AND ((behavior control/ or psychological theory/) OR (("behaviour change techniques" or "behavior change techniques" or "BCT" or "behaviour change technique" or "behavior change technique" or "behavioral change strategies" or "behavioural change strategies" or "behaviour change wheel" or "behavior change wheel" or "behavioural theory" or "behavioral theory" or "behaviour change theory" or "behavior change theory" or "health behaviour change" or "behaviour change" or "behavior change" or "digital behaviour change	3,254 (+689 since last search)

	intervention" or "digital behavior change intervention" or "DBCI" or "behaviour change intervention").ti,ab))	
CINAHL ^a	(MH (Treatment Adherence and Compliance OR Patient Participation OR Patient Compliance) OR AB (Engag* OR “user engagement” OR immersion OR flow OR involvement OR presence OR adherence OR attrition compliance OR maintenance OR acceptability OR satisfaction)) AND (MH (Telemedicine OR Mobile Applications OR Internet-Based Intervention) OR AB (“mHealth” OR “eHealth” OR “mobile health” OR “eHealth” OR telehealth OR mobile OR phone OR smartphone OR cell OR digital OR “app” OR “apps” OR application* OR digital OR web OR internet)) AND (MH (Behavior Control OR Psychological Theory) OR AB (“behavior change techniques” or “behavior change techniques” or “BCT” or “behavior change technique” or “behavior change technique” or “behavioral change strategies” or “behavioral change strategies” or “behavior change wheel” or “behavior change wheel” or “behavioral theory” or “behavioral theory” or “behavior change theory” or “behavior change theory” or “health behavior change” or “behavior change” or “behavior change” or “digital behavior change intervention” or “digital behavior change intervention” or “DBCI” or “behavior change intervention”))	755 (+83 since last search)
APA PsycArticles ^a	(mainsubject(Treatment Adherence and Compliance OR Patient Participation OR Patient Compliance) OR ab(Engag* OR “user engagement” OR immersion OR flow OR involvement OR presence OR adherence OR attrition compliance OR maintenance OR acceptability OR satisfaction)) AND (mainsubject(Telemedicine OR Mobile Applications OR Internet-Based Intervention) OR ab(“mHealth” OR “eHealth” OR “mobile health” OR “eHealth” OR telehealth OR mobile OR phone OR smartphone OR cell OR digital OR “app” OR “apps” OR application* OR digital OR web OR internet)) AND (mainsubject(Behavior Control OR Psychological Theory) OR ab(“behavior change techniques” or “behavior change techniques” or “BCT” or “behavior change technique” or “behavior change technique” or “behavioral change strategies” or “behavioral change strategies” or “behavior change wheel” or “behavior change wheel” or “behavioral theory” or “behavioral theory” or “behavior change theory” or “behavior change theory” or “health behavior change” or “behavior change” or “behavior change” or “digital behavior change intervention” or “digital behavior change intervention” or “DBCI” or “behavior change intervention”))	69 (+3 since last search)
Science Direct ^b	("engagement") AND ("mobile" OR "digital" OR "mHealth") AND ("BCT" OR "behavior theory" OR "behavioral theory" OR "behavior change intervention")	1,762 (+471 since last search)
Cochrane Library ^c (Trials)	(Engag* OR “user engagement” OR immersion OR flow OR involvement OR presence OR adherence OR attrition compliance OR maintenance OR acceptability OR satisfaction) AND (mHealth OR eHealth OR “mobile health” OR telehealth OR mobile OR phone OR smartphone OR cell OR digital OR “app” OR “apps” OR application* OR digital OR web OR internet) AND (“behavior change techniques”	2,354 (+575 since last search)

	or “behavior change techniques” or “BCT” or “behavior change technique” or “behavior change technique” or “behavioral change strategies” or “behavioral change strategies” or “behavior change wheel” or “behavior change wheel” or “behavioral theory” or “behavioral theory” or “behavior change theory” or “behavior change theory” or “health behavior change” or “behavior change” or “behavior change” or “digital behavior change intervention” or “digital behavior change intervention” or “DBCI” or “behavior change intervention”)	
Web of Science ^{c,d}	(Engag* OR “user engagement” OR immersion OR flow OR involvement OR presence OR adherence OR attrition compliance OR maintenance OR acceptability OR satisfaction) AND (“mHealth” OR “eHealth” OR “mobile health” OR “eHealth” OR telehealth OR mobile OR phone OR smartphone OR cell OR digital OR “app” OR “apps” OR application* OR digital OR web OR internet) AND (“behavior change techniques” or “behavior change techniques” or “BCT” or “behavior change technique” or “behavior change technique” or “behavioral change strategies” or “behavioral change strategies” or “behavior change wheel” or “behavior change wheel” or “behavioral theory” or “behavioral theory” or “behavior change theory” or “behavior change theory” or “health behavior change” or “behavior change” or “behavior change” or “digital behavior change intervention” or “digital behavior change intervention” or “DBCI” or “behavior change intervention”)	2,990 (+575 since last search)
Total		25,506
Total	(increase since previous search on 12 July 2022)	4,321
Total	(with 4,159 duplicates removed by EndNote)	21,346

^aCINAHL and APA PsycArticles do not have a Title/Abstract option, so the search was limited to abstracts

^bScience Direct only allows 8 Boolean Operators to be used in a search, so key terms from each theme were selected and used in the search

^cCochrane Library and Web of Science do not use MeSH terms, so only keywords were searched

^dWeb of Science does not have a Title/Abstract search limiter, so ‘Topic’ was used (searches Title, Abstract, author keywords, and Keywords Plus)

Appendix 2-F. Revised EndNote screening for systematic review update

Pass ^a	Search string	# of references remaining (<i>update</i>)	Previously included articles present? ^b
1	Year = GREATER OR EQUAL TO 2012	17,247	Yes
2	Title = NOT (review OR protocol OR chapter OR congress OR conference OR poster OR abstracts)	14,938	Yes
3 ^{c,d}	Any Field = (theor OR behaviour change OR behavior change OR BCT OR model OR social cognitive theory OR planned behavior OR transtheoretical OR IMB)	9,795	Yes
4	Any Field = (digital OR mobile OR app OR internet OR web OR phone OR telehealth OR mHealth OR tablet)	7,237	Yes
5	Any Field = (engag OR adher OR immers OR involve OR adopt OR maint OR sustain OR app us OR acceptab)	5,076	Yes
6 ^c	Title OR Abstract OR Keywords = (behav)	4,797	Yes
7 ^c	Title OR Abstract OR Keywords = (engag OR adher OR app us* OR maint OR acceptability)	4,316	Yes
8	Any Field = NOT (text message OR SMS)	3,646	Yes
9 ^c	Title OR Abstract OR Keywords = (phone OR mobile OR digital OR mHealth OR app OR tablet)	3,193	Yes
10	Title = (behav OR engag OR adher OR mobile OR app OR mHealth)	1,900	Yes

- ^aEach pass was conducted on the subset of studies retrieved in the previous pass
- ^bSix articles previously identified as relevant were used throughout the screening process to ensure that search terms were not unintentionally eliminating relevant articles [123,201,202,210,226,232]
- ^cArticles without an abstract stored in EndNote were included automatically in this pass to avoid their exclusion
- ^dFour specific theories that were not captured with the general ‘behaviour theory’ terms were included in this pass based on a scoping review which identified theories of behaviour change. These were the four that were cited at least 10 times in the included articles. [494]

Appendix 2-G. Study characteristics

First author (year)	Study type (general)	Study type (specific)	Country	Sample demographics / inclusion criteria	Analysed sample size	Study duration
<i>Originally included studies</i>						
Ahonkhai (2021) [219]	Development	Development	Nigeria	Adolescent and young adult-HIV patients living in LMICs, targeting those with poor adherence to antiretroviral medication	N/A	N/A
Baumel (2018) [197]	App content analysis	Content analysis	N/A	Apps that were self-guided and cost-free	52 apps, 32 websites	18 months
Buman (2016) [220]	Randomised study (incl. RCT)	Multiphase optimization strategy (MOST) research design (full-factorial randomisation $2 \times 2 \times 2$)	US	US Veterans (35-60 years) who were overweight/obese ($BMI \geq 25 \text{ kg/m}^2$) with a fasting glucose of $\geq 100 \text{ mg/dL}$	26	8 weeks
Caon (2022) [221]	Quantitative descriptive	Descriptive study	England, Scotland, Italy, Spain	Adolescents (13-16 years)	357	6 months
Cueto (2019) [211]	Quantitative non-randomised	Retrospective Cohort Study	US	Children and adolescents (5-18 years) who were overweight or obese ($BMI \geq 85^{\text{th}}$ percentile)	1120	1 of 3 commitment periods: 4 weeks, 12 to 16 weeks, or 24 weeks
Curtis (2019) [222]	Qualitative	Qualitative Study	Germany	Children and adolescents (12-18 years) with sickle cell disease	10	N/A
D'Addario (2020) [208]	Qualitative	Focus group study	Italy	Adults (30-50 years) with insufficient physical activity	13	N/A
Davis (2019) [209]	Randomised study (incl. RCT)	Quasi-experimental (randomised)	US	University students (18-29 years)	83	N/A
Delmas (2020) [196]	Quantitative non-randomised	Regression analysis	US	Real world users	2740	83 weeks
DeSmet (2019) [210]	Quantitative descriptive	Cross-Sectional Survey Study	Belgium	Working-age adults (22-55 years) living in Ghent	86	N/A
Du (2016) [223]	Randomised study (incl. RCT)	2x2 Factorial Design Feasibility Study	US	Adults (18+ years) who were physically capable of performing the exercises in this study	89	8 weeks

Duff (2018) [195]	Qualitative	Qualitative usability study	Ireland	Focus groups: Adult HeartSmart participants Feasibility testing: Adults (18+ years) with clinically manifested but stable cardiovascular disease	Focus groups: 26 Feasibility testing: 17	2 weeks
Garnett (2015) [182]	Mixed-methods	Delphi study	UK and Netherlands	International academic experts with backgrounds in health, psychology, and addiction	7	N/A
Hochsmann (2019) [224]	Randomised study (incl. RCT)	Randomized Controlled Trial	Basel metropolitan area	Overweight adults (45-70 years of age) with type 2 diabetes	35	24 weeks
Hales (2016) [225]	Mixed-methods	Mixed methods pilot study	US	Overweight or obese adults (BMI 25-49.9 kg/m ²)	9	2 months
Lin (2021) [226]	Quantitative descriptive	Profile analysis	Canada	Teens (13-17 years) and their parents (with no dietary or physical activity restrictions)	301	4.5 months
Nuijten (2021) [227]	Randomised study (incl. RCT)	Crossover Experimental Study	Netherlands	Preadolescents (11-13 years)	310 - 99% (288/290) of students and 95.5% (22/23) of teachers	12 weeks
Nuijten (2022) [228]	Randomised study (incl. RCT)	RCT	Belgium	Staff members of 7 governmental organizations	111	2 months
Nurmi (2020) [229]	Qualitative	Feasibility (think aloud and SSI) study	Finland	Adults (25-63 years)	10	N/A
Smith (2022) [123]	Qualitative	Development (focus groups)	UK	Women diagnosed with gestational diabetes at Oxford University Hospitals NHS Foundation Trust	10	N/A
Songtaweasin (2020) [230]	Qualitative	Qualitative study	Thailand	Young men who have sex with men (18-21 years)	23 (focus group participants) + 15 key informant interviews	N/A
Sporrel (2020) [231]	Development	Development study	Netherlands	N/A	N/A	N/A
Szinay (2021) [201]	Qualitative	Qualitative Interview Study	UK	Adults (18+ years)	17	N/A
Tang (2015) [202]	Qualitative	Qualitative Interview Study	UK	English-speaking adults (18-40 years) living in Exeter and Plymouth	19	3 weeks
Thornton (2021) [207]	Mixed-methods	Development and usability study	Australia	Secondary school students (grades 7-9) participating in	232 (downloaded app)	N/A

				the Health4Life trial		
Valle (2020) [212]	Quantitative non-randomised	Micro-randomized trial	US	Young adults (18-35 years) who were overweight/obese (BMI=25–40 kg/m ²) with fewer than 150 min per week of moderate-to-vigorous physical activity	52	12 weeks
Whiteley (2018) [232]	Mixed-methods	Qualitative development and usability study	US	Adolescents and young adults (14-26 years) with HIV receiving antiretroviral therapy	20	N/A
Whiteley (2019) [233]	Mixed-methods	Qualitative development and usability study	US	Young men who have sex with men (18-35 years)	20	N/A
<i>Additional studies included after the update</i>						
Abdul Khalil (2023) [234]	Qualitative	Semi-structured interview study	Malaysia	Healthy adults with experience using a diet-related mobile app	20	N/A
Attwood (2017) [200]	Mixed-methods	Mixed-methods evaluation	UK	App was downloaded more by women (59%) than men and by 35-44 year olds (31%) than other ages	119,713 (users downloading app in period)	13 months
Beck (2021) [235]	Development	Development with focus groups	Australia	Adult recovering addicts	N/A	N/A
Bell (2020) [198]	Quantitative descriptive	Longitudinal observational study	UK	Adults who drink excessively (AUDIT score ≥ 8)	19,233 existing users	615 days
Bell (2023) [236]	Randomised study (incl. RCT)	Micro-randomised trial	UK	Adults who drink in excess (AUDIT score 8 or above); mean age 44 years, 44% male, 46% female	566 (349)	30 days
Brindal (2019) [237]	Randomised study (incl. RCT)	RCT	Australia	Adults (20-67 years) with BMI between (20-61kg/m ²), 75% female	61 (36 MotiMate, 25 control)	24 weeks
Burgermaster (2023) [238]	Qualitative	Sequential multiple assignment randomised trial	US	Adults (18-65 years) with Type 2 Diabetes	35	N/A
Chang (2023) [239]	Randomised study (incl. RCT)	RCT	US	Young adults (18-29 years)	85 (43 treatment, 42 control)	4 weeks
Curtis (2015) [240]	Development	Development with focus groups	UK	Parents with overweight and healthy weight children aged 5-11 years	46	N/A

Daniels (2023) [241]	Development	Development with co-design, interviews, and prototype testing	Belgium	Older adults (≥ 65 years) able to be physically active	Empathy and defining phase: In-depth interviews (n=22) Ideate and prototyping phase: Co-creative workshops (n=16, 6) Initial prototype testing of the alpha version 65	N/A
Direito (2015) [203]	Randomised study (incl. RCT)	3-arm parallel RCT	New Zealand	Insufficiently active healthy young people (14-17 years); mean age 15.7, 57% female	51 (immersive app intervention (n=17), nonimmersive app intervention (n=16), or the control group (n=18))	8 weeks
Domin (2022) [213]	Qualitative	Focus group study	Luxembourg	Adolescents (16-18), 50% female	8	N/A
Edney (2020) [242]	Randomised study (incl. RCT)	3-group cluster RCT	Australia	Adults (18-65 years) completing <150 minutes of MVPA per week; mean age 41 years, 74% female	Same as initial	100 days
Edwards (2018) [243]	Development	Development with focus groups	UK	Teen and adult smokers (15-67 years), 53% female	73	N/A
Ehlers (2014) [244]	Quantitative descriptive	Survey study	US	Women (30-64 years)	120	N/A
Ekezie (2021) [245]	Qualitative	Interview and focus group study	UK	Women with previous gestational diabetes	23	Not specified
Fatima (2023) [246]	Quantitative descriptive	Development with pilot testing	UK	Variety of demographics over several pilots, including children, teens, and adults	Pilot 1: 25 Pilot 2: 29 Pilot 3: ? Pilot 4: 6 Pilot 5: 12 Pilot 6: 17	Pilot 1: 2 weeks Pilot 2: ? Pilot 3: ? Pilot 4: 5 weeks Pilot 5: 2 weeks Pilot 6: 70 days
Flaherty (2021) [205]	Qualitative	Observational, phenomenological exploration	Ireland	Women (30-45 years)	6 (midway point) 10 (endpoint)	8 - 11 weeks
Garnett (2019) [247]	Development	Development	UK	Adults who drink excessively	N/A	N/A
Hendrie (2019) [248]	Development	Development with pilot testing	Australia	Adults (20-90 years) who had previously participated /	consumer survey (n = 1068), usability	7 days (pilot)

				expressed interest in online nutrition surveys; mean age 56 years, 84% female	testing session (n = 11), pilot effectiveness and usability trial (n = 283)	
Hendrie (2020) [194]	Quantitative descriptive	Community cohort study	Australia	Adults (31-70 years); mean age 48 years, 84% female	1224 completed the 21-day survey, 273 completed the 90-day survey	90 days
Kekkonen (2023) [249]	Qualitative	Qualitative study within RCT	Finland	Adult microentrepreneurs	59	8 weeks
Lawrason (2023) [250]	Randomised study (incl. RCT)	2-arm repeated-measures RCT	Canada and US	Adults with spinal cord injury who can walk	20	8 week
Lim (2022) [251]	Quantitative descriptive	Prospective cohort study	Singapore	Adults with diabetes or prediabetes; median age 52 years	160	6 months
Litterbach (2017) [252]	Qualitative	Qualitative study	Australia	Mothers of infants	24	Not specified
Lubbans (2014) [253]	Mixed-methods	Subset analysis of intervention group of a cluster RCT	Australia	Adolescent boys in Grade 7; mean age 13 years	361 (42 analysed in focus groups in this paper)	20 weeks
McClure (2023) [254]	Randomised study (incl. RCT)	2-arm pilot parallel RCT	United States	Adults who smoke 10+ cigarettes a day and were ambivalent about quitting	57	3 months
Montilva-Monsalve (2023) [255]	Qualitative	Interview study	UK	Adults with recently diagnosed (<12 mo) and long-standing T2D (>15 mo)	8 patients with recently diagnosed T2D 15 patients with long-standing T2D	N/A
Mummah (2016) [214]	Development	Development with interviews	US	Adults (18-50 years), BMI 28-40kg/m ²	(qualitative interviews, n = 18) (qualitative interviews, n = 14; questionnaire, n = 41)	N/A
Mummah (2016) [215]	Randomised study (incl. RCT)	Pilot RCT	US	Overweight adults; mean age 42 years, mean BMI 32kg/m ²	17	12 weeks
Nahum-Shani (2021) [256]	Randomised study (incl. RCT)	Proof-of-concept micro-RCT	US	Adolescents and young adults (14-24 years); mean age 21 years, 54% female	68 (MRT analysis), 59 (follow-up assessment)	30 days
Nicholas (2021) [257]	Mixed-methods	Development and mixed-methods feasibility study (subset of cluster RCT)	Australia	Insufficiently physically active employees; mean age 44 years, mean BMI 29.02kg/m ²	walkers (n = 34) and Peer Leaders (n = 7)	16 weeks
Patterson (2024) [258]	Qualitative	Qualitative analysis	Australia	Adults; mean age 59, mostly male	15	6 months

Perski (2017) [204]	Qualitative	Think aloud and interview study	UK	Adults that smoke cigarettes daily or drink excessively (AUDIT score ≥ 5)	20	N/A
Pfaffli Dale (2019) [259]	Quantitative descriptive	Survey study	Canada	Teens and adults; mean age 36 years, 65% female	30,538	9 months
Resnick (2021) [260]	Mixed-methods	Mixed-methods pilot study	US	Primary care patients; mean age 51 years, 76% female	41	8 week
Roberts (2019) [206]	Qualitative	Qualitative study	UK	Adult Breast, Prostate, and Colorectal Cancer Survivors	32	2 weeks
Rohde (2019) [261]	Development	Development	Germany	Adolescents and young adults (14-25 years); mean age 18 years, 48% female	210	N/A
Sakane (2023) [216]	Randomised study (incl. RCT)	Pilot RCT	Japan	Overweight and obese men (40-69 years)	74	12 week
Sanders (2023) [262]	Quantitative descriptive	Intervention mapping and user testing study	UK	Inactive adults; mean age 53 years, 53% female, mean BMI 27kg/m ²	15	28 days
Schulte (2022) [217]	Randomised study (incl. RCT)	2-arm parallel RCT	Netherlands	Young adults (18-30 years) wanting to reduce alcohol consumption	ITT 503, per-protocol (int group) 128	6 weeks
Simons (2018) [263]	Mixed-methods	Development, usability, acceptability, and feasibility study	Australia	Lower-educated working young adults	11 (usability) 16 (acceptability/feasibility)	
Stancu (2022) [264]	Quantitative descriptive	Discrete choice experiments (full-factorial 2x2 design)	Germany and Spain	Adults (18-81 years), mean age 47 years	719 respondents from Germany (n = 360) and Spain (n = 359)	N/A
Tombor (2019) [265]	Randomised study (incl. RCT)	Randomised full-factorial experiment (2x2x2x2x2)	UK	Pregnant smokers (18+) who were interested in quitting and set a quit date	494	4 weeks?
Whale (2021) [218]	Development	Development	UK	Children and young people (10-19) with urinary incontinence or urgency and clinicians	23 young people, 8 clinicians	4-6 weeks
Willms (2023) [199]	Development	Development and prototype testing	Canada	Adults; 67% female	6	N/A
Zimmerman (2023) [266]	Quantitative descriptive	Observational study	US	Veterans Health Administration (VHA) primary care patients	56	3 months

Appendix 2-H. Analysis of potential associations between BCTs and engagement and the evidence for the association

First author (year)	BCTs inferred or explicit in paper?	Included BCTs with potential associations with engagement	Evidence of association	Evidence and analysis - key points	RoB/QA ^a
Delmas (2020) [196]	Inferred	7.1 Prompts/cues	Evidenced in study	Notifications were a significant predictor of checking the app ($\beta = 0.890$, $SE = 0.026$, $p < 0.01$) and sharing air quality information with others ($\beta = 0.680$, $SE = 0.086$, $p < 0.01$)	****
Du (2016) [223]	Inferred	3.1 Social support (unspecified)	Evidenced in study	- Comparison Team and Solo mobile app conditions found Team condition had significantly higher compliance ($M = 0.49$, $SD = 0.35$) than Solo condition ($M = 0.30$, $SD = 0.39$; $t 50.82=1.94$, $P=.05$) - Team condition participants were 66% more likely to engage longer with the app than Solo condition participants	Some concerns (RoB)
Garnett (2015) [182]	Explicit in paper	1.1 Goal setting (behaviour) 1.4 Action planning 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour	Evidenced in study	- Expert consensus process identified top 4 BCTs most likely to reduce alcohol consumption - Highest ranked engagement strategies were ease of use, design aesthetic, and feedback	****
Lin (2021) [226]	Explicit in paper	1.1 Goal setting (behaviour) 1.4 Action planning ^b 3.1 Social support (unspecified) 5.3 Information about social and environmental consequences ^b 6.1 Demonstration of the behaviour 7.1 Prompts/cues ^b 8.1 behavioural practice/rehearsal ^b 8.3 Habit formation ^b	Evidenced in study	- Factor modelling compared engagement profiles (uninvolved, dabblers, engaged, and keeners) - App features with the highest engagement across the different profiles were aims (BCT = goal setting), social wall (BCT = social support and demonstration of behaviours), and virtual coach (BCT = social support)	****
Nuijten (2021) [227]	Inferred	3.1 Social support (unspecified)	Evidenced in study	- Compared intergroup (between class competition, between class competition with teachers separate) and intragroup conditions (within class competition) - Significantly more app visits by students in intergroup condition (between classes) than students in intragroup competition (M	High RoB

				diff = +0.469 days, $p < 0.001$) Students in intergroup condition (between classes) also completed more activities than other two conditions, but this was not significant	
Nuijten (2022) [228]	Inferred	1.1 Goal setting (behaviour)	Evidenced in study	- Participants who set a physical activity goal visited the app significantly more than participants who did not (M diff = +2.176 days for maintenance goal, $P < .001$; +1.625 days for improvement goal, $P = .005$) - Participants who set a goal also registered significantly more activities than those who did not (M diff = +1.535 activities for maintenance goal, $P = .03$; +3.258 activities for improvement goal, $P < .001$)	Some concerns (RoB)
Szinay (2021) [201]	Inferred	1.1 Goal setting (behaviour) 1.4 Action planning 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) 3.2 Social support (practical) 5.1 Information about health consequences 7.1 Prompts/cues 8.3 Habit formation 10.3 Non-specific reward	Evidenced in study	Factors identified in interviews as most important for participants' engagement included: knowledge (e.g. user guidance and statistical information), memory, attention, and decision processes (e.g. reduced cognitive load), environmental resources (e.g. tailored technology) and social influences (e.g. peer and professional support)	*****
Tang (2015) [202]	Explicit in paper	1.1 Goal setting (behaviour) 1.4 Action planning 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) 3.2 Social support (practical) 5.5 Anticipated regret 5.6 Information about emotional consequences 7.1 Prompts/cues 10.3 Non-specific reward 10.9 Self-reward	Evidenced in study	Interviews identified several BCTs that were related to engagement by at least 2 participants, including: goal setting, if-then planning, self-monitoring, feedback, awareness of emotional consequences, symbolic rewards, and social contact/support	*****
<i>Additional studies included after the update</i>					
Bell (2020) [198]	Explicit in paper	7.1 Prompts / cues	Evidenced in study	Strong association observed because the receipt of a notification and use of the app within the next hour	*****

Bell (2023) [236]	Explicit in paper	7.1 Prompts / cues	Evidenced in study	Compared to not receiving a notification, participants were 3.5x more likely to open the app within an hour of receiving a notification	Low RoB
Brindal (2019) [237]	Inferred	Control app 2.3 <i>Self-monitoring of behaviour^b</i> 2.4 <i>Self-monitoring of outcomes of behaviour^b</i> MotiMate 1.4 Action planning 2.2 Feedback on behaviour 2.3 <i>Self-monitoring of behaviour^b</i> 2.4 <i>Self-monitoring of outcomes of behaviour^b</i> 3.3 Social support (emotional) 7.1 Prompts / cues	Evidenced in study	- Users of the intervention app were engaged for significantly longer than those of the control app - Users of the intervention app were significantly more likely to recommend the app to a friend compared to the control app ($\chi^2_{22}=10.7$, $P<.01$)	Some concerns (RoB)
Chang (2023) [239]	Inferred	10.1 <i>Material reward^b</i>	Evidenced in study	No evidence that financial incentives can independently increase engagement with an app	Low RoB
Direito (2015) [203]	Inferred	2.2 Feedback on behaviour 2.3 Self-monitoring 2.5 Monitoring of outcome of behaviour 2.7 Feedback on outcome of behaviour 3.1 Social support (unspecified) 4.1 Instruction on how to perform behaviour 6.2 <i>Social comparison^b</i>	Evidenced in study	- Most used and liked features were related to self-monitoring and feedback - Least used and liked features were social networking related	Low RoB
Domin (2022) [213]	Inferred	1.1 Goal setting 1.4 Action planning 2.2. Feedback on behaviour 2.1 Monitoring of behaviour by others (activity tracking) 2.4 <i>Self-monitoring of outcomes of behaviour (mood and sleep tracking)^b</i> 2.5 <i>Monitoring of outcomes of behaviour^b</i> 3.1 <i>Social support (unspecified)^b</i> 4.1 <i>Instruction on how to perform the behaviour (coach)^c</i> 6.2 <i>Social comparison^c</i> 7.1 <i>Prompts/cues (reminders)^b</i> 10.3 <i>Non-specific reward^c</i>	Evidenced in study	- Different feature preferences for users with low physical activity (PA) vs those with higher levels of PA: low PA group liked social support and comparison but high PA group did not - Features liked and perceived as useful by focus groups generally included: goal setting and action planning, instruction on how to perform the behaviour, personalisation, monitoring of behaviour, and feedback - Disliked features included: social networking, avatar coach, and rewards - Features with mixed preferences included: monitoring of outcomes of behaviour and reminders	*****

Edney (2020) [242]	Inferred	2.1 Monitoring of behaviour by others 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) 6.2 Social comparison 7.1 Prompts/cues ^b 10.3 Non-specific reward ^b	Evidenced in study	- Participant feedback indicated that the self-monitoring and leaderboard features were motivating, but were not as interesting in the rewards, mini challenges, and newsfeed - Gamified and non-gamified app users engaged for over two-thirds of possible days (M = 72 and 67 / 100)	Low RoB
Ehlers (2014) [244]	Inferred	1.1 Goal setting 1.2 Problem solving 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) 4.1 Instruction on how to perform behaviour 6.2 Social comparison 7.1 Prompts/cues ^b 10.3 Non-specific rewards	Evidenced in study	- Participant feedback indicated that self-monitoring, goal setting, and feedback were their favourite and most useful features - Participants were less interested in features for motivation, self-efficacy, social support, problem solving, listing benefits, and providing reminders - Social support and comparison, rewards, and information and suggestions were suggested to make app engagement more interesting or enjoyable	***
Ekezie (2021) [245]	Inferred	1.1 Goal setting 2.1 Monitoring of behaviour by others 3.1 Social support (unspecified) 6.2 Social comparison 7.1 Prompts/cues 10.3 Non-specific reward	Evidenced in study	- Social support and comparison had mixed feedback (important for some participants but not all) - Monitoring and achieving behavioural goals and being rewarded were perceived as enjoyable	*****
Fatima (2023) [246]	Explicit in paper	1.1 Goal setting 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) 5.1 Information about health consequences ^b 6.2 Social comparison 8.3 Habit formation 10.2 Material reward 10.3 Non-specific reward ^b	Evidenced in study	- App use and feedback identified features supporting habit formation, self-monitoring, feedback, and rewards as preferred across age groups, although there were differences in preferences by age - Teenagers particularly liked avatars, social comparison and competition, goals, monitoring, and rewards	***
Flaherty (2021) [205]	Explicit in paper	App A 1.3 Goal setting (outcome) 2.4 Self-monitoring of outcomes ^b 3.1 Social support (unspecified) ^b 7.1 Prompts/cues ^b 8.2 Behavioural substitution ^b 11.3 Conserving mental resources ^b App B 1.3 Goal setting (outcome) 2.4 Self-monitoring of outcomes ^b 3.1 Social support (unspecified) ^b 4.2 Information about antecedents ^b 6.3 Information about others approval ^b 7.1 Prompts/cues ^b	Evidenced in study	- Goal setting was associated with engagement - Negative emotions associated with prompts highlighting failure to achieve goals could decrease or increase motivation to engage, depending on the user - App A provided less feedback than App B which was associated with less engagement	*****

		8.2 Behavioural substitution ^b 9.3 Comparative imagining of future outcomes ^b 10.3 Non-specific reward ^b 10.6 Non-specific incentive ^b 12.1. Restructuring the physical environment ^b 12.2. Restructuring the social environment ^b 12.3. Avoidance/reducing exposure to cues for the behavior ^b 12.4. Distraction ^b			
Lim (2022) [251]	Inferred	1.1 Goal setting 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 2.4 Self-monitoring of outcomes of behaviour 2.7 Feedback on outcomes of behaviour 3.1 Social support (unspecified) 3.2 Social support (practical) 7.1 Prompts/cues	Evidenced in study	- High app engagement reported over intervention period (median engagement > 90%) - Features with greatest engagement: "step tracking (95.6%), meal logging (76.6%), and communication with the dietitian within the app's chat system (50%)" - Features with least engagement: "[random blood glucose] RBG monitoring (18%), fasting blood glucose monitoring (19%), and weight charting (26%)"; however, users were only instructed to report these twice weekly	****
Litterbach (2017) [252]	Inferred	1.2 Problem solving ^b 3.1 Social support (unspecified) ^b 3.2 Social support (practical) ^b 4.1 Instruction on how to perform behaviour ^b 4.2 Information about antecedents ^b 5.1 Information about health consequences ^b 6.1 Demonstration of the behaviour ^b 7.1 Prompts/cues 9.1 Credible source ^d 10.6 Non-specific incentive ^b	Evidenced in study	Qualitative feedback indicated app credibility and notifications were important for app use	*****
Lubans (2014) [253]	Explicit in paper	1.1 Goal setting 1.4 Action planning 2.3 Self-monitoring of behaviors 3.1 Social support (unspecified) ^b 5.1 Information about health consequences ^b 7.1 Prompts/cues 8.1 Behavioural practice / rehearsal ^b 8.7 Graded tasks ^b	Evidenced in study	- Goal setting was reported as most used feature (70%), followed by self-monitoring (62% training technique, 49% steps and fitness challenges) ~50% agreed that notifications helped support engagement with behaviour, although feedback also found this feature "a nuisance"	***
Montilva-Monsalve (2023) [255]	Explicit in paper	1.1 Goal setting ^b 1.3. Goal setting (outcome) ^b 1.4. Action planning ^b 1.5. Review behaviour goal(s) ^b 1.6. Discrepancy between current behaviour and goal ^b 2.2. Feedback on behaviour 2.3. Self-monitoring of behaviour 2.4. Self-monitoring of outcome(s)	Evidenced in study	Qualitative evaluation identified BCTs that should be included (all but 8.7 Graded Tasks were already included)	*****

		<i>of behaviour^b</i> 2.5. <i>Monitoring of outcome(s) of behaviour without feedback^b</i> 2.6. <i>Biofeedback^b</i> 2.7. <i>Feedback on outcome(s) of behaviour</i> 3.1. <i>Social support (unspecified)^b</i> 3.2. <i>Social support (practical)^b</i> 4.1. <i>Instruction on how to perform the behaviour</i> 4.2. <i>Information about Antecedents^b</i> 5.1. <i>Information about health consequences^b</i> 6.1. <i>Demonstration of the behaviour</i> 6.3. <i>Information about others' approval^b</i> 7.1. <i>Prompts/cues</i> 7.3. <i>Reduce prompts/cues^b</i> 8.2. <i>Behaviour substitution^b</i> 8.3. <i>Habit formation^b</i> 8.6. <i>Generalisation of target behaviour^b</i> 8.7 <i>Graded tasks^d</i> 9.1. <i>Credible source^b</i> 11.1. <i>Pharmacological support^b</i> 11.2. <i>Reduce negative emotions^b</i> 11.3. <i>Conserving mental resources</i> 12.1. <i>Restructuring the physical environment</i> 12.2. <i>Restructuring the social environment^b</i> 12.3. <i>Avoidance/reducing exposure to cues for the behaviour</i> 13.1. <i>Identification of self as role model^b</i> 13.2. <i>Framing/reframing^b</i> 15.1. <i>Verbal persuasion about capability</i>			
Pfaeffli Dale (2019) [259]	Inferred	7.1 <i>Prompts/cues^b</i> 10.1 <i>Material incentive</i> 10.2 <i>Material reward</i>	Evidenced in study	- Authors concluded that incentives supported engagement (quiz completion rate of 38% was considered by authors to be high relative to other digital health interventions) - Notifications did not appear to support behavioural engagement	*****
Resnick (2021) [260]	Inferred	1.1 <i>Goal setting</i> 1.4 <i>Action planning</i> 1.5 <i>Review behavioural goals</i> 2.2 <i>Feedback on behaviour</i> 2.3 <i>Self-monitoring of behaviour</i> 3.1 <i>Social support (unspecified)</i> 6.2 <i>Social comparison</i> 7.1 <i>Prompts/cues</i> 10.4 <i>Social rewards</i>	Evidenced in study	- App check-ins had decent engagement (M=5.7 out of a possible 8) 83% of users shared progress on app with friends (vs. 44% sharing with all users) - Almost two thirds of users (64%) engaged with app beyond weekly check ins - Participant feedback indicated desire for additional feedback and facilitated social connection	***

				Authors conclude that "reminders, social comparison, social support, goal setting, personalised feedback" could support app and behavioural engagement	
Roberts (2019) [206]	Explicit in paper	Human 1.1 Goal setting (behavior) 2.2 Feedback on behavior 2.3 Self-monitoring of behavior 6.2 <i>Social comparison^b</i> 7.1 Prompts/cues 10.3 <i>Nonspecific reward^b</i> The Walk 2.2 Feedback on behavior 10.3 <i>Nonspecific reward^c</i> 10.6 <i>Nonspecific incentive^c</i> The Johnson & Johnson Official 7 Minute Workout 1.4 Action planning 2.3 Self-monitoring of behavior 4.1 Instruction on how to perform behavior 6.1 Demonstration of the behavior 7.1 Prompts/cues 8.7 <i>Graded tasks^b</i> 9.1 <i>Credible source^b</i> Gorilla Workout 4.1 Instruction on how to perform behavior 6.1 Demonstration of the behavior 7.1 Prompts/cues 8.7 <i>Graded tasks^b</i>	Evidenced in study	- Qualitative feedback indicated that features including video demonstrations and goal setting, self-monitoring, and feedback were liked by app users - Participants provided mixed feedback on prompts/cues depending on tone, timing, and relevance - Mixed, leaning negative, feedback about incentives, rewards, and gamification (depending on implementation in specific apps)	*****
Sakane (2023) [216]	Inferred	1.1 Goal setting behaviour 1.3 Goal setting outcome 2.2 Feedback on behaviour 2.3 Self-monitoring 2.4 Self-monitoring of outcomes of behaviour 2.7 Feedback on outcomes of behaviour 5.1 Information about health consequences	Evidenced in study	- Significantly greater adherence to self-monitoring of weight and blood pressure in intervention vs. control group (M=79.3% vs. 68.6%, p=.01 and M=78.0% vs. 58.8%, p=.001) - Higher adherence to daily steps goal in intervention vs. control group (M=75.7% vs. 68.2%, p=.09)	Some concerns (RoB)
Schulte (2022) [217]	Inferred	Intervention (app) 1.1 <i>Goal setting^b</i> 2.2 <i>Feedback on behaviour^b</i> 2.3 <i>Self-monitoring of behaviour^b</i> 3.1 <i>Social support (unspecified - MI)^b</i> 4.1 <i>Instruction on how to perform behaviour^b</i> 6.2 <i>Social comparison^b</i> 7.1 <i>Prompts/cues^b</i> Control (educational website) 5.1 <i>Information about health consequences^b</i>	Evidenced in study	No significant difference between groups on outcomes (including total TWEETS engagement score and affective, cognitive, and behavioural components of TWEETS)	Some concerns (RoB)

Simons (2018) [263]	Explicit in paper	1.1 Goal-setting 1.5 Review behavior goals 2.2 Feedback on behavior 2.3 Self-monitoring ^b 3.1 Social support (unspecified) ^b 4.1 Instruction on how to perform the behavior 5.1 Information about health consequences 7.1 Prompts/cues	Evidenced in study	- Participant feedback indicated that notifications supported engagement - Tips, information, goals (and ability to adapt them), and graphical feedback on progress were well-received features; goal setting and feedback were considered motivating	**
Tombor (2019) [265]	Explicit in paper	Identity module (control: 13.5) 3.1 Social support (unspecified) ^b 9.3 Comparative imagining of future outcomes ^b 13.1 Identification of self as a role model ^b 13.2 Framing/reframing ^b 13.3 Incompatible beliefs ^b 13.4 Valued self-identity ^b 13.5 Identity associated with changed behaviour ^b 15.4 Self talk ^b Stress management module (control: 11.2, 12.6) 1.2 Problem solving ^b 1.4 Action planning ^b 8.3 Habit formation ^b 11.2 Reduce negative emotions ^b 11.3 Conserving mental resources ^b 12.6 Body changes ^b Health information module (control: 5.1) 5.1 Information about health consequences ^b 5.3 Information about social and environmental consequences 5.6 Information about emotional consequences F2F support module (control: 3.1) 3.1 Social support (unspecified) ^b 3.2 Social support (practical) ^b Behaviour substitution module (control: 4.2, 8.2, 12.4) 1.2 Problem solving ^b 1.4 Action planning ^b 4.2 Information about antecedents ^b 8.2 Behaviour substitution ^b 8.3 Habit formation ^b 12.4 Distraction ^b	Evidenced in study	- Authors report low engagement compared to another pregnancy health-related app - Main effects: only health information module had higher days of engagement significantly associated with intervention vs control - Combination effects: use of the full versions of identify and behavioural substitution modules had "a small but statistically significant effect on smoking abstinence"	Some concerns (RoB)
Ahonkhai (2021) [219]	Inferred	1.1 Goal setting (behaviour) 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) 6.2 Social comparison 10.3 Non-specific reward 14.2 Punishment	Hypothesised	- Identified key barriers to adherence from literature and interviews: forgetfulness, poor executive functioning, and poor social support - Hypothesised that goal setting, feedback, rewards, social comparison, social support, gamification, fun/playfulness, avatars will increase engagement	N/A

Curtis (2019) [222]	Explicit in paper	1.2 Problem solving 1.3 Goal setting (outcome) 1.4 Action planning 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.2 Social support (practical) 3.3 Social support (emotional) 4.2 Information about antecedents 5.1 Information about health consequences 6.1 Demonstration of the behaviour 6.2 Social comparison 7.1 Prompts/cues 8.1 behavioural practice/rehearsal 8.3 Habit formation 11.2 Reduce negative emotions 12.5 Adding objects to the environment	Hypothesised	- Identified key barriers to adherence from literature and interviews: capability (remembering), motivation (beliefs in effectiveness, confidence in self-managing), opportunity (limited time, changing social support) - Theoretical behavioural analysis (COM-B and TDF) used to propose BCTs to address barriers and support medication adherence	*****
D'Addario (2020) [208]	Explicit in paper	1.1 Goal setting (behaviour) 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) <i>6.2 Social comparison^c</i>	Hypothesised	- Focus groups identified preferences for autonomy and self-regulation features: suggests BCTs such as goal setting, feedback, and self-monitoring are associated with higher engagement - Social support and self-comparison were perceived as important for motivation - Participants did not want competition and social comparison for physical activity	*****
Duff (2018) [195]	Explicit in paper	1.1 Goal setting (behaviour) 2.3 Self-monitoring of behaviour 3.2 Social support (practical) 5.1 Information about health consequences	Hypothesised	- Identified most frequently used BCTs from a systematic review and incorporated in app design - Focus groups indicated that participants wanted to monitor progress, felt prompts and cues would support awareness and motivation, and would like social interaction	*****
Hales (2016) [225]	Inferred	2.2 Feedback on behaviour 7.1 Prompts/cues 3.1 Social support (unspecified) 3.3 Social support (emotional) 4.1 Instruction to perform the behaviour 6.2 Social comparison 10.2 Material reward 10.3 Non-specific reward	Hypothesised	Key themes from focus groups indicated that participants wanted to see others' progress and send them encouragement, found notifications and podcast information helpful, and suggested adding an incentive system and increasing praise / positive feedback	N/A

Smith (2022) [123]	Explicit in paper	1.1 Goal setting 1.4 Action planning 1.5 Review behaviour goals 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.3 Social support (emotional) 4.1 Instruction to perform the behaviour 5.1 Information about health consequences 7.1 Prompts and cues 9.1 Credible source 15.1 Written persuasion about capabilities	Hypothesised	- Focus groups with key stakeholders identified the most appropriate BCTs to achieve engagement with behaviour (physical activity) - Hypothesise that motivational interviewing (social support) and personalised reminders (prompts/cues) will support engagement with intervention	****
Songtaweesin (2020) [230]	Inferred	7.1 Prompts/cues 10.2 Material reward 10.3 Non-specific reward	Hypothesised	Focus groups and interviews identified recommendations for improving app engagement: dynamic content, notifications, fun features, earning redeemable points, and choose-your-own adventure storytelling	*****
Sporrel (2020) [231]	Inferred	1.1 Goal setting (behaviour) 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 4.1 Instruction to perform the behaviour 5.6 Information about emotional consequences 7.1 Prompts/cues 10.3 Non-specific reward	Hypothesised	- Focus groups and literature review identified preferences for reminders, feedback about goal progress, messages focusing on positive emotional outcomes, personalisable goals, and instruction videos - Social strategies not necessarily associated with increased motivation to engage in physical activity	N/A
Thornton (2021) [207]	Explicit in paper	1.1 Goal setting (behaviour) 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 7.1 Prompts/cues 10.3 Non-specific reward	Hypothesised	- Only a quarter of students reported wanting to use the app frequently, but provided positive feedback about behaviour tracking, goal setting, reminders, and personalised feedback features of the app - No consensus on whether competition would be helpful or detrimental	*
Whiteley (2018) [232]	Inferred	1.2 Problem solving 3.1 Social support (unspecified) 3.2 Social support (practical) 5.1 Information about health consequences 5.2 Salience of consequences 10.3 Non-specific reward 12.1 Restructuring the physical environment	Hypothesised	Interviews with youth indicated a desire for informational game content with comprehensive detail, future orientation, positive reinforcement, and that promote collaboration with healthcare providers and provided strategies	**

Whiteley (2019) [233]	Inferred	1.2 Problem solving 3.1 Social support (unspecified) 3.2 Social support (practical) 5.1 Information about health consequences 5.2 Salience of consequences 8.7 Graded tasks 10.3 Non-specific reward	Hypothesised	- Similar findings to Whiteley (2018) [232] - Interviews also indicated that participants wanted levels that become increasingly difficult (for a sense of accomplishment)	**
<i>Additional studies included after the update</i>					
Beck (2021) [235]	Explicit in paper	1.4 Action planning 2.2 Feedback on behaviour 2.3 Self-monitoring 7.1 Prompts/cues 10.3 Non-specific reward	Hypothesised	Potential BCTs to include were identified from preliminary focus groups and literature	N/A
Curtis (2015) [240]	Explicit in paper	1.1 Goal setting 1.9 Commitment 2.1 Monitoring of behavior by others without feedback 2.2 Feedback on behavior 2.3 Self-monitoring of behavior 3.1 Social support (unspecified) 3.2 Social support (practical) 3.3 Social support (emotional) 4.1 Instruction on how to perform the behavior 5.1 Information about health consequences 5.3 Information about social and environmental consequences 7.1 Prompts/cues 8.1 Behavioral practice/rehearsal 8.3 Habit formation 10.3 Non specific reward 12.1 Restructuring the physical environment 12.5 Adding objects to the environment 13.1 Identification of self as role model 13.4 Valued self-identity	Hypothesised	Potential BCTs to include were identified from preliminary focus groups and literature	N/A
Daniels (2023) [241]	Inferred	1.1 Goal setting 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.2 Social support (practical) 3.3 Social support (emotional) 4.1 Instruction on how to perform the behaviour 5.1 Information about health consequences 5.6 Information about emotional consequences 7.1 Prompts/cues	Hypothesised	Participant involvement in co-design and prototype testing indicated that social support, tailored notifications, encouragement, and information about health consequences were important features to include	N/A

Edwards (2018) [243]	Explicit in paper	1.1 Goal setting (behavior) 1.2 Problem solving 1.3 Goal setting (outcome) 1.6 Discrepancy between current behavior and goal 2.2 Feedback on behavior 2.3 Self-monitoring of behavior 2.7 Feedback on outcomes of behavior 3.1 Social support (unspecified) 3.2 Social support practical 4.2 Information about antecedents 5.1 Information about health consequences 5.2 Salience of consequences 5.3 Information about social and environmental consequences 5.6 Information about emotional consequences 7.1 Prompts or cues 8.2 Behavior substitution 8.7 Graded tasks 9.1 Credible source 9.2 Pros and cons 9.3 Comparative imagining of future outcomes 10.3 Nonspecific reward 10.4 Social reward 10.6 Nonspecific incentive 10.7 Self-incentive 10.9 Self reward 10.10 Reward (outcome) 11.1 Pharmacological support 11.2 Reduce negative emotions 12.1 Restructuring the physical environment 12.2 Restructuring the social environment 12.3 Avoidance or reducing exposure to cues for the behavior 12.4 Distraction 12.5 Adding objects to the environment 13.5 Identity associated with changed behavior 15.1 Verbal persuasion about capability 15.3 Focus of past success	Hypothesised	Focus groups indicated that the BCTs included in the app prototype were felt to be engaging and motivating	N/A
Garnett (2019) [247]	Explicit in paper	7.1 Prompts/cues Normative feedback module 2.7 Feedback on outcomes of behaviour 6.2. Social comparison Goal setting module 1.1 Goal setting (behaviour) Cognitive bias training module 7.8 Associative learning 8.1 Behavioural practice/rehearsal 8.3 Habit formation Self-monitoring and feedback module 1.5 Review behaviour goals	Hypothesised	Included BCTs were selected based on a scoping review, expert consensus study, content analysis, and a theory-based iterative design process to support engagement	N/A

		1.6 Discrepancy between current behaviour and goal 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 2.4 Self-monitoring of outcomes of behaviour 2.7 Feedback on outcomes of behaviour 5.2 Salience of consequences 5.6 Information about emotional consequences 10.3 Non-specific reward 10.4 Social reward 10.9 Self-reward Action planning module 1.4 Action planning 9.1 Credible source Identity change module 5.1 Information about health consequences 5.2 Salience of consequences 5.3 Information about social and environmental consequences 5.5 Anticipated regret 5.6 Information about emotional consequences 9.2 Pros and cons 13.2 Framing/reframing 13.3 Incompatible beliefs 13.4. Valued self-identity 13.5. Identity associated with changed behaviour			
Hendrie (2019) [248]	Explicit in paper	1.1 Goal setting (behavior) 1.4 Action planning 1.5 Review behavior goals 1.6 Discrepancy between current behavior and goal 1.7 Review outcome goals 2.2 Feedback on behavior 2.3 Self-monitoring of behavior 5.1 Information about health consequences 5.3 Information about social and environmental consequences 5.6 Information about emotional consequences 6.2 Social comparison 8.1 Behavioral practice or rehearsal 8.3 Habit formation 10.3 Nonspecified reward 10.6 Nonspecific incentive 14.5 Rewarding completion	Hypothesised	- Theory-based process conducted to select BCTs - Pilot feedback indicated that most users liked the app and would use it regularly (78% reported that they would use the app once a day after 3 days of use) - Users particularly liked tracking feature	N/A

Mummah (2016) [214]	Explicit in paper	1.1. Goal setting (behavior) 1.6. Discrepancy between current behavior and goal 1.9. Commitment 2.2. Feedback on behavior 2.3. Self-monitoring of behavior 4.1. Instruction on how to perform a behavior 5.1. Information about health consequences 6.2. <i>Social comparison^b</i> 6.3. Information about others' approval 7.1. Prompts/cues 7.3. Reduce prompts/cues 9.1. Credible source 10.4 Social reward 10.5. Social incentive 13.1. Identification of self as role model 13.2. Framing/reframing 13.5. Identity associated with changed behavior 15.1. Verbal persuasion about capability	Hypothesised	- Authors selected BCTs for inclusion based on themes generated from qualitative interviews - User feedback suggested that self-monitoring, feedback, information about health consequences, instruction on how to perform behaviour, and prompts/cues would be particularly important for engagement - Mixed opinions were reported about a social comparison feature: competition could motivate some participants but discourage others	N/A
Rohde (2019) [261]	Explicit in paper	1.1 Goal setting 1.4 Action planning 1.6 Discrepancy between current behavior and goal 2.2 Feedback on behavior 2.3 Self-monitoring (behavior) 3.1 Social support (unspecified) 4.1 Instruction on how to perform a behavior 5.1 Information about health consequences 6.2 Social comparison 7.1 Prompts/cues 8.7 Graded tasks 10.3 Non-specific reward 15.1 Verbal persuasion about capability	Hypothesised	BCW-based BCTs selection process was conducted with the aim of supporting engagement	N/A
Willms (2023) [199]	Explicit in paper	1.1 Goal setting 1.4 Action planning 1.5 Review behavior goals 2.3 Self-monitoring of behavior 3.2 Social support (practical) 3.3 Social support (emotional) 5.1 Information about health consequences 5.4 Monitoring of emotional consequences 7.1 Prompts/cues 8.1 Behavioral practice or rehearsal 8.3 Habit formation 8.7 Graded tasks 10.1 Material incentive for behavior 12.1 Restructuring the physical environment 12.4 Distraction 12.5 Adding objects to the	Hypothesised	- Two thirds of participants reported finding the app useful and enjoyable - Participants reported being interested in the financial incentives	N/A

		environment 13.3 Incompatible beliefs 13.4 Valued self-identity			
Baumel (2018) [197]	Inferred	2.2 Feedback on behaviour 10.3 Non-specific rewards	Inferred by reviewers	- Two items of the Therapeutic Persuasiveness domain scale are aligned with BCTs: rewards and ongoing feedback - Both were significantly positively associated with mobile app user retention after 30 days ($r=0.26$, $p=0.03$ and $r=0.33$, $p=0.01$, respectively) - Neither were significantly associated with usage time	N/A
Buman (2016) [220]	Explicit in paper	Sleep component: 1.1 Goal setting (behaviour) 1.2 Problem solving 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 4.2 Information about antecedents 7.3 Reduce prompts/cues Sedentary behaviour component: 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 5.1 Information about health consequences 7.1 Prompts/cues 8.2 behaviour substitution Exercise component: 1.1 Goal setting (behaviour) 1.2 Problem solving 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 4.1 Instruction to perform the behaviour 5.1 Information about health consequences 7.1 Prompts/cues	Inferred by reviewers	- Steady use of sleep and sedentary app components over intervention period (~30min/week), but not exercise component - implies potential association between BCTs included in these components and engagement - Participants disliked exclusion of 'lighter intensity' exercise activities - implies importance of BCT 8.7 Graded tasks^c	Some concerns (RoB)
Caon (2022) [221]	Explicit in paper	7.1 Prompts/cues ^b 10.3 Non-specific rewards ^b	Inferred by reviewers	- Just over 1/3 of participants (37.5% used app for more than 2 weeks) - No direct association between engagement and BCTs included in app	*****

Cueto (2019) [211]	Explicit in paper	1.1 Goal setting (behaviour) 1.2 Problem solving 1.3 Goal setting (outcome) 1.4 Action planning 1.5 Review behaviour goal(s) 1.7 Review outcome goals 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 2.5 Monitoring of outcomes of behaviour 3.1 Social support (unspecified) 8.2 Behaviour substitution 8.3 Habit formation	Inferred by reviewers	High overall program retention (79.9%) implies that included BCTs associated with engagement	****
Davis (2019) [209]	Explicit in paper	Noom Walk Pedometer: 2.3 <i>Self-monitoring of behavior</i> ^b 6.3 <i>Information about others' approval</i> ^b MapMyFitness: 1.1 Goal setting (behaviour) 1.5 Review behavioural goals 2.3 <i>Self-monitoring of behavior</i> ^b 2.4 Self-monitoring of outcomes of behaviour 4.1 Instruction on how to perform the behaviour 6.2 Social comparison 6.3 <i>Information about others' approval</i> ^b 7.1 Prompts/cues 8.7 Graded tasks 10.3 Non-specific reward	Inferred by reviewers	- Comparison of apps with different numbers of BCTs found the app with more BCTs had a significantly higher mean rating on the uMARS engagement subscale ($t(81) = 6.71, p < .001, g = 1.15$, 95% confidence interval (1.02, 1.87)) - Implies that BCTs present only in MapMyFitness might contribute to engagement - Potential confounding factors: apps could differ on items other than BCTs (eg. usability, aesthetics)	Low RoB
DeSmet (2019) [210]	Explicit in paper	2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 4.1 Instruction on how to perform behaviour 5.1 Information on health consequences, self-monitoring	Inferred by reviewers	- Survey found that most preferred BCTs were: information on behaviour health outcomes, obtaining insights into their healthy lifestyles, self-monitoring, feedback, instructions, or tips - Lowest ranked were: social support and social comparison	***
Hochsman (2019) [224]	Explicit in paper	1.1 Goal setting (behaviour) 1.4 Action planning 2.2 Feedback on behaviour 7.1 Prompts/cues 10.3 Non-specific reward	Inferred by reviewers	- Intervention group had a significant improvement in intrinsic physical activity motivation (IMI total score) compared to control group (adjusted difference of 8.15 points, 95% CI 0.90-15.39; $P = .03$) - Intervention group also had a significantly increase in the 'interest/enjoyment' subscale (adjusted difference = 2.03 points, 95% CI 0.04-4.09; $P = .049$) - Intervention group had a weekly average of 131.1 (SD 48.7) minutes of in-game walking and 15.3 (SD 24.6) minutes of strength training	Some concerns (RoB)

				Implies potential positive association between included BCTs and engagement	
Nurmi (2020) [229]	Explicit in paper	1.1 Goal setting (behaviour) 1.3 Goal setting (outcome) 1.4 Action planning 1.5 Review behaviour goal(s) 1.6 Discrepancy between current behaviour and goal 1.7 Review outcome goals 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 2.6 Biofeedback 3.1 Social support (unspecified - motivational interviewing) 10.4 Social reward 15.2 Mental rehearsal of successful performance 15.3 Focus on past success	Inferred by reviewers	- Interviews and theory identified key features relating to autonomy support and change talk - Participants reported that features related to autonomy (eg. personalised goals, tailored feedback) and relatedness (motivational interviewing features) helped support motivation to engage	*****
Valle (2020) [212]	Inferred	1.1 Goal setting (behaviour) ^b 2.2 Feedback on behavior ^b 2.3 Self-monitoring of behavior ^b 3.1 Social support (unspecified) ^b 7.1 Prompts/cues ^b	Inferred by reviewers	- No direct association evidenced between BCTs and engagement measures - Past success appeared to be associated with increased engagement - BCT 15.3 Focus on past success^c could potentially help support engagement, but possible that people who had greater past success were more committed for other reasons	***
Additional studies included after the update					
Abdul Khalil (2023) [234]	Inferred	2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) 4.1 Instruction on how to perform behaviour (<i>macro engagement</i>) 5.1 Information about health consequences 6.2 Social comparison 6.3 Information about others' approval 7.3 Prompts / cues 10.3 Non-specific reward	Inferred by reviewers	- Factors identified in interviews as associated with increasing self-awareness and thus motivation to engage with the app or behaviour included: self-monitoring, feedback, encouragement, social comparison, rewards, feedback on their behaviour, information about consequences, and reminders - Factors associated with app adherence included social	*****

				support and reminders	
Attwood (2017) [200]	Explicit in paper	2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour	Inferred by reviewers	Self-monitoring and feedback were the only included BCTs that were linked with engagement in the qualitative feedback from participants	**
Burgermaster (2023) [238]	Inferred	All versions 2.3 Self-monitoring of behaviour 2.4 Self-monitoring of outcomes External Regulation version (rewards) <i>1.1 Goal setting^b</i> 10.2 Material reward Introjected Regulation version (RDF) <i>1.1 Goal setting^b</i> 2.2 Feedback on behaviour Identified Regulation (SA) 1.5 Review behavioural goals Integrated Regulation (FORC) 5.1 Information about health consequences Not included in apps 3.1 Social support (unspecified) ^d	Inferred by reviewers	- Different types of app users engaged with different types of features and had different reactions to particular features, depending on their type of motivation and regulation (SDT) - Features that were highlighted as being used and appreciated for users with internal motivation and regulation included self-monitoring, feedback, reviewing behavioural goals, and informational support - Externally motivated and regulating users needed more support and encouragement to stay engaged	*****
Hendrie (2020) [194]	Explicit in paper	<i>1.1 Goal setting (behavior)^b</i> <i>1.4 Action planning^b</i> <i>1.5 Review behavior goals^b</i> <i>1.6 Discrepancy between current behavior and goal^b</i> <i>1.7 Review outcome goals^b</i> 2.2 Feedback on behavior 2.3 Self-monitoring of behavior <i>5.1 Information about health consequences^c</i> <i>5.3 Information about social and environmental consequences^c</i> <i>5.6 Information about emotional consequences^c</i> <i>6.2 Social comparison^c</i> <i>8.1 Behavioral practice or rehearsal^c</i> <i>8.3 Habit formation^c</i> <i>10.3 Nonspecified reward^c</i> <i>10.6 Nonspecific incentive^c</i> <i>14.5 Rewarding completion^c</i>	Inferred by reviewers	- Most commonly used app feature was Veg Lookup (~60% of users on a given day): based on feature description, associated BCTs inferred to be feedback and self-monitoring - Least commonly used app features were Challenges, Meal Ideas, and Learn (~5-25% of users on a given day): based on feature description, associated BCTs inferred to be rewards, incentives, feedback, behavioural practice, habit formation, review outcome goals, and information about consequences	****
Kekkonen (2023) [249]	Inferred	<i>1.1 Goal setting^b</i> 2.3 Self-monitoring of behaviour <i>3.1 Social support (unspecified)^c</i> 6.2 Social comparison 7.1 Prompts/cues <i>8.1 Behavioural practice/rehearsal^b</i> <i>9.1 Credible source^b</i> <i>10.4 Social reward^b</i>	Inferred by reviewers	- Lack of or malfunctioning notifications associated with reported lack of engagement with the app - Social comparison feature suggested to improve motivation to use app - Only 5% of users wanted social networking feature in app - Self-monitoring could help increase engagement, but mostly if it was sensor-based and required minimal user input	*****

Lawrason (2023) [250]	Explicit in paper	1.1 Goal setting 1.2 Problem solving 1.4 Action planning 1.5 Review behaviour goals 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) 4.1 Instruction on how to perform behaviour 4.2 Information on antecedents 5.1 Information about health consequences 10.6 Non-specific incentive ^d 15.1 Verbal persuasion about capability 15.3 Focus on past success 15.4 Self-talk	Inferred by reviewers	- Authors reported that users spent "adequate time" on the app - Users recommended incentives and integration with other apps to support community engagement as well as positive feedback and peer mentors (social support) - Users reported liking self-monitoring and feedback	Some concerns (RoB)
McClure (2023) [254]	Inferred	Both versions: 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) - CBT 4.1 Instruction on how to perform behaviour 5.3 Information about social and environmental consequences 10.2 Material reward 10.3 Non-specific reward 15.4 Self-talk Enhanced version only: 1.2 Problem solving 4.2 Information about antecedents 8.1 Behavioural practice / rehearsal 8.3 Habit formation 12.3 Avoidance/reducing exposure to cues for behaviour 13.2 Framing / reframing	Inferred by reviewers	- Significantly greater usage (number of sessions; 19.9 vs. 7.3, $p < .001$) in enhanced version than control but no difference in days of use (43.0 vs. 41.8, $p = .90$) - Reported satisfaction and use of features was similar between groups	Some concerns (RoB)
Mummah (2016) [215]	Explicit in paper	1.1. Goal setting (behavior) 1.6. Discrepancy between current behavior and goal 1.9. Commitment 2.2. Feedback on behavior 2.3. Self-monitoring of behavior 4.1. Instruction on how to perform a behavior 5.1. Information about health consequences 6.2. <i>Social comparison</i> ^b 6.3. Information about others' approval 7.1. Prompts/cues 7.3. Reduce prompts/cues 9.1. Credible source 10.4 Social reward 10.5. Social incentive 13.1. Identification of self as role model 13.2. Framing/reframing 13.5. Identity associated with changed behavior 15.1. Verbal persuasion about capability	Inferred by reviewers	- Participants in pilot study generally found app enjoyable and not boring - Authors reported that app "achieved reasonably high rates of engagement" [215] - Notifications were associated with increased self-monitoring via app	Some concerns (RoB)

Nahum-Shani (2021) [256]	Inferred	Reciprocity condition: 3.3 Social support (unspecified) 7.1 Prompts/cues Non-monetary reinforcement condition A - entertaining content: 10.3 <i>Non-specific reward^b</i> Non-monetary reinforcement condition B - personalised info: 2.2 Feedback on behaviour	Inferred by reviewers	• Microrandomised trial reports supported impact of reciprocity condition but not entertaining content on self-reporting the day prompt was received (although neither statistically significant) • Personalised info had a positive effect specifically for users at high risk of disengaging (no self-reporting the day before) but a negative effect on those at low risk of disengaging (self-reported the day before)	Low RoB
Nicholas (2021) [257]	Explicit in paper	1.1 Goal setting behavior 1.2 Problem solving 1.3 Goal setting outcome 1.4 Action planning 1.5 Review behavioral goals 1.6 Discrepancy between current behavior goal 2.2 Feedback on behavior 2.3 Self-monitoring of behavior 3.1 Social support unspecified 3.2 <i>Social support (practical)^b</i> 3.3 Social support (emotional) 4.1 <i>Instruction on how to perform the behavior^b</i> 5.1 <i>Information about health consequences^b</i> 5.4 <i>Monitoring of emotional consequences^b</i> 7.1 Prompts/cues 8.7 <i>Graded tasks^b</i> 10.4 Social reward	Inferred by reviewers	• App received moderate score for engagement subscale of uMARS (M=2.95/5) • Feedback suggested that goal setting, motivational messages, and self-monitoring helped support engagement with behaviour	*****
Patterson (2024) [258]	Explicit in paper	1.4 Action planning 1.6 <i>Discrepancy between current behaviour and goal^b</i> 2.7 Feedback on outcome(s) of behaviour 4.2 Information about antecedents 7.1 Prompts/cues 8.1 <i>Behavioural practice/rehearsal^b</i> 8.2 <i>Behaviour substitution^b</i> 8.3 <i>Habit formation^b</i> 8.4 <i>Habit reversal^b</i> 9.1 <i>Credible source^b</i> 10.4 <i>Social reward^b</i> 12.1 <i>Restructuring the physical environment^b</i> 12.2 <i>Restructuring the social environment^b</i> 12.3 <i>Avoidance/reducing exposure to cues for the behaviour^b</i> 13.1 <i>Identification of self as a role model^b</i>	Inferred by reviewers	• Qualitative feedback suggested that users felt goals and ability to track and view progress in app, social support, and personalisation based on their own data and reasons for sedentary behaviour were important for engagement • Notifications were felt to help support engagement, but personalisation could increase this	*****
Perski (2017) [204]	Inferred	1.1 Goal setting 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 5.1 <i>Information about health consequences^c</i>	Inferred by reviewers	• Qualitative evaluation of users' preferences for features to support engagement in a variety of apps generated several themes that could be mapped to	*****

		6.3 Information about others approval 7.1 Prompts/cues 9.1 Credible source 10.2 Material reward 10.4 Social reward		BCTs • BCTs inferred as relating to themes influencing selection of apps included credible source, information about others approval • BCTs inferred as being incorporated in features deemed important for engagement included: self-monitoring, feedback, goal setting, social or material rewards, prompts/cues • Negatively-framed information about health consequences could discourage engagement	
Sanders (2023) [262]	Explicit in paper	1.1 Goal setting 1.2 Problem solving 1.5 Review behavioural goals 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) 4.1 Instruction on how to perform behaviour 5.1 Information about the health consequences 6.1 Demonstration of behaviour 6.2 Social comparison 5.2 Salience of consequences 7.1 Prompts/cues 8.3 Habit formation 8.7 Graded tasks 9.1 Credible source 10.4 Social reward 13.2 Framing / reframing 15.3 Focus on past success	Inferred by reviewers	• Over 28 day intervention period, users engaged with app a mean number of 77 times for ~13 minutes/week • Analysis of app use data indicated that users primarily landed on default dashboard page initially then engaged with pages about "my stats" and "my goals" (incorporating "BCTs of goal setting, feedback on progression toward goals, and self-monitoring")	***
Stancu (2022) [264]	Inferred	1.4 Action planning (menu flexibility) 2.2 Feedback on behaviour (informative feedback) 3.1 Social support (evaluative feedback) 10.3 Non-specific reward	Inferred by reviewers	• Discrete choice experiments demonstrated user preferences for feedback and social support • Different user preferences in Spain and Germany • Users also preferred personalisability - e.g. the ability to self-select their rewards and personalised the recommended menu	****
Whale (2021) [218]	Explicit in paper	1.1 Goal setting 1.2 Problem solving 1.4 Action planning 1.5 Review behaviour goals 1.6 Discrepancy between current behaviour and goal 1.8 Behavioural contract 1.9 Commitment 2.1 Monitoring of behaviour by others 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 3.1 Social support (unspecified) 4.1 Instruction on how to perform the behaviour 5.1 Information about health	Inferred by reviewers	• Participant feedback about app was generally positive • Participants, particularly older ones, found self-monitoring and feedback on progress (visual or summary) helpful • Participants, particularly younger ones, found that the rewards feature supported engagement	N/A

		consequences of the behaviour 7.1 Prompts/cues 8.7 Graded tasks 10.3 Non-specific Rewards 11.2 Reduce negative emotions			
Zimmermann (2023) [266]	Inferred	1.1 Goal setting 1.2 Problem solving 2.2 Feedback on behaviour 2.3 Self-monitoring of behaviour 2.4 Self-monitoring of outcomes of behaviour 3.1 Social support (unspecified) 3.2 Social support (practical) 3.3 Social support (emotional)	Inferred by reviewers	- Over the 90 day intervention period, there was a median of 72.5 days of app use (range: 0-90) - The self-monitoring feature was used by most participants - Authors suggest that health coach feature supported user engagement, as users who engaged with this feature were more engaged with program overall	****

^aOverall RoB2 assessments are reported as 'high', 'low', or 'some concerns' and overall MMAT assessments are reported as * (0-5); studies that were not assessed (e.g. development focus, content analyses) are noted as 'N/A'

^bThe authors deemed there was insufficient evidence of a potential association with engagement; they were not included in Table 2-6 (italicised grey text)

^cThe evidence suggested that these BCTs were not associated with engagement; they were not included in Table 2-6 (italicized red text)

^dThese BCTs were not included in the apps evaluated, but were deemed to have potential associations with engagement based on the papers' reported findings and included in Table 2-6

Appendix 2-I. MMAT quality assessment* [192]

<i>Qualitative studies</i>							
First author (year)	S1. Are there clear research questions?	S2. Do the collected data allow to address the research questions?	1.1. Is the qualitative approach appropriate to answer the research question?	1.2. Are the qualitative data collection methods adequate to address the research question?	1.3. Are the findings adequately derived from the data?	1.4. Is the interpretation of results sufficiently substantiated by data?	1.5. Is there coherence between qualitative data sources, collection, analysis and interpretation?
Curtis (2019) [222]	N - but aim was reported	Y	Y	Y	Y	Y	Y
D'Addario (2020) [208]	N - but aim was reported	Y	Y	Y	Y	Y	Y
Duff (2018) [195]	N - but aim was reported	Y	Y	Y	Y	Y	Y
Songtaweessin (2020) [230]	N - but aim was reported	Y	Y	Y	Y	Y	Y
Nurmi (2020) [229]	Y	Y	Y	Y	Y	Y	Y
Smith (2022) [123]	N - but aim was reported	Y	Y	Y	Can't tell - method of data analysis not described in detail	Y	Y
Szinay (2021) [201]	N - but aim was reported	Y	Y	Y	Y	Y	Y
Tang (2015) [202]	Y	Y	Y	Y	Y	Y	Y
<i>Update</i>							
Abdul Khalil (2023) [234]	N - but aim was reported	Y	Y	Y	Y	Y	Y
Burgermaster (2023) [238]	Y	Y	Y	Y	Y	Y	Y
Domin (2022) [213]	Y	Y	Y	Y	Y	Y	Y
Ekezie (2021) [245]	N - but aim was reported	Y	Y	Y	Y	Y	Y
Flaherty (2021) [205]	Y	Y	Y	Y	Y	Y	Y
Kekkonen (2023) [249]	Y	Y	Y	Y	Y	Y	Y

Litterbach (2017) [252]	N - but aim was reported	Y	Y	Y	Y	Y	Y
Montilva-Monsalve (2023) [255]	Y	Y	Y	Y	Y	Y	Y
Patterson (2024) [258]	Y	Y	Y	Y	Y	Y	Y
Perski (2017) [204]	Y	Y	Y	Y	Y	Y	Y
Roberts (2019) [206]	N - but aim was reported	Y	Y	Y	Y	Y	Y
Quantitative non-randomised studies							
First author	S1. Are there clear research questions?	S2. Do the collected data allow to address the research questions?	3.1. Are the participants representative of the target population?	3.2. Are measurements appropriate regarding both the outcome and intervention (or exposure)?	3.3. Are there complete outcome data?	3.4. Are the confounders accounted for in the design and analysis?	3.5. During the study period, is the intervention administered (or exposure occurred) as intended?
Cueto (2019) [211]	N - but aims and hypotheses reported	Y	Y - retrospective cohort analysis of real world users	Y	Y - users with incomplete data were excluded	N - not in design, but adjustments made in analysis	Y - no interference from researchers, as data collected retrospectively
Delmas (2020) [196]	Y	N - outcome measures related to engagement, not assessing effectiveness of education	Y	Y	Y	N	Y - real world use
Valle (2020) [212]	N - but aim was reported	Y	N - ~80% were female, most highly educated	Y	Y	Y - to some extent	N - but necessary to examine outcome in question (explained in section 2.2.1)
Quantitative descriptive studies							
First author	S1. Are there clear research questions?	S2. Do the collected data allow to address the research questions?	4.1. Is the sampling strategy relevant to address the research question?	4.2. Is the sample representative of the target population?	4.3. Are the measurements appropriate?	4.4. Is the risk of nonresponse bias low?	4.5. Is the statistical analysis appropriate to answer the research question?

Caon (2022) [221]	N - but aim was reported	Y	Y	Y	Y - app use data commonly used to assess engagement, although this has limitations	Y - engagement / non-response is an outcome measure	Y
DeSmet (2019) [210]	Y	Y	Y	N - primarily well-educated	Y	N	Y
Lin (2021) [226]	N - but aims were reported	Y	Y	N - not necessarily inclusive of marginalised / disadvantaged groups, or other countries / populations	Y	Y - engagement / non-response is an outcome measure	Y
<i>Update</i>							
Bell (2020) [198]	N - but aim was reported	Y	Y - included data from all existing users meeting criteria	Y	Y	Y - N/A	Y
Ehlers (2014) [244]	N - but aim was reported	Y	Y - sample of convenience but recruitment targeted for population of interest	N - "age distribution of the sample was positively skewed... sample was disproportionately Caucasian, educated, and affluent" [244]	Y - not validated but developed based on specific theoretical frameworks	Can't tell - demographics of completers vs. non-completers not specified	Y
Fatima (2023) [246]	N - but aim was reported	Y	Y - sample of convenience but recruitment targeted for population of interest	Can't tell - a variety of age groups included, but minimal other demographic data provided	Y	Can't tell - unclear how many were recruited vs. completed pilot	Y
Hendrie (2020) [194]	Y	Y	Y - sample of convenience but reflects real-world adoption/use	N - females and young people overrepresented compared to population	Y	Y - engagement / non-response is an outcome measure	Y
Lim (2022) [251]	N - but aim was reported	Y	Y - subset analysis of RCT	Can't tell - paper does not discuss generalisability of sample to target population	Y	Y - engagement / non-response is an outcome measure	Y
Pfaeffli Dale (2019) [259]	N - but aim was reported	Y	Y	Y - real-world study of existing app users	Y	Y - engagement / non-response is an outcome measure	Y
Sanders (2023) [262]	N - but aim was reported	Y	N - convenience sample of participants previously involved in program, may be biased	Y - physically inactive, mix of gender and ethnicity	Y	Can't tell - unclear how many were recruited	Y

Stancu (2022) [264]	N - but aim was reported	Y	Y - convenience sample, independently recruited	Can't tell - few demographic characteristics presented, appears representative for gender	Y	Y	Y
Zimmermann (2023) [266]	N - but aim was reported	Y	Y - self-selection reflects real-world adoption/use	Can't tell - paper does not discuss generalisability of sample to target population	Y	Y - engagement / non-response is an outcome measure	Y
Mixed-methods studies							
First author	S1. Are there clear research questions?	S2. Do the collected data allow to address the research questions?	5.1. Is there an adequate rationale for using a mixed methods design to address the research question?	5.2. Are the different components of the study effectively integrated to answer the research question?	5.3. Are the outputs of the integration of qualitative and quantitative components adequately interpreted?	5.4. Are divergences and inconsistencies between quantitative and qualitative results adequately addressed?	5.5. Do the different components of the study adhere to the quality criteria of each tradition of the methods involved?
Garnett (2015) [182]	Y	Y	Y	Y	Y	Y - N/A	Y
Hales (2016) [225]	Y	Y	N - authors did not explicitly describe	N - qual and quant results reported separately, no section discussing how they inform each other	N - see 5.2	N - not addressing the same questions	Can't tell - process of identifying and analysing themes not described in detail
Thornton (2021) [207]	N - but aims were reported	Y	N - authors did not explicitly describe	Y	N - not in detail	Y - N/A	Can't tell - process of identifying and analysing themes not described in detail
Whiteley (2018) [232]	N - but aim was reported	Y	N - authors did not explicitly describe	Y	N - not in detail	Y - N/A	Y
Whiteley (2019) [233]	N - but aim was reported	Y	N - authors did not explicitly describe	Y	N - not in detail	Y - N/A	Y
Update							
Attwood (2017) [200]	N - but aim was reported	Y	Y	N - qual and quant results reported separately, no section discussing how they inform each other	N - see 5.2	N	Y
Lubbans (2014) [253]	N - but aim was reported	Y	N - authors did not explicitly describe	Y	N - not in detail	Y - N/A	Y
Nicholas (2021) [257]	N - but aim was reported	Y	Y	Y	Y	Y - N/A	Y

Resnick (2021) [260]	N - but aim was reported	Y	N - authors did not explicitly describe	Y	N - not in detail	Y - N/A	Y
Simons (2018) [263]	N - but aim was reported	Y	N - authors did not explicitly describe	N - qual and quant results reported separately, no section discussing how they inform each other	N - not in detail	Y - N/A	Y

**NOTE: 3 studies not included from the first phase (2 development studies, 1 content analysis) and 10 from the update (development studies)*

Appendix 3-A. DBCI Engagement Scale [277,279]

Please answer the following questions with regards to your most recent use of the [Drink Less] app.

How strongly did you experience the following? (Measured on a 7-point scale with end-points and middle anchored: not at all; moderately; extremely)

1. Interest
2. Intrigue
3. Focus
4. Inattention
5. Distraction
6. Enjoyment
7. Annoyance
8. Pleasure

9. How much time (in minutes) do you roughly think that you spent on the app?

10. Which of the app's components do you remember visiting? (Indexed as a proportion of available modules (e.g. $5/7 * 100 = 71.4\%$))

- Calendar (Self-monitoring/feedback)
- Create and view goals (Goal setting)
- What has and hasn't worked (Self-monitoring/feedback)
- Create and view action plans (Action planning)
- Your hangover and you (Self-monitoring/feedback)
- Review your drinking (Normative feedback)
- Dashboard (Self-monitoring/feedback)
- Game (Cognitive bias re-training)
- Drink + me (Identity change)
- Useful information (Other)
- Other (Other)
- Can't remember (Other)

Appendix 3-B. Twente Engagement with Ehealth Technologies Scale (TWEETS)

[278]

Item	Thinking about using [the technology] the last week, I feel that:	Construct
1	[this technology] is part of my daily routine	Behaviour
2	[this technology] takes me little effort to use	Behaviour
3	I'm able to use [this technology] as often as needed (to achieve my goals)	Behaviour
4	[this technology] makes it easier for me to work on [my goal]	Cognition
5	[this technology] motivates me to [reach my goal]	Cognition
6	[this technology] helps me to get more insight into [my behaviour relating to the goal]	Cognition
7	I enjoy using [this technology]	Affect
8	I enjoy seeing the progress I make in [this technology]	Affect
9	[This technology] fits me as a person	Affect

Appendix 4-A. Case Study 1: Functional overview of NoObesity version 1 apps
(adapted from [353])

App function	Purpose	Outcome	Family app?	HCP app?	BCTs^a
Set family goals	This function enables families to set behavioural goals to support them in their health and wellbeing. This function takes them through a Specific, Measurable, Action-oriented, Realistic, Timed, Evaluated, Reviewed (SMARTER) process to make plans and set goals.	Families set goals using a SMARTER process that is part of the Making Every Contact Count (MECC) program.	Y	N	1.1, 1.2, 1.4, 1.6
Record family progress	This function allows families to record how they have been progressing with family goals and enables them to evaluate and review their goals if they have not progressed as expected.	Families are able to record their progress and review their goals.	Y	N	2.3
Update photo	This function is in place for families to be able to upload pictures of themselves doing healthy things. This could be a meal they have cooked, pictures of them on a health walk, etc.	This is used as an engagement tool allowing families to upload photos of things they have done as a way of encouraging them to come back to the app.	Y	N	-
Family survey	This function supports families to take stock of where they are around a number of health-related areas (water, fruit, vegetables, physical activity, sleep, screen time, toothbrushing, happiness, body size). These questions can be completed for all members of the family. Once the survey is completed, the app will prompt families to think about setting a goal around one of the areas of the health survey (if they have scored low on it).	It enables collection of baseline data to allow for comparison of change.	Y	N	1.6, 2.2
Parent's survival guide	This function supports families in thinking about how they can overcome the challenges that they face when supporting their children around healthy eating, etc.	Families gain solutions on how they could approach issues and challenges using a MECC approach.	Y	N	1.2

Healthy choices & healthy challenges	These two functions are developed in the form of games within the app that HCPs can play as an approach to learning in a fun way about what makes a healthy diet and what level of physical activity is required to burn off calories.	It enables families to learn about the Eatwell Plate and about physical activity and provides HCPs with knowledge to support families.	Y	Y	4.1
Useful links	This function provides links to websites for HCPs to access for furthering their own learning and Continuing Professional Development.	It enables families and HCPs to learn about resources available; enables HCPs to be better informed on possible signposting.	Y	Y	-
HC Points / Certification	This function encourages families and HCPs to do all of the above as they gain points and awards on further exploring the content of the app.	Families and HCPs have used the app to develop their knowledge and skills in supporting the prevention and management of childhood obesity.	Y	Y	10.3
Link accounts	This function enables families and HCPs to connect their accounts so that the HCP can see the families' data and progress.	HCPs will have a better understanding of the families' behaviours and progress, enabling them to provide more tailored advice in appointments.	Y	Y	3.1
How to help families	This function enables the HCP to follow a family (through a story), where the family comes into contact with a number of different health care workers. The app then provides the HCP with information on how interactions with the family can be improved for their own practice.	HCPs will have an increased awareness of the potential opportunities that could be missed in supporting families around health and wellbeing. HCPs will be enabled to reflect on their own practice when coming into contact with families. HCPs will see (read) how a MECC approach can be used for their own practice in supporting families.	N	Y	4.1, 6.1
Common issues	This function provides some of the common issues HCPs face when supporting families around healthy eating, diet, and activity. This function offers HCPs with some solutions on how they can overcome the common issues and challenges they face. The solutions are developed with MECC principles, which allow HCPs to use MECC skills in supporting families.	HCPs are able to learn about solutions they can use to overcome issues and challenges and are enabled to practise these MECC skills.	N	Y	4.1

"No specific behaviour change theory was used in the design of the app, the BCTs have been identified here post hoc by the authors to aid future analyses and comparisons of interventions"

Appendix 4-B. Case Study 1: Alignment of NASSS and RE-AIM frameworks

[341–344]

NASSS domains	RE-AIM dimensions	Key considerations	Key outcomes
<i>Domain 1: The condition or illness</i>	<i>Reach</i>	Both emphasise importance of contextual factors and evaluating technology in real user population	- Sample characteristics (demographics) - Participation and drop-out numbers and rate
Characteristics of users (clinical, socio-cultural, etc.)	Number, proportion, and representativeness participants (user)		
<i>Domain 2: The technology</i>	<i>Implementation</i>	- Both focus on key issues of usability, acceptability, and other factors associated with individual use - RE-AIM captures issue of whether the intervention was delivered as intended	- System Usability Scale - Key themes about usability, acceptability, and app features - Barriers - Suggestions for improvement
Features of technology associated with usability, acceptability, trustworthiness, and sustainability	Fidelity of intervention delivery (system), use of intervention strategies (user)		
<i>Domain 3: The value proposition</i>	<i>Effectiveness (or Efficacy)</i>	- Both emphasise the importance of evidence to support adoption of intervention (safety, positive impact on intended outcomes) - NASSS considers impact of intervention for suppliers as well as users, and how that factors into the intervention's sustainability	- Key themes about motivation and perceived impact on self-efficacy and behaviours - Ratings of app motivation strategies - Self-reported success at goals (in app and survey) - Self-reported self-efficacy (in app and survey)
Evidence of impact and value, for users and suppliers (incl. safety, efficacy, benefit, affordability)	Impact on intended outcomes		
<i>Domain 4: The adopter system</i>	<i>Adoption</i>	Both examine adoption by staff /	Key themes about adoption of

Adoption (and continued use) by staff and users; reasons for non-adoption / abandonment; assumptions about use (capability, opportunity)	Number, proportion, and representativeness of places/people willing to adopt intervention (system)	setting delivering the intervention • NASSS includes adoption by users, which is captured by RE-AIM ('Reach') • Key questions around adoption: how much, how long, why, why not, what is required?	HCP link and impact of name on adoption • Ratings of HCP link and app name
Domain 5: The organization	N/A	• Domains 5 and 6 both related to various elements of the adoption and maintenance of the intervention, so have been analysed independently from the RE-AIM framework	• Key themes from qualitative interview with HEE representatives
Organization's capacity and readiness for adoption and scale-up; recognition of need for and availability of dedicated resources; disruption to routines; work to implement	N/A		
Domain 6: The wider context	N/A		
Political, economic, regulatory, professional, sociocultural factors	N/A		
Domain 7: Embedding and adaptation over time	Maintenance	• Both capture institutional / organisational maintenance • RE-AIM also captures individual maintenance • In this study, the organisational aspects of ongoing maintenance were	• Self-reported intention to continue use of app • Self-reported length of use (if discontinued) • App use data from system (length of use,

Feasibility of ongoing adaptation; organisational resilience (willingness to reflect and adapt)	Institutionalisation of intervention in routine processes (system), long-term impact on outcomes (user)	closely aligned with key themes in Domains 5 and 6, and were grouped with them in the analysis	most recent use date)

Appendix 4-C. Case Study 1: SRQR checklist

Page #

Title and abstract

Title - Concise description of the nature and topic of the study Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended	1
Abstract - Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	2

Introduction

Problem formulation - Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	4-5
Purpose or research question - Purpose of the study and specific objectives or questions	5

Methods

Qualitative approach and research paradigm - Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/ interpretivist) is also recommended; rationale**	5-6
Researcher characteristics and reflexivity - Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	9
Context - Setting/site and salient contextual factors; rationale**	7
Sampling strategy - How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale**	6-7
Ethical issues pertaining to human subjects - Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	7
Data collection methods - Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and	7-8

modification of procedures in response to evolving study findings; rationale**	
Data collection instruments and technologies - Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	7-8, Table 1
Units of study - Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	6-7, 10, Figure 1, Table 2
Data processing - Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/de-identification of excerpts	8-9
Data analysis - Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale**	9
Techniques to enhance trustworthiness - Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale**	9

Results/findings

Synthesis and interpretation - Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	9-23
Links to empirical data - Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	Textbo x 1-6

Discussion

Integration with prior work, implications, transferability, and contribution(s) to the field - Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/generalizability; identification of unique contribution(s) to scholarship in a discipline or field	23-28
Limitations - Trustworthiness and limitations of findings	26-27

Other

Conflicts of interest - Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	40

Funding - Sources of funding and other support; role of funders in data collection, interpretation, and reporting	40

Appendix 4-D. Case Study 1: TREND checklist

Paper Section/Topic	Item No.	Descriptor	Reported?	
			✓	Pg #
TITLE and ABSTRACT				
Title and Abstract	1	Information on how units were allocated to interventions		N/A
		Structured abstract recommended	X	
		Information on target population or study sample	X	
INTRODUCTION				
Background	2	Scientific background and explanation of rationale	X	
		Theories used in designing behavioural interventions		N/A
METHODS				
Participants	3	Eligibility criteria for participants, including criteria at different levels in recruitment/sampling plan (e.g., cities, clinics, subjects)	X	
		Method of recruitment (e.g., referral, self-selection), including the sampling method if a systematic sampling plan was implemented	X	
		Recruitment setting	X	
		Settings and locations where the data were collected	X	
Interventions	4	Details of the interventions intended for each study condition and how and when they were actually administered, specifically including:	X	
		○ Content: what was given?	X	
		○ Delivery method: how was the content given?	X	
		○ Unit of delivery: how were subjects grouped during delivery?		N/A
		○ Deliverer: who delivered the intervention?		N/A
		○ Setting: where was the intervention delivered?	X	
		○ Exposure quantity and duration: how many sessions or episodes or events were intended to be delivered? How long were they intended to last?	X	
		○ Time span: how long was it intended to take to deliver the intervention to each unit?		N/A
		○ Activities to increase compliance or adherence (e.g., incentives)		N/A
Objectives	5	Specific objectives and hypotheses	X	
Outcomes	6	Clearly defined primary and secondary outcome measures	X	
		Methods used to collect data and any methods used to enhance the quality of measurements	X	
		Information on validated instruments such as psychometric and biometric properties	X	
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules	X	

Assignment method	8	Unit of assignment (the unit being assigned to study condition, e.g., individual, group, community)	X	
		Method used to assign units to study conditions, including details of any restriction (e.g., blocking, stratification, minimization)		N/A
		Inclusion of aspects employed to help minimize potential bias induced due to non-randomization (e.g., matching)		N/A
Blinding (masking)	9	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to study condition assignment; if so, statement regarding how the blinding was accomplished and how it was assessed		N/A
Unit of Analysis	10	Description of the smallest unit that is being analysed to assess intervention effects (e.g., individual, group, or community)	X	
		If the unit of analysis differs from the unit of assignment, the analytical method used to account for this (e.g., adjusting the standard error estimates by the design effect or using multilevel analysis)		N/A
Statistical methods	11	Statistical methods used to compare study groups for primary methods outcome(s), including complex methods for correlated data	X	
		Statistical methods used for additional analyses, such as subgroup analyses and adjusted analysis		N/A
		Methods for imputing missing data, if used		N/A
		Statistical software or programs used	X	
RESULTS				
Participant flow	12	Flow of participants through each stage of the study: enrollment, assignment, allocation and intervention exposure, follow-up, analysis (a diagram is strongly recommended)	X	
		○ Enrollment: the numbers of participants screened for eligibility, found to be eligible or not eligible, declined to be enrolled, and enrolled in the study	X	
		○ Assignment: the numbers of participants assigned to a study condition	X	
		○ Allocation and intervention exposure: the number of participants assigned to each study condition and the number of participants who received each intervention	X	
		○ Follow-up: the number of participants who completed the follow-up or did not complete the follow-up (i.e., lost to follow-up), by study condition	X	
		○ Analysis: the number of participants included in or excluded from the main analysis, by study condition	X	
		Description of protocol deviations from study as planned, along with reasons		N/A
Recruitment	13	Dates defining the periods of recruitment and follow-up	X	
Baseline data	14	Baseline demographic and clinical characteristics of participants in each study condition	X	
		Baseline characteristics for each study condition relevant to specific disease prevention research		N/A
		Baseline comparisons of those lost to follow-up and those retained, overall and by study condition		N/A
		Comparison between study population at baseline and target population		N/A

		of interest		
Baseline equivalence	15	Data on study group equivalence at baseline and statistical methods used to control for baseline differences		N/A
Numbers analysed	16	Number of participants (denominator) included in each analysis for each study condition, particularly when the denominators change for different outcomes; statement of the results in absolute numbers when feasible	X	
		Indication of whether the analysis strategy was “intention to treat” or, if not, description of how non-compliers were treated in the analyses		N/A
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each estimation study condition, and the estimated effect size and a confidence interval to indicate the precision	X	6-7
		Inclusion of null and negative findings	X	
		Inclusion of results from testing pre-specified causal pathways through which the intervention was intended to operate, if any		N/A
Ancillary analyses	18	Summary of other analyses performed, including subgroup or restricted analyses, indicating which are pre-specified or exploratory		N/A
Adverse events	19	Summary of all important adverse events or unintended effects in each study condition (including summary measures, effect size estimates, and confidence intervals)		N/A
DISCUSSION				
Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias, imprecision of measures, multiplicative analyses, and other limitations or weaknesses of the study	X	
		Discussion of results taking into account the mechanism by which the intervention was intended to work (causal pathways) or alternative mechanisms or explanations		N/A
		Discussion of the success of and barriers to implementing the intervention, fidelity of implementation	X	
		Discussion of research, programmatic, or policy implications	X	
Generalizability	21	Generalizability (external validity) of the trial findings, taking into account the study population, the characteristics of the intervention, length of follow-up, incentives, compliance rates, specific sites/settings involved in the study, and other contextual issues	X	
Overall evidence	22	General interpretation of the results in the context of current evidence and current theory	X	

Appendix 4-E. Case Study 1: SSI guide and study questionnaires

SSI guide

Introduce interviewer

Before we start, is it okay with you if I audio-record this session?

- The reason we record is so we can go back and remember what was said and what wasn't said. Destroyed as soon as study is done, any publications don't use names
 - *Ask again for confirmation once recording is on*
 - If no: "That's fine, is it okay if I type notes as we talk and share those notes with you once we have finished the interview so you can make sure they are accurate?"

Purpose of our investigation is to evaluate the app - we weren't involved in development, we just want to figure out what is good and can proceed and what needs to improve (any positive / negative feedback is useful). Your feedback is very valuable for us to evaluate the app.

Today we're going to be going through some questions to discuss your experiences using the app. There are six main aspects about the app that you will be asked about, and then you will have the opportunity to provide feedback on anything else you want to talk about that you feel we haven't covered. If there is anything you don't understand during the interview, feel free to ask!

Before we get into those questions, I just wanted to remind you that all of your answers will be kept confidential and stored on a password-protected computer that only the research team can access. The audio recording that we're taking today will be deleted as soon as it has been transcribed, and any identifying information (like your name) will be removed from the audio recording before it is sent to the transcription service.

So before we get started, do you have any questions?

**Indented questions were used as prompts to encourage the participant to provide more detail, when necessary*

1) General usability

- How would you describe your and your family's experience using the 'NoObesity' app?
 - **How much did you like or dislike using the app? Why?*
 - *How easy or difficult did you find the app to use? Why?*
- What has happened as a result of using the app with your family?
 - *Has your behaviour changed? If so, how?*

- *Can you tell me a bit more about how the app helped you and your family support that change?*

2) Self-monitoring and goal-setting

- What do you think of the app function that helps to monitor your family's habits (the 'Record Family Progress' section)?
- Did the app help you set goals? If so, how come?

3) Physical activity and healthy eating support

- Did the app support your family to be more active? If so, how?
- Did the app support your family to eat healthy? If so, how?

4) Weight and health assessment

- What did you think about the app's weight and health measurements (the 'Family Survey' section)?
 - *Did you use it? Did you like or dislike it? Why?*

5) Personalised feedback and motivational strategies

- What did you think of the app's feedback?
 - *This can include any information you were provided by the app in response to something you did on the app*
 - *For example, feedback from the app on health information that you added to the family survey section / getting points in challenges*
- Did the app motivate your family?
 - *What helped or didn't help?*
 - *How come?*

6) Social support and health care expert involvement

- Did you share and collaborate with any others who were also using the app? Eg. any HCPs?
 - *If yes, how?*
 - *What did you like or dislike about that and why?*
 - *If no, why not?*
- Was your health care worker involved in using the app? If so, can you explain how?
 - *Hypothetical (if possible, would that be something you'd use? What would be useful about that for you or a HCP you're working with?)*
 - **Context:** app was developed as a facilitation tool because HEE found it was hard for HCPs to speak with families about weight, and they didn't have good training/tools to talk about in the context of family and children, felt uncomfortable because sensitive topic
 - *What are your reflections on that?*

- *How would you respond to getting feedback from / interacting with doctor through app?*

7) Feedback

- We found there were a lot of strong views (positive and negative) about the name of the app in early feedback during recruitment, what is your view?
 - *Name of the app came from a competition HEE had with children*
- What would you suggest to make the app better?
 - *Can you tell me more about that, and how you think it would improve the app?*
 - *Are there any other suggestions that you think would make the app better?*
- Do you plan to continue using the app after the completion of the study?
- Is there anything else you'd like to mention?

If you have any questions or comments that you want to add after the interview is over, you are welcome to email me or another researcher on the information sheet.

What next:

- Three months from now - send a survey via email, just to get further feedback on your thoughts and experiences
- At the conclusion at that point, we will be analysing feedback from these sessions and surveys to provide recommendations to HEE because they plan to create a new version of the app
- We will publish data in academic journals, but all data will be anonymous
- Also creating a website and public materials, which will be sent to everyone to overview
- Amazon voucher (100 pounds) - everyone who completes study

Do you have any questions?

Demographic Questionnaire

1. AGE: In what year were you born?
 - a. (If refuse to answer): Would you mind indicating which of these age bands you are in: 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84, 85+
2. GENDER: How do you currently describe your gender identity?
3. FAMILY:
 - a. Are you...? Single, Married, Living together with a partner, Divorced/separated, Widowed
 - b. How many adults live in your household (people age 18 or more)? [WRITE IN]
 - c. How many children (people age 17 or less) live in your household? [WRITE IN]
4. ETHNICITY: To which of these groups do you consider you belong?
 - a. ASIAN: of Indian origin
 - b. ASIAN: of Pakistani origin
 - c. ASIAN: of Bangladeshi origin
 - d. ASIAN: of Chinese origin
 - e. ASIAN: of any other origin (WRITE IN) _____
 - f. BLACK: of African origin
 - g. BLACK: of Caribbean origin
 - h. BLACK: of other origin (WRITE IN) _____
 - i. WHITE: of British origin
 - j. WHITE: of Scottish origin
 - k. WHITE: of Welsh origin
 - l. WHITE: of any other origin (WRITE IN) _____
 - m. OTHER (WRITE IN)
 - n. Don't know
5. URBAN: Would you describe the place where you live as a big city, the suburbs or outskirts of a big city, a small city or town, a country village, or, a farm or home in the country?
 - a. A big city
 - b. The suburbs or outskirts of a big city
 - c. A small city or town
 - d. A country village
 - e. A farm or home in the country

- f. Other
 - g. Don't know
6. INCOME: The incomes of households differ a lot in Britain today. Here is a table showing the range of incomes that people have. Which option best represents the total income of your household before tax? Please select one answer only.
- a. Up to £12.500 per year
 - b. £12.500 up to £20.000 per year
 - c. £20.000 up to £30.000 per year
 - d. £30.000 up to £40.000 per year
 - e. £40.000 up to £50.000 per year
 - f. £50.000 up to £60.000 per year
 - g. £60.000 up to £70.000 per year
 - h. £70.000 up to £80.000 per year
 - i. Over £80.000 per year
 - j. Prefer not to say
7. EDUCATION: What is the highest educational or vocational qualification that you have or that you will receive if you complete your next set of exams?
- a. No qualifications
 - b. 5 or more GCSE grades A-C
 - c. 4 or less GCSE grade A-C
 - d. GCSE grade D-G
 - e. 5 or more Scottish Standard Grades, grades 1-3
 - f. 4 or less Scottish Standard Grades, grades 1-3
 - g. 5 or more O Grades, grades 1-3 (Scottish Ordinary Grades)
 - h. 4 or less O Grades, grades 1-3 (Scottish Ordinary Grades)
 - i. Scottish Standard Grades, grades 4-7
 - j. Scottish Highers (either SCE or SQC)
 - k. Scottish Certificate Sixth Year Studies
 - l. SVQ level 1 or 2 (Scottish Vocational Qualifications)
 - m. SVQ level 3 (“)
 - n. SVQ level 4 (“)
 - o. SVQ level 5 (“)
 - p. CSEs
 - q. 5 or more O levels

- r. 4 or less O levels
- s. GCE A levels or equivalent
- t. NVQ level 1 or 2
- u. NVQ level 3 or 4
- v. NVQ level 5
- w. GNVQ Foundation
- x. GNVQ Intermediate
- y. GNVQ Advanced
- z. Certificate or Diploma of Higher Education
- aa. HND (Higher National Diploma)
- bb. Bachelor's degree
- cc. Graduate Certificates and Diploma
- dd. Post-degree professional qualification (eg. banking accountancy, architecture)
- ee. Master's Degree
- ff. Doctoral Degree
- gg. Don't know

8. EMPLOYMENT: Which of these descriptions best describes your current situation?

- a. Working Full time (30 hours a week or more)
- b. Working Part time (8-29 hours a week)
 - i. If (a) or (b): Apart from working, do you also study? (yes / no / DK)
- c. Retired
- d. Unemployed
- e. Permanently sick or disabled
- f. In community or military service
- g. Undergraduate Student
- h. Post graduate student
- i. In full time education (not higher degree)
- j. In part time education (not higher degree)
 - i. If (g)-(j): And apart from studying, do you also work? (yes / no / DK)
- k. Doing housework, looking after children or other persons

Final survey

System Usability Scale (SUS)

The 10 SUS statements below were presented twice, using a 5-point Likert scale (1=strongly disagree and 5=strongly agree) with the following instructions: (1) “Please rate your agreement with the following statements from your experience using the NoObesity app” and (2) “Please rate the following statements from your child(ren)'s perspective of using the NoObesity app.”

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

Impact of app

**Measured with a 5 point Likert scale*

1. Did the NoObesity app help you set goals?
 - a. *Strongly disagree - strongly agree
2. Did you achieve your goals?
 - a. *Strongly disagree - strongly agree
3. What did you think about the monitoring function (Record Family Progress)?
 - a. *Not at all useful - extremely useful
4. Did the NoObesity app support your family to be more active?
 - a. *Strongly disagree - strongly agree
5. Did the NoObesity app support your family to eat healthier?
 - a. *Strongly disagree - strongly agree
6. Please think about up to 3 goals you set on NoObesity and rate for each how confident you are that you can do those activities as of now. If you did not set 3 goals, please select 'Not Applicable' for Goals 2 and/or 3. Rate your degree of confidence by recording a number from 0 to 100 on the scale below (0=cannot do at all, 50=moderately certain can do, 100=highly certain can do). (*Self-efficacy*)

7. Did you set up the initial family profile (Family Survey section)?
 - a. Yes / no
8. If yes, how frequently did you go back to look at the Family Survey section?
 - a. Never / once or twice during study / 1-4 times a month / 2-6 times a week / daily
9. What did you think about the HC points / trophies?
 - a. *Dislike a great deal - like a great deal
10. What did you think of the information the app provided as feedback?
 - a. *Not at all useful - extremely useful
11. How effective was the NoObesity app at motivating your family to be more active?
 - a. *Not at all effective - extremely effective
12. How effective was the NoObesity app at motivating your family to eat healthier?
 - a. *Not at all effective - extremely effective
13. Please rate the following features on how effective they were at motivating your family (*Not at all effective - extremely effective):
 - a. goal-setting
 - b. Notifications
 - c. Monitoring behaviour (tracking progress)
 - d. Points / trophies
 - e. Family photo
 - f. Feedback from Family Survey about current behaviour
 - g. Games
 - h. Suggestions for healthier eating
 - i. Suggestions of activities
 - j. Linking to a healthcare worker
 - k. Doing something together as a family
 - l. Other (please specify)
14. How likely would you be to link this app with your GP or another healthcare provider?
 - a. *Extremely unlikely - extremely likely
15. How comfortable would you be interacting with a doctor on this app?
 - a. *Extremely uncomfortable - extremely comfortable
16. How useful do you think it would be to interact with a doctor on this app?
 - a. *Not at all useful - extremely useful

Bandura - self-efficacy questions

1. Please briefly list any barriers to your use of the app.
2. Please briefly list any barriers to your efforts to meet your goal(s).

RE-AIM related questions

1. What did you think about the name of the app (“NoObesity”)?
 - a. *Dislike a great deal - like a great deal
2. If you were seeing NoObesity in the app store for the first time, how likely would you be to download it?
 - a. *Extremely unlikely - extremely likely
3. Are you still using the NoObesity app?
 - a. Yes / no
4. If no, how long (approximately) did you use the NoObesity app for before stopping?
5. If no, why did you stop using the app?

General feedback

1. Please briefly list any suggestions you have to make the app better:
2. Do you have any other feedback about the app you want to share?

Appendix 4-F. Case Study 1: COM-B and TDF coding framework

COM-B	Definition [170]	TDF	Definition [373]	Examples (engagement with mHealth) [184,201]
Physical capability	Physical skill, strength or stamina	Skills	An ability or proficiency acquired through practice	App literacy: Technological competency
Psychological capability	Knowledge or psychological skills, strength or stamina to engage in the necessary mental processes	Knowledge	An awareness of the existence of something	- App awareness: Knowledge of the existence of health and well-being apps - User guidance: Instructions on how to effectively use the app - Health information: Educational information related to health and well-being aspects - Statistical information: A visual or numerical summary of progress
		Memory, attention, and decision processes	The ability to retain information, focus selectively on aspects of the environment and choose between two or more alternatives	- Well-designed reminders: The ability to customize reminders - Less cognitive load: The app is not too time consuming, easy to use, and requires minimal input - Reminders: Preferably customizable notification-type messages
		Behavioural regulation	Anything aimed at managing or changing objectively observed or measured actions	- Self-monitoring: The ability of the app to help self-regulation of the target behaviour - Established routines: Regularity in using the app - Routines: The ability to support routine or habit formation - Safety netting: Retaining the app for a potential upcoming event in the future - Stepping-stone: App as a first step in the behaviour change process
Physical opportunity	Opportunity afforded by the environment involving time, resources, locations, cues,	Environmental context and resources	Any circumstance of a person's situation or environment that discourages or encourages	- Availability and accessibility: The ability to use a smartphone anytime anywhere - Low cost: the price of the app - Interactive and positive tone: Encouraging communication style

	physical 'affordance'		the development of skills and abilities, independence, social competence and adaptive behaviour	● Personalization to needs: The possibility to use an app that is tailored to a user's needs ● Tailoring: Innovative features, adaptability, and interactive and 2-way communication between the app and user
Social opportunity	Opportunity afforded by interpersonal influences, social cues and cultural norms that influence the way that we think about things, e.g. the words and concepts that make up our language	Social influences	Those interpersonal processes that can cause individuals to change their thoughts, feelings, or behaviours	● Peer support: Including social interaction with users with similar needs within the app or within their community; a choice to connect to social media platforms, competitions, and challenges with others or with themselves ● Social support (practical): Possibility to contact health professionals and practitioners within the app ● Recommendations: Suggestions received from other users ● Health practitioner support: Possibility to get in touch with health professionals and practitioners within the app ● Community networking: Social interaction with users with similar needs within the app or within their community ● Social media: A choice to connect to social media platforms ● Social competition: Competitive nature of the app with others or with themselves ● Personification of the app: Applying human attributes to the app
Reflective motivation	Reflective processes involving plans (self-conscious intentions) and evaluations (beliefs about what is good and bad)	Social / professional role and identity	A coherent set of behaviours and displayed personal qualities of an individual in a social or work setting	
		Beliefs about capabilities	Acceptance of the truth, reality or validity about an ability, talent or facility that a person can put to constructive	● Self-confidence: Perceived capability to change one's behaviour using an app

			use	
		Optimism	The confidence that things will happen for the best or that desired goals will be attained	
		Beliefs about consequences	Acceptance of the truth, reality, or validity about outcomes of a behaviour in a given situation	● Perceived utility of the app: Discrepancy of what the users are looking for and what the app offers ● Commitment: The level of commitment while engaging with an app to change the behaviour and achieve set goals
		Intentions	A conscious decision to perform a behaviour or a resolve to act in a certain way	
		Goals	Mental representations of outcomes or end states that an individual wants to achieve	● Goal setting: Establishing what the user would like to accomplish ● Action planning: Establishing how the user would like to achieve set goals
Automatic motivation	Automatic processes involving emotional reactions, desires (wants and needs), impulses, inhibitions, drive states and reflex responses	Reinforcement	Increasing the probability of a response by arranging a dependent relationship, or contingency, between the response and a given stimulus	● Feedback: Feedback regarding the user's performance ● Rewards: Tangible (eg, objects and discounts) and intangible (eg, badges and certificates) rewards in response to the user's effort and gamification elements ● Encouragement: Additional ways to provide reinforcement (eg, encouraging messages)
		Emotion	A complex reaction pattern, involving experiential, behavioural, and physiological elements, by which the individual attempts to deal	● Curiosity: Desire to acquire knowledge and skills to use a behaviour change tool ● Positive emotions: Triggered by the included user guidance, statistical information, additional health information, embedded professional support, community networking possibilities, tracking features, and rewards

			with a personally significant matter or event	● Negative emotions: Triggered by lack of user guidance, invasive push notifications, cognitive overload, and unrevealed in-app costs ● Mixed emotions: Triggered by reminders

Appendix 4-G. Case Study 2: CONSORT-AI Checklist [375]

Section	Item	CONSORT 2010 Item ^a	CONSORT-AI Item		Page No
Title and Abstract					
Title and Abstract	1a	Identification as a randomised trial in the title	CONSORT-AI 1a,b Elaboration	(i) Indicate that the intervention involves artificial intelligence/machine learning in the title and/or abstract and specify the type of model.	Title page
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)		(ii) State the intended use of the AI intervention within the trial in the title and/or abstract.	Title, abstract
Introduction					
Background and objectives	2a	Scientific background and explanation of rationale	CONSORT-AI 2a (i) Extension	Explain the intended use of the AI intervention in the context of the clinical pathway, including its purpose and its intended users (e.g. healthcare professionals, patients, public).	3
	2b	Specific objectives or hypotheses			3
Methods					
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio			4
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons			6-7
Participants	4a	Eligibility criteria for participants	CONSORT-AI 4a (i) Elaboration	State the inclusion and exclusion criteria at the level of participants.	4-5
			CONSORT-AI 4a (ii) Extension	State the inclusion and exclusion criteria at the level of the input data.	N/A
	4b	Settings and locations where the data were collected	CONSORT-AI 4b Extension	Describe how the AI intervention was integrated into the trial setting, including any onsite or offsite requirements.	5
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	CONSORT-AI 5 (i) Extension	State which version of the AI algorithm was used.	5
			CONSORT-AI 5 (ii) Extension	Describe how the input data were acquired and selected for the AI intervention.	5

			CONSORT-AI 5 (iii) Extension	Describe how poor quality or unavailable input data were assessed and handled.	N/A
			CONSORT-AI 5 (iv) Extension.	Specify whether there was human-AI interaction in the handling of the input data, and what level of expertise was required of users.	5
			CONSORT-AI 5 (v) Extension	Specify the output of the AI intervention	5
			CONSORT-AI 5 (vi) Extension	Explain how the AI intervention's outputs contributed to decision-making or other elements of clinical practice.	N/A
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed			5
	6b	Any changes to trial outcomes after the trial commenced, with reasons			6-7, 18
Sample size	7a	How sample size was determined			6
	7b	When applicable, explanation of any interim analyses and stopping guidelines			N/A
Randomisation					
Sequence generation	8a	Method used to generate the random allocation sequence			5
	8b	Type of randomisation; details of any restriction (such as blocking and block size)			5
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned			5
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions			5
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how			5
	11b	If relevant, description of the similarity of interventions			N/A
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes			6
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses			6

Results					
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome			7, Figure 2
	13b	For each group, losses and exclusions after randomisation, together with reasons			Figure 2
Recruitment	14a	Dates defining the periods of recruitment and follow-up			5,7
	14b	Why the trial ended or was stopped			18
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group			8, Table 3
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups			Table 3
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)			10-13
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended			N/A
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory			13-16
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	CONSORT-AI 19 Extension	Describe results of any analysis of performance errors and how errors were identified, where applicable. If no such analysis was planned or done, explain why not.	17, N/A
Discussion					
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses			17-18
Generalisability	21	Generalisability (external validity, applicability) of the trial findings			18
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence			17-19
Other Information					
Registration	23	Registration number and name of trial registry			1
Protocol	24	Where the full trial protocol can be accessed, if available			N/A

Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	CONSORT-AI 25 Extension.	State whether and how the AI intervention and/or its code can be accessed, including any restrictions to access or re-use.	20

Appendix 4-H. Case Study 2: SSI topic guide

Hello ____, it's nice to meet you. I really appreciate you taking the time to be involved in this evaluation, it will be really helpful. My name is Madison Milne-Ives and I am a researcher at the Centre for Health Technology at the University of Plymouth who is working on the Wysa evaluation project.

Purpose of our investigation is to evaluate the Wysa app - we weren't involved in development, we just want to figure out if it would be a useful tool to bring into the NHS mental health service provision in addition to standard care. Your feedback is very valuable for us.

I just want to remind you that all of your answers will be kept confidential and stored on a password-protected computer that only we can access. My questions are about your experience using the Wysa app, not about your mental health, but if there are any questions you don't want to answer, that's completely fine, just let me know and we'll move on. We can also stop at any point you want and you don't need to give us a reason.

Before we start, is it okay with you if I audio-record this session?

- The reason we record is so we can go back and remember what was said and what wasn't said. The audio recording that we're taking today will be deleted as soon as it has been transcribed, and any identifying information (like your name) will be removed.
- NOTE: ask again for confirmation once recording is on
- If no: "That's fine, is it okay if I type notes as we talk and share those notes with you once we have finished the interview so you can make sure they are accurate?"

Today we're going to be going through some questions to discuss your experiences using the app. If there is anything you don't understand during the interview, feel free to ask!

Before we get started, do you have any questions?

Questions

1. Thinking back to when you first got the app, can you tell me about what you felt and thought about using it? (*TFA - affective attitude, engagement - affective, cognitive*)
2. What motivated you to use the app? Why did you want to try it? (*engagement - affective*)
 - a. What did you hope to gain / how did you think it would benefit you? (*TFA - perceived effectiveness, anticipated*)
3. How much effort did it take to use Wysa? What made it easy or difficult to use? (*TFA - burden*)
 - a. How confident were you in your ability to use the app? (*TFA - self-efficacy*)
 - b. How well did you understand how Wysa worked? (*TFA - intervention coherence*)

4. How did using the app - or specifically the general talk feature or the self-care exercise packs - make you feel? (*engagement - micro, affective*)
 - a. What aspects of the app did you find most and least enjoyable? Why?
5. What did you think about the app's chat function and exercise pack content? (*engagement - micro, cognitive*)
 - a. What did you find most interesting / attention catching? Why?
 - b. What was most boring? Why?
 - c. What affected your interest or the attention you paid to the app?
6. Did you find the program and resources credible and trustworthy? What elements of the program fostered that trust/distrust for you? (*engagement - affective (social trust), cognitive*)
7. What concerns, if any, did you have about using Wysa? (*TFA - ethicality*)
 - a. To what degree did the app's features and content align with your values and thoughts about mental health support?
8. How much confidence did you have in Wysa's ability to provide mental health support? (*TFA - perceived effectiveness*)
9. What, if anything, did you have to give up to engage with Wysa? Were there any costs or drawbacks to using the app? (*TFA - opportunity costs*)
10. Do you think the intervention influenced you beyond when you were using the app? If so, how? (*engagement - behavioural, TFA - perceived effectiveness*)
 - a. For example, having more skills and strategies or feeling more capable of managing stress, sleeping better, etc.
11. Can you describe how you incorporated the app into your daily life over the past 3 months? (*engagement - micro, behavioural*)
 - a. How did you use the app? What features did you use the most / least? Why?
12. How did your use of the app change over time? (*engagement - micro*)
 - a. In what circumstances were you more likely to engage with the app? Ie. Why did you log on when you did? (*engagement - micro*)
 - b. What made you want to open up the app again?
 - c. Why did you continue to use the app or discontinue use of the app?
 - d. If you stopped using the app, why would you / did you start using it again?
13. Do you have any other feedback you'd like to provide that you don't feel like we have discussed?

Thank you very much for taking the time to participate, it's been really helpful to hear more about your experience and we really appreciate it. Do you have any questions for me at this stage?

Appendix 4-I. Case Study 2: Initial codebook used in preliminary qualitative coding

Code	Definition
<i>Engagement components</i>	
Affective engagement	Motivation, positive and negative feelings (e.g. enjoyment), sensory pleasure of UX [42,45,51]
Cognitive engagement	Challenge, interest, thought, attention [42,45,51]
Behavioural engagement	Interactivity, usage (amount, duration, frequency, etc.) [42,45,51]
Micro engagement	Engagement with the mobile health app itself (UX / BCTs) [14]
Macro engagement	“Engagement and identification with the wider intervention goals” [14], engagement with the targeted health behaviour [50]
<i>Acceptability constructs (TFA)</i>	
Affective attitude	How an individual feels about the intervention
Burden	The perceived amount of effort that is required to participate in the intervention
Ethicality	The extent to which the intervention has good fit with an individual's value system
Intervention coherence	The extent to which the participant understands the intervention and how it works
Opportunity costs	The extent to which benefits, profits, or values must be given up to engage in the intervention
Perceived effectiveness	The extent to which the intervention is perceived to be likely to achieve its purpose
Self-efficacy	The participant's confidence that they can perform the behaviour(s) required to participate in the intervention
Other	Utterances that answer the research question but do not necessarily reflect the TFA constructs

Appendix 4-J. Case Study 3: DECIDE-AI reporting guideline [382]

Item #	Theme	Recommendation	Page #
1–17	AI-specific reporting items		
1–X	Generic reporting items		
Title and abstract			
1	Title	Identify the study as early clinical evaluation of a decision support system based on AI or machine learning, specifying the problem addressed.	Title page
I	Abstract	Provide a structured summary of the study. Consider including: intended use of the AI system, type of underlying algorithm, study setting, number of patients and users included, primary and secondary outcomes, key safety endpoints, human factors evaluated, main results and conclusions.	1
Introduction			
2	Intended use	a) Describe the targeted medical condition(s) and problem(s), including the current standard practice, and the intended patient population(s).	3
		b) Describe the intended users of the AI system, its planned integration in the care pathway, and the potential effect, including patient outcomes, that it is intended to have.	3-5
II	Objectives	State the study objectives.	3
Methods			
III	Research governance	Provide a reference to any study protocol, study registration number, and ethics approval.	3-4, 6
3	Participants	a) Describe how patients were recruited, stating the inclusion and exclusion criteria at both patient and data level, and how the number of recruited patients was decided.	4-5
		b) Describe how users were recruited, stating the inclusion and exclusion criteria, and how the intended number of recruited users was decided.	4-5
		c) Describe steps taken to familiarize the users with the AI system, including any training received before the study.	3, 5-6

4	AI system	a) Briefly describe the AI system, specifying its version and type of underlying algorithm used. Describe, or provide a direct reference to, the characteristics of the patient population on which the algorithm was trained and its performance in preclinical development/validation studies.	3
		b) Identify the data used as inputs. Describe how the data were acquired, the process needed to enter the input data, the pre-processing applied, and how missing/low-quality data were handled.	Fig 1
		c) Describe the AI system outputs and how they were presented to the users (an image may be useful).	Fig 1
5	Implementat ion	a) Describe the settings in which the AI system was evaluated.	4
		b) Describe the clinical workflow/care pathway in which the AI system was evaluated, the timing of its use, and how the final supported decision was reached and by whom.	5
IV	Outcomes	Specify the primary and secondary outcomes measured.	6
6	Safety and errors	a) Provide a description of how significant errors/malfunctions were defined and identified.	5-6
		b) Describe how any risks to patient safety or instances of harm were identified, analyzed, and minimized.	5-6
7	Human factors	Describe the human factors tools, methods or frameworks used, the use cases considered, and the users involved.	N/A
V	Analysis	Describe the statistical methods by which the primary and secondary outcomes were analyzed, as well as any pre-specified additional analyses, including subgroup analyses and their rationale.	6
8	Ethics	Describe whether specific methodologies were used to fulfil an ethics-related goal (such as algorithmic fairness) and their rationale.	N/A
VI	Patient Involvement	State how patients were involved in any aspect of: the development of the research question, the study design, and the conduct of the study.	N/A
Results			
9	Participants	a) Describe the baseline characteristics of the patients included in the study and report on input data missingness.	7-9
		b) Describe the baseline characteristics of the users included in the study.	7-9
10	Implementat ion	a) Report on the user exposure to the AI system, on the number of instances the AI system was used, and on the users' adherence to the intended implementation.	Fig 3
		b) Report any significant changes to the clinical workflow or care pathway caused by the AI system.	N/A
VII	Main results	Report on the pre-specified outcomes, including outcomes for any comparison group if applicable.	9-15

VIII	Subgroups analysis	Report on the differences in the main outcomes according to the pre-specified subgroups.	13-14
11	Modifications	Report any changes made to the AI system or its hardware platform during the study. Report the timing of these modifications, the rationale for each, and any changes in outcomes observed after each of them.	16
12	Human–computer agreement	Report on the user agreement with the AI system. Describe any instances of and reasons for user variation from the AI system’s recommendations and, if applicable, users changing their mind based on the AI system’s recommendations.	9-10
13	Safety and errors	a) List any significant errors/malfunctions related to: AI system recommendations, supporting software/hardware, or users. Include details of: (i) rate of occurrence, (ii) apparent causes, (iii) whether they could be corrected, and (iv) any significant potential effects on patient care.	10-11
		b) Report on any risks to patient safety or observed instances of harm (including indirect harm) identified during the study.	10-11
14	Human factors	a) Report on the usability evaluation, according to recognized standards or frameworks.	11-13
		b) Report on the user learning curves evaluation.	N/A
Discussion			
15	Support for intended use	Discuss whether the results obtained support the intended use of the AI system in clinical settings.	15-16
16	Safety and errors	Discuss what the results indicate about the safety profile of the AI system. Discuss any observed errors/malfunctions and instances of harm, their implications for patient care, and whether/how they can be mitigated.	15
IX	Strengths and limitations	Discuss the strengths and limitations of the study.	16-17
Statements			
17	Data availability	Disclose if and how data and relevant code are available.	20
X	Conflicts of interest	Disclose any relevant conflicts of interest, including the source of funding for the study, the role of funders, any other roles played by commercial companies, and personal conflicts of interest for each author.	19-20

AI-specific items are numbered in Arabic numerals; generic items are numbered in Roman numerals.

Appendix 4-K. Case Study 3: SRQR checklist [362]

Page no(s).

Title and abstract

Title - Concise description of the nature and topic of the study Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended	N/A
Abstract - Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	1

Introduction

Problem formulation - Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	2
Purpose or research question - Purpose of the study and specific objectives or questions	2

Methods

Qualitative approach and research paradigm - Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/ interpretivist) is also recommended; rationale**	3
Researcher characteristics and reflexivity - Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	6
Context - Setting/site and salient contextual factors; rationale**	3-5
Sampling strategy - How and why research participants, documents, or events were selected; criteria for deciding when no	5

further sampling was necessary (e.g., sampling saturation); rationale**	
Ethical issues pertaining to human subjects - Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	3
Data collection methods - Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale**	5
Data collection instruments and technologies - Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	5
Units of study - Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	6
Data processing - Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/de-identification of excerpts	6
Data analysis - Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale**	6
Techniques to enhance trustworthiness - Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale**	6

Results/findings

Synthesis and interpretation - Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	12-14

Links to empirical data - Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	12-14

Discussion

Integration with prior work, implications, transferability, and contribution(s) to the field - Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/generalizability; identification of unique contribution(s) to scholarship in a discipline or field	17
Limitations - Trustworthiness and limitations of findings	17-18

Other

Conflicts of interest - Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	20-21
Funding - Sources of funding and other support; role of funders in data collection, interpretation, and reporting	20

Appendix 4-L. Case Study 3: Dora's symptom assessments with examples of significant and insignificant symptoms

Symptom assessed	Examples of significant symptoms	Examples of insignificant symptoms
Eye Redness	Diffuse redness Progressively worsening redness.	White eye If the eye is only red in the corner. Occasional redness. Redness that has resolved.
Eye Pain	Constant or worsening pain Deep or severe pain.	Pain that is described as gritty or dry. Pain that has now resolved. Mild pain only when applying eye drops.
Vision Concerns	Worsening vision Concerns about vision Inability to read a number plate from 20 metres Distorted vision	Happy with their vision Blurred vision but only for reading, with good distance vision.
Flashing lights in vision	Any flashing lights in vision.	No flashing lights in vision
Floaters in vision	Any new floaters in vision.	Floaters that have been present since before the cataract operation (no new floaters)

Appendix 4-M. Case Study 3: Dora and clinician decision matrix for the patient care management outcome conveyed at the end of the follow-up assessment

		Clinician decision		
		Pass ^a	Fail ^b	Insufficient information
Dora decision	Pass ^a	No clinician callback	Clinician callback	Interrupt call/ call back
	Fail ^b	No clinician callback	Clinician callback	Interrupt call / call back

^aNo further clinician review needed

^bFurther clinician review needed to determine management

Appendix 4-N. Case Study 3: Criteria for a patient to ‘fail’ the Dora call

A positive response, or failure to understand the patient response to:

- Significant redness, and/or
- Significant pain, and/or
- Significant visual issues and/or
- Significant flashing lights and/or
- Significant floaters and/or

AND/OR

A failure to answer any patient questions to their satisfaction

AND/OR

The patient requested a human clinician review

OR

The Dora call did not complete^a

^aCall completion was defined as the conversation progressing to completing all clinically relevant parts including all symptom assessments and second eye preference modules.

Appendix 4-O. Case Study 3: post-call survey

Blank, Grant, OxIS 2019 Questionnaire. All Parts (January 30, 2019). Available at SSRN: <https://ssrn.com/abstract=3522118>

Brooke J. Systems Usability Scale. In: Jordan P, Thomas B, Weerdmeester W, editors. Usability Evaluation In Industry Taylor & Francis; 1996. p. 189–194.

Parmanto B, Lewis AN Jr, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). *Int J Telerehabil* 2016 Jul 1;8(1):3–10. PMID:27563386

Demographic Questions

1. AGE: In what year were you born?
 - a. (If refuse to answer): Would you mind indicating which of these age bands you are in: 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84, 85+
2. GENDER*: How do you currently describe your gender identity?
 - a. *Note: in the OxIS, gender was assessed ‘by observation’. The wording of this question is therefore taken from a different source (Hughes, Camden, & Yangchen, 2016)
3. FAMILY:
 - a. Are you...? Single, Married, Living together with a partner, Divorced/separated, Widowed
 - b. How many adults live in your household (people age 18 or more)? [WRITE IN]
 - c. How many children (people age 17 or less) live in your household? [WRITE IN]
4. ETHNICITY: To which of these groups do you consider you belong?
 - a. ASIAN: of Indian origin
 - b. ASIAN: of Pakistani origin
 - c. ASIAN: of Bangladeshi origin
 - d. ASIAN: of Chinese origin
 - e. ASIAN: of any other origin (WRITE IN) _____
 - f. BLACK: of African origin
 - g. BLACK: of Caribbean origin
 - h. BLACK: of other origin (WRITE IN) _____
 - i. WHITE: of British origin
 - j. WHITE: of Scottish origin
 - k. WHITE: of Welsh origin

- l. WHITE: of any other origin (WRITE IN) _____
 - m. OTHER (WRITE IN)
 - n. Don't know
5. URBAN: Would you describe the place where you live as a big city, the suburbs or outskirts of a big city, a small city or town, a country village, or, a farm or home in the country?
- a. A big city
 - b. The suburbs or outskirts of a big city
 - c. A small city or town
 - d. A country village
 - e. A farm or home in the country
 - f. Other
 - g. Don't know
6. INCOME: The incomes of households differ a lot in Britain today. Here is a table showing the range of incomes that people have. Which option best represents the total income of your household before tax? Please select one answer only.
- a. Up to £12.500 per year
 - b. £12.500 up to £20.000 per year
 - c. £20.000 up to £30.000 per year
 - d. £30.000 up to £40.000 per year
 - e. £40.000 up to £50.000 per year
 - f. £50.000 up to £60.000 per year
 - g. £60.000 up to £70.000 per year
 - h. £70.000 up to £80.000 per year
 - i. Over £80.000 per year
 - j. Prefer not to say
7. EDUCATION: What is the highest educational or vocational qualification that you have or that you will receive if you complete your next set of exams?
- a. No qualifications
 - b. 5 or more GCSE grades A-C
 - c. 4 or less GCSE grade A-C
 - d. GCSE grade D-G

- e. 5 or more Scottish Standard Grades, grades 1-3
- f. 4 or less Scottish Standard Grades, grades 1-3
- g. 5 or more O Grades, grades 1-3 (Scottish Ordinary Grades)
- h. 4 or less O Grades, grades 1-3 (Scottish Ordinary Grades)
- i. Scottish Standard Grades, grades 4-7
- j. Scottish Highers (either SCE or SQC)
- k. Scottish Certificate Sixth Year Studies
- l. SVQ level 1 or 2 (Scottish Vocational Qualifications)
- m. SVQ level 3 (“)
- n. SVQ level 4 (“)
- o. SVQ level 5 (“)
- p. CSEs
- q. 5 or more O levels
- r. 4 or less O levels
- s. GCE A levels or equivalent
- t. NVQ level 1 or 2
- u. NVQ level 3 or 4
- v. NVQ level 5
- w. GNVQ Foundation
- x. GNVQ Intermediate
- y. GNVQ Advanced
- z. Certificate or Diploma of Higher Education
- aa. HND (Higher National Diploma)
- bb. Bachelor's degree
- cc. Graduate Certificates and Diploma
- dd. Post-degree professional qualification (eg banking accountancy, architecture, etc.)
- ee. Master's Degree
- ff. Doctoral Degree
- gg. Don't know

8. EMPLOYMENT: Which of these descriptions best describes your current situation?
- a. Working Full time (30 hours a week or more)
 - b. Working Part time (8-29 hours a week)
 - i. If (a) or (b): Apart from working, do you also study? (yes / no / DK)
 - c. Retired
 - d. Unemployed
 - e. Permanently sick or disabled
 - f. In community or military service
 - g. Undergraduate Student
 - h. Post graduate student
 - i. In full time education (not higher degree)
 - j. In part time education (not higher degree)
 - i. If (g)-(j): And apart from studying, do you also work? (yes / no / DK)
 - k. Doing housework, looking after children or other persons

System Usability Scale

Rated on a 5-point Likert scale, from 'Strongly Disagree' (1) to 'Strongly Agree' (5)

- 9. I think that I would like to use this system frequently.
- 10. I found the system unnecessarily complex.
- 11. I thought the system was easy to use.
- 12. I think that I would need the support of a technical person to be able to use this system.
- 13. I found the various functions in this system were well integrated.
- 14. I thought there was too much inconsistency in this system.
- 15. I would imagine that most people would learn to use this system very quickly.
- 16. I found the system very awkward to use.
- 17. I felt very confident using the system.
- 18. I needed to learn a lot of things before I could get going with this system.

Telehealth Usability Questionnaire

Rated on a 7-point Likert scale, from 'Disagree' (1) to 'Agree' (7). It has been adapted slightly to fit the particular context of this study.

19. Telehealth improves my access to healthcare services.
20. Telehealth saves me time traveling to a hospital or specialist clinic.
21. Telehealth provides for my healthcare needs.
22. It was simple to use this system.
23. It was easy to learn to use the system.
24. The way I interact with this system is pleasant.
25. I like using the system.
26. The system is simple and easy to understand.
27. This system is able to do everything I would want it to be able to do.
28. I could easily talk to **DORA**.
29. I could hear **DORA** clearly.
30. I felt I was able to express myself effectively.
31. I think the visits provided over the telehealth system are the same as in-person visits.
32. Whenever I made a mistake using the system, I could recover easily and quickly.
33. The system gave error messages that clearly told me how to fix problems.
34. I feel comfortable communicating with **DORA**.
35. Telehealth is an acceptable way to receive healthcare services.
36. I would use telehealth services again.
37. Overall, I am satisfied with this telehealth system.

Appendix 4-P. Case Study 3: SSI guide

Hello, this is Madison, calling from the University of Plymouth; we had an interview scheduled about the 'Dora' automated follow-up phone call you received, is now still an okay time? We really appreciate you taking the time to be involved in this evaluation, it will be really helpful.

The purpose of our study is to get a better understanding of your experience and how you felt about talking to Dora on the phone after your cataract surgery. Your feedback is very valuable for us to evaluate the Dora technology - there are no right or wrong answers. This will probably take around half an hour, is that okay?

Before we start, is it okay with you if I audio-record this session?

- The reason we record is so we can go back and remember what was said and what wasn't said. Destroyed as soon as study is done, any publications don't use names
- *NOTE: ask again for confirmation once recording is on*
- If no: "That's fine, is it okay if I type notes as we talk and share those notes with you once we have finished the interview so you can make sure they are accurate?"

Before we get started, I just wanted to remind you that all of your answers will be kept confidential and protected - we're not from the company, so we are happy with all feedback.

So before we get started, do you have any questions?

Prompts / probes:

- 'Can you tell me more about that?'
- 'Can you tell me about ____ in more detail?'
- 'What happened then?'
- 'Uh huh?', 'Yes?', 'Go on'
- 'Can you give me an example of . . . ?'
- 'What do you mean by...?'
- 'In what way?'
- 'Why do you say that?'

- ‘Can I check, have I understood this correctly?’

Interview Questions (based on TFA constructs)

1. What did you expect talking to DORA to be like before you received the call?
 - a. How much experience have you had talking to automated calls before?
2. How did you feel about the experience of talking to Dora? (*affective attitude*)
 - a. How did you feel about the experience after talking to Dora? Any different?
3. What aspects of talking to Dora did you like or dislike?
4. How much effort did it take for you to interact with Dora? (*burden*)
 - a. What, if anything, made it difficult for you to use Dora?
 - b. How did things go with the technological aspects (sound, connection) of the call? How did that affect your experience?
5. What concerns, if any, did you have about talking to Dora? (*ethicality*)
6. Can you tell me about your understanding of how Dora works? (*intervention coherence*)
7. How well do you think Dora would work if used more extensively? (*perceived effectiveness*)
 - a. What is your opinion on the suitability of technology like Dora doing surgery follow-up in the long run?
8. How much confidence did you have in DORA’s ability to do the follow-up assessment well? (*perceived effectiveness*)
 - a. Before / after call?
9. How comfortable did you feel talking to Dora? (*self-efficacy*)
 - a. How able did you feel to communicate your experience / symptoms to Dora?
10. What, if anything, did you gain or lose by doing a follow-up assessment with Dora? (*opportunity costs*)
 - a. Are there things you feel you are giving up by not coming in person for a follow-up assessment?

b. What, if anything, do you feel you gained from using Dora?

11. Do you think that Dora could help you / other patients? If no, why? / If yes, how?

12. Do you think that Dora could harm you / other patients? If no, why? / If yes, how?

13. Do you think that Dora could help / harm society as a whole? If no, why? / If yes, how?

*** General feedback**

1. How willing would you be to use DORA again?
2. If you had the choice between no follow-up, a face-to-face appointment, or DORA, which would you choose and why?
3. Do you have any suggestions for how Dora could be improved?
4. Is there anything else you'd like to mention?

Thank you so much, very helpful. What's going to happen next:

- We will be analysing the feedback from these sessions and surveys as well as the data about how Dora compares to a human doctor to evaluate Dora's usefulness
- We will publish data in academic journals, but all data will be anonymous

Do you have any questions?

- If you have any (more) questions or comments, you are welcome to email me or another researcher on the participant information sheet

Appendix 5-A. Case Study 1: Thematic framework

Theme	Subtheme	Subtheme 2	Codes
Facilitators / positive feedback	Motivating / engaging	App was motivating	App helped motivate (1) Accountability App increased awareness of behaviour App increased awareness of weight Liked recording progress? (1) Recording progress was motivating? (1) Prompted behaviours
		App was engaging	Engagement / use of app? (1) Engaging kids? (1) Interactivity? (1) Kids used the app (2) Adoption into daily routine
	Perceived behavioural benefits	-	Ate more healthy (1) More active / outside more (1) Increased water consumption Tried new things
	Suggestions were helpful	-	App supported self-efficacy? (1) Overcoming barriers helps with self-efficacy Suggestions supported self-efficacy (1) Liked activity suggestions? (1) Liked healthy eating swaps? (1) Liked parent survival guide? (1) Links to other resources were useful Goals helped with planning (1) Liked goal suggestions / found helpful Guidance (1 - goal setting)
	Good usability / acceptability	-	Easy to use (1 - goal setting) User friendly / easy to use (2) Kid friendly (2) Aesthetic (1)
	Liked features	Liked family aspect	Liked family involvement / doing together
		Liked games	Kids liked games? (1) Participant like games? (1) Educational
		Liked goal setting	Flexibility (1) Goals helped with planning Liked goal setting? (2) Simple goals Thinking about barriers was useful
		Liked app feedback	Notifications useful / lacking? (1) Feedback useful? (1)

			Liked points trophies? (1)
		Elements liked in Family Survey	Personalisation (1 - family survey) Inclusion of other healthy behaviours Liked oral health addition
		Liked NoObesity name	Medical name is good Opinion of name? (2) Prevent from downloading? (0)
	Positive opinions about HCP link	-	Accountability Comfort discussing with HCP? (2) Communicating via app (1) Could integrate with other health services / family HCP feedback useful? (1) Want human connection through app Would link with HCP? (1)
Barriers / negative feedback	Not motivating / engaging	General	App helped motivate (0) Engagement / use of app (0) Recording progress was motivating (0)
		Not enough variety - repetitive	Repetitive (games) Repetitive (goal setting) Simple goals (0)
		Not enough variety - limited content	Limited things to do <i>Liked recording progress (0)</i> Boring Nothing new Mostly used to record progress Not realistic (healthy swaps) Looked for suggestions / ideas outside of app Liked healthy eating swaps (0) Flexibility (0 - goal setting)
		Not memorable	<i>Liked recording progress (0)</i> Not memorable (family survey) Forgot to keep up
		Not engaging for kids	Interactivity (0) Engaging kids (0) Kids used the app (0,1)
		Feedback was lacking	Feedback useful (0) Lack of feedback Liked points / rewards (0) Notifications useful/lacking? (0)
		Issues with notifications	Daily notifications too much Didn't take much notice of notifications Noise of notification frustrating Notifications came at bad times

	No perceived benefits from app	Didn't improve health behaviour	Didn't help Already doing behaviour More active / outside more (0) Ate more healthy (0) Healthy eating harder than activity
		Didn't support self-efficacy	App supported self-efficacy (0) Can't overcome some barriers
	Poor usability / acceptability	Poor usability	User friendly / easy to use (0,1) Kid friendly (0,1) Differences within family
		Poor acceptability	Kids liked games (0) Participants liked games (0) Liked goal setting (0,1) Aesthetic (0) Poor acceptability
		Lack of clarity	Points not clear / need more meaning Unclear how it would work Confused about who it was for
		Lack of guidance	Lack of guidance (usability) Easy to use (0 - goal setting) Guidance (0 - goal setting) No help function
		Specific complaints	Specific complaints Technical difficulties Specific complaints (games) Can't change inputted data Inaccuracies
	Barriers	Barriers to app use	<i>Child's age</i> <i>Covid-19</i> Device
		Barriers to behaviour	Accessibility <i>Child's age</i> <i>Covid-19</i> Habits / routines Lack of accountability Lack of motivation (convenience) Lack of skills Lack of time Obesity management different for kids vs adults Other health conditions Picky eaters Sibling teasing Weather
	Sensitive topic	Issues with name	Can see why others wouldn't like Make name more positive Negative connotations

			Opinion of name (0,1) Prevent from downloading (1,2)
		Weight / health measurements	Discomfort tracking kids weight Discomfort tracking own weight
		Concerns with linking with HCP	Comfort discussing with HCP (0,1) Sensitive topic App data didn't reflect true behaviour Felt that it was inappropriate Would link with HCP (0) Communicating via app (0) HCP feedback useful? (0) Privacy / security concerns Wouldn't want to bother HCP Would link with HCP (0,1) Depends on clinician
Suggestions for improvement	More content and more variety of content	-	More variety More information / education Home workouts Local activities Ways to overcome personal barriers Recipes / meal planning / shopping support Vegetarian / vegan options
	More engagement for kids	-	More family / kid engagement Competition would motivate kids Points / rewards could be motivating (esp for kids)
	Improve self-monitoring feature	-	Trends (incl. retrospective review and recording, visualise progress) Simplify recording progress Add notes
	Improve notifications	-	Prompts / reminders Targeted notifications
	Improve usability	-	Interface redesign Date of birth rather than age Back button functionality Intro / guide to use
	Make app more positive	-	Positive feedback / encouragement Focus on feelings not measurements
	Increased personalisation	-	Personalisation Separate goals
	Integration	-	Could integrate with other health services / family Integration with other health devices Include / promote in schools
Misc.	RE-AIM (maintenance)		Continue use after study?

Appendix 5-B. Case Study 1: Thematic framework mapping factors to COM-B

Theme	Subtheme	Codes	Factor summary	Short description of factor	Factor association with theme	COM-B component
Engagement - motivation	Poor acceptability	Kids liked games (0) Participants liked games (0) Liked goal setting (0,1) Aesthetic (0) Poor acceptability	Aesthetic appeal	Visual interface and colour scheme could affect participants' feelings about the app	Positive	Motivation (automatic)
Engagement - motivation	Weight / health measurements	Discomfort tracking kids weight Discomfort tracking own weight	Discomfort tracking weight	Parents did not want an emphasis on weight and measurements, particularly for their children	Negative	Motivation (automatic) Motivation (reflective)
Engagement - motivation	Feedback was lacking	Feedback useful (0) Lack of feedback Liked points / rewards (0) Notifications useful/lacking? (0)	Feedback / visualisations	Participants all valued feedback from the app about their progress and how their current behaviour compared to healthy recommendations, but many found the app's feedback lack.	Positive	Capability (psychological) Motivation (automatic)
Engagement - motivation	Improve self-monitoring feature	Trends (incl. retrospective review and recording, visualise progress) Simplify recording progress Add notes				
Engagement - motivation	Integration	Could integrate with other health services / family Integration with other health devices Include / promote in schools	Integration	A couple participants suggested that enabling the app to connect with wearables or linking it with existing health service delivery could be helpful for engagement and impact	Positive	Opportunity (physical)
Engagement - motivation	Liked games	Kids liked games? (1) Participant like games? (1) Educational	Interactivity / gamification	Participants generally liked the idea of having gamification in the app, but many found the current games repetitive and the points/rewards system confusing. Including friendly competition, rewards, and games was a common suggestion to improve children's engagement	Positive	Motivation (automatic)
Engagement - motivation	More engagement for kids	More family / kid engagement Competition would motivate kids Points / rewards could be motivating (esp for kids)				
Engagement - motivation	App was engaging	Engagement / use of app? (1) Engaging kids? (1) Interactivity? (1) Kids used the app (2) Adoption into daily routine	<i>Descriptive sub-theme (not a factor)</i>	-	-	-
Engagement - motivation	Not engaging for kids	Interactivity (0) Engaging kids (0) Kids used the app (0,1)	<i>Descriptive sub-theme (not a factor)</i>	-	-	-
Engagement - motivation	Continue use after study?	-	<i>Descriptive sub-theme (not a factor)</i>	-	-	-
Engagement - motivation	Liked NoObesity name	Medical name is good Opinion of name? (2) Prevent from downloading? (0)	Negative / positive connotations	The connotations of the app were important for some of the participants,	Either	Motivation (automatic)

Engagement - motivation	Issues with name	Can see why others wouldn't like Make name more positive Negative connotations Opinion of name (0,1) Prevent from downloading (1,2)		who felt that a more positive focus on health and behaviour was better than a focus on weight and obesity		
Engagement - motivation	Make app more positive	Positive feedback / encouragement Focus on feelings not measurements				
Engagement - motivation	Issues with notifications	Daily notifications too much Didn't take much notice of notifications Noise of notification frustrating Notifications came at bad times	Notifications	Notifications were seen as helpful reminders to use the app, but parents wanted to be able to decide when was a convenient time to receive them	Either	Capability (psychological) Motivation (automatic) Motivation (reflective)
Engagement - motivation	Improve notifications	Prompts / reminders Targeted notifications				
Engagement - motivation	Not enough variety - repetitive	Repetitive (games) Repetitive (goal setting) Simple goals (0)	Novelty / variety	Participants wanted new content to encourage them to keep looking at the app	Positive	Capability (psychological) Motivation (automatic)
Engagement - motivation	Not enough variety - limited content	Limited things to do Liked recording progress (0) Boring Nothing new Mostly used to record progress Not realistic (healthy swaps) Looked for suggestions / ideas outside of app Liked healthy eating swaps (0) Flexibility (0 - goal setting)				
Engagement - motivation	Not memorable	Liked recording progress (0) Not memorable (family survey) Forgot to keep up				
Engagement - motivation	More content and more variety of content	More variety More information / education Home workouts Local activities Ways to overcome personal barriers Recipes / meal planning / shopping support Vegetarian / vegan options				
Engagement - motivation	Elements liked in Family Survey	Personalisation (1 - family survey) Inclusion of other healthy behaviours Liked oral health addition	Personalisation	Participants liked the option to put in a family photo and expressed desire to make the app more personal to them	Positive	Capability (psychological)
Engagement - motivation	Increased personalisation	Personalisation Separate goals				
Engagement - usability	Good usability	Easy to use (1 - goal setting) User friendly / easy to use (2) Kid friendly (2)	Ease of use	Most participants felt that the app was easy to use, but some noted that reducing the effort required to use the app would make them more likely to use it	Positive	Capability (psychological)
Engagement - usability	Poor usability	User friendly / easy to use (0,1) Kid friendly (0,1) Differences within family				
Engagement - usability	Improve usability	Interface redesign Date of birth rather than age Back button functionality Intro / guide to use				

Engagement - usability	Lack of clarity	Points not clear / need more meaning Unclear how it would work Confused about who it was for	Lack of clarity / guidance	Several participants felt that a lack of clarity hindered their use of the app and suggested incorporating a guide or introduction to clarify its aims and function	Negative	Capability (psychological) Motivation (automatic)
Engagement - usability	Lack of guidance	Lack of guidance (usability) Easy to use (0 - goal setting) Guidance (0 - goal setting) No help function				
Engagement - usability	Specific complaints	Specific complaints Technical difficulties Specific complaints (games) Can't change inputted data Inaccuracies	Technical difficulties	Some participants reported issues of the app freezing, getting locked out, or not being able to click on the goals that interfered with their use	Negative	Opportunity (physical)
Engagement - usability	Barriers to app use	<i>Child's age</i> <i>Covid-19</i> <i>Device</i>	Unable to use	Children's engagement with the app was limited by their age	Negative	Opportunity (physical)
Perceived impact - motivation	App was motivating	App helped motivate (1) Accountability App increased awareness of behaviour App increased awareness of weight Liked recording progress? (1) Recording progress was motivating? (1) Prompted behaviours	Awareness of behaviour	Goal setting and self-monitoring features reminded participants to pay attention to their current behaviour and whether it aligned with their goals (helped to hold themselves accountable)	Positive	Motivation (automatic) Motivation (reflective) Capability (psychological)
Perceived impact - motivation	Liked family aspect	Liked family involvement / doing together	Doing as a family	Participants liked the prompt to consider what was healthy for the family as a whole and do healthy activities together with everyone involved	Positive	Opportunity (social)
Perceived impact - motivation	Not motivating	App helped motivate (0) Engagement / use of app (0) Recording progress was motivating (0)	<i>Descriptive sub-theme (not a factor)</i>	-	-	-
Perceived impact - self-efficacy	More content and more variety of content	More variety More information / education Home workouts Local activities Ways to overcome personal barriers Recipes / meal planning / shopping support Vegetarian / vegan options	App suggestions (for behaviours)	Healthy eating and activity swaps helped give parents easy ideas for improving their behaviours	Positive	Capability (psychological)
Perceived impact - self-efficacy	Suggestions were helpful	App supported self-efficacy? (1) Overcoming barriers helps with self-efficacy Suggestions supported self-efficacy (1) Liked activity suggestions? (1) Liked healthy eating swaps? (1) Liked parent survival guide? (1) Links to other resources				

		were useful Goals helped with planning (1) Liked goal suggestions / found helpful Guidance (1 - goal setting)				
Perceived impact - self-efficacy	Liked goal setting	Flexibility (1) Goals helped with planning Liked goal setting? (2) Simple goals Thinking about barriers was useful	Goal setting support	Some participants noted finding the app's suggestions for goals and prompts to consider barriers helpful for setting goals and making plans	Positive	Capability (psychological)
Perceived impact - self-efficacy	Didn't support self-efficacy	App supported self-efficacy (0) Can't overcome some barriers	<i>Descriptive sub-theme (not a factor)</i>	-	-	-
Perceived impact - self-efficacy	Liked app feedback	Notifications useful / lacking? (1) Feedback useful? (1) Liked points trophies? (1)	Useful feedback	Feedback from the app	Positive	Capability (psychological) Motivation (automatic) Motivation (reflective)
Perceived impact - self-efficacy	Liked app feedback	Notifications useful / lacking? (1) Feedback useful? (1) Liked points trophies? (1)	Notifications	Prompts from the app were generally seen as helpful to remind the users to engage with the behaviour	Either	Capability (psychological) Motivation (automatic) Motivation (reflective)
Perceived impact - health behaviours	Perceived behavioural benefits	Ate more healthy (1) More active / outside more (1) Increased water consumption Tried new things	<i>Descriptive sub-theme (not a factor)</i>	-	-	-
Perceived impact - health behaviours	Didn't improve health behaviour	Didn't help Already doing behaviour More active / outside more (0) Ate more healthy (0) Healthy eating harder than activity	Already doing behaviour	Participants who were already active or eating mostly healthy didn't think the app had an impact on their behaviour	Negative	Motivation (reflective)
Perceived impact - health behaviours	Barriers to behaviour	Accessibility Other health conditions	Lack of accessibility / affordability	Participants highlighted several barriers relating to accessibility, including the availability of local, affordable activities, safe spaces for children to be active outside, and the cost of healthy food	Negative	Opportunity (physical)
Perceived impact - health behaviours	Barriers to behaviour	Child's age Obesity management different for kids vs adults Picky eaters Sibling teasing	Family dynamics	A variety of different family dynamics were mentioned that influenced family's weight-related behaviours, including children's willingness to eat certain foods, their age and ability to do certain activities, and sibling teasing	Either	Capability (physical) Capability (psychological) Opportunity (social)

Perceived impact - health behaviours	Barriers to behaviour	Covid-19	Covid-19	Covid-19 restrictions limited families options for activities that they could do and affected their motivation to do healthy behaviours	Negative	Opportunity (physical)
Perceived impact - health behaviours	Barriers to behaviour	Lack of skills	Lack of skills / ability	While none of the participants reported a lack of skills as a personal barrier, it was mentioned as a potential barrier that could hinder some people in preparing healthy meals	Negative	Capability (physical) Capability (psychological)
Perceived impact - health behaviours	Barriers to behaviour	Lack of motivation (convenience) Lack of accountability Habits / routines	Lack of motivation	A lack of motivation - just "not wanting to" - was often a barrier to making an extra effort to do a healthy behaviour	Negative	Motivation (automatic) Motivation (reflective)
Perceived impact - health behaviours	Barriers to behaviour	Lack of time	Lack of time	Finding time to prepare healthy meals or do activities as a family was a barrier	Negative	Opportunity (physical)
Perceived impact - health behaviours	Barriers to behaviour	Weather	Bad weather	Temperature, daylight, and weather conditions affected participants' motivation and ability to go outside to be active	Negative	Motivation (automatic) Motivation (reflective) Opportunity (physical)
Perceived impact - communication with HCP	Positive opinions about HCP link	Accountability	Accountability	Some participants mentioned that having a sense of accountability to a HCP who could see their progress would help motivate them to do their goals	Positive	Motivation (reflective) Opportunity (social)
Perceived impact - communication with HCP	Positive opinions about HCP link Concerns with linking with HCP	Comfort discussing with HCP? (2) Communicating via app (1) Want human connection through app Would link with HCP? (1) App data didn't reflect true behaviour Comfort discussing with HCP (0,1) Depends on clinician Sensitive topic Would link with HCP (0,1) Felt that it was inappropriate HCP feedback useful? (0) Wouldn't want to bother HCP	Comfort / trust with HCP	Many participants felt that the benefit of having a linked HCP would depend on how comfortable they felt discussing weight-related issues with them, with one participant noting that they would worry that the recorded progress on the app wouldn't reflect their true behaviour and would give the wrong impression	Positive	Motivation (automatic) Motivation (reflective) Opportunity (social)
Perceived impact - communication with HCP	Concerns with linking with HCP	Communicating via app (0) Privacy / security concerns	Privacy / security concerns	One participant mentioned that they would want to understand how personal data was used and protected before sharing anything via the app	Negative	Motivation (reflective)

Perceived impact - communication with HCP	Positive opinions about HCP link	Could integrate with other health services / family HCP feedback useful? (1)	Useful feedback (from HCP)	Most participants felt that having the HCP be able to provide feedback based on their family's goals and progress and make suggestions would be helpful	Positive	Capability (psychological) Motivation (automatic) Motivation (reflective) Opportunity (physical)

Appendix 5-C. Case Study 1: Thematic framework mapping factors to engagement components and patterns of engagement

Factor	Factor association	COM-B component	TDF	Ppt	Quotes	Eng. component	Pattern of engagement
Engagement							
Aesthetic appeal	Positive	Motivation (auto)	Emotion	5	I didn't find the colors, the green, I didn't find it appealing. Like, there are some apps that I go on because they just appeal to me and I just find the colors vibrant and I just enjoy being on the app. This app, I didn't find that. ... Aesthetically, it was not pleasing to... If I was in a horrible mood, it didn't make me in a better mood ... it's just a very depressing app	Affective Cognitive Behavioural	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro)
				4	Yeah. I think it looks really good. It looks engaging. You know, if you show this to a child, they would find it interesting because of the iconography and how it looks.		Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
Discomfort tracking weight	Negative	Motivation (auto) Motivation (ref)	Emotion Beliefs conseq.	10	I don't think my daughter, who is eight, would have liked seeing, if she had been on the slightly heavier side or whatever, I don't think she would have liked seeing it. I think she is getting to an age where she is more aware of it.	Affective Cognitive Behavioural	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro)
				12	I am thinking as me as a parent, I didn't want to put the details in there and be too hung up on measurement, if you know what I mean.		Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
Feedback / visualisations	Positive	Capability (psych) Motivation (auto)	Reinforcement Knowledge	4	I think the feedback you got in that family survey thing is helpful because it's equating it to what you're actually doing, what you might think, "Well, okay. I eat enough fruit in the day. Because you think you eat one apple." But obviously, if you put that in there, it's going to say, "Well, you're not eating enough fruit, so you need to think about it."	Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
				5	If there was a graph or something, I could see that it's declining, my water intake, it would make me want to act on that. But I couldn't see my progress... it was not visually engaging enough for me to think ooh,		Behavioural (micro) → Cognitive → Behavioural (macro)

					look, my graph's going down. I haven't been drinking as much water as I said I'd like to. Let me get back onto it.		
Integration	Positive	Opportunity (phys)	Envr. Social	1 4	If [my kids] could connect their smartwatch to that, and their step counts, and be in competition and that kind of thing. Yes, they would love it. Obviously, you're having all your weight checks and that if they feel that there's a problem, that they will refer you to these healthy eating things. So that would really help, I suppose, at that point to click in and say, "Well, look, use this. We'll link it to them, and put it in at that point because that's trying to get it as early as possible.	Affective Cognitive Behavioural	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro) Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro) Context → Cognitive → Behavioural (macro)
Interactivity / gamification	Positive	Motivation (auto)	Envr. Emotion, Reinforcement Goals Social	14 9	It's like human nature, we all want to get these wee awards and get these trophies and earn these points, whether that's just a personality or whatever it is. I know that as soon as you set a goal on something, I'm going to want to achieve it and I'm going to want to do the best at it that I can. Something to make it more interactive, more fun. I don't know, like daily challenges for the children or I don't know. You've now been eating fruits for a whole month consistently. You've unlocked 10 new character. It's more like game driven, but informative game driven.	Affective Cognitive Behavioural	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro) Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro) Behavioural (micro) → Affective → Behavioural (macro) Behavioural (micro) → Cognitive → Behavioural (macro)

Negative / positive connotations	Either	Motivation (auto)	Envr. Emotion, Reinforcement	5 4 12	Again, I think that it could just have been a bit more encouraging. So if I went days, for example, without doing well, I'll just put it like that, it just said oh, success recorded or something like that, was coming up. If we'd [inaudible 00:06:49] badly, it could have been a bit more attractive, a bit more engaging. And I know that it's... technology, but I just think that it could offer a lot more, but it just wasn't offering it. The word [obesity] is what it is and it's a word that is used. It's obviously got lots of negative connotations. Possibly, in, for me, because I think that, everything's connected to how you feel about yourself, how you enjoy going out in the fresh air and the things that you're eating, they're all connected. So, for me, how you feel would have been better.	Affective Behavioural	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro) Affective → Behavioural (macro)
Notifications	Either	Capability (psych) Motivation (auto) Motivation (ref)	Memory.. Emotion	3 8	I think it's more the timing. I don't mind the notifications coming up, and if it was in the evening and it reminded me sort of when my little girl was in bed, then I probably would have gone on to the app, and ticked it off, and kept up with it. I think it was more the time in. Not necessarily pestering you with notifications, like notifications would be good to remind you to log it. But if the feedback was just giving you in the app when it was open, I think that'd be really good because I think people are motivated by that	Affective Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Affective → Subsequent Behavioural (micro) Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
Novelty / variety	Positive	Capability (psych) Motivation (auto)	Memory.. Emotion Beliefs consq. Envr.	2 15	It's very limited in what you can and cannot do with it. They've got a little bit of games, but after you've done that a couple of times, you're not bothered in it anymore." There was nothing new, there was no change to it. Actually, had there been some new content on there, it would have encouraged me to come back more and more and go, "Oh, what more can I learn? Can I beat it again on learning something else?"	Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
Personalisation	Positive	Capability (psych)	Envr.	15	That was something else I really liked as well actually, the photo. To make it personalized to us. That was one of the first things I did to make it really like it was about us. So no, I did like that aspect as	Affective Behavioural	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro)

					well, to be able to actually personalize it. I suppose that's the only other thing, if there was any other ways of personalizing it a little bit more to make it our family and things like that.		
Ease of use	Positive	Capability (psych)	Memory. ..	9	It was very easy. It wasn't complicated. It wasn't something that I needed to fill out a lot of details or try to really use my brain to think, "Why is this not working?"	Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
				3	So, maybe it was just like a one click kind of thing for that day, or if it could come off as like a question or something, "Have you drank your water for today?" And, you can just click Yes, rather than going through and tick it on each day. Or, have to be in your three or four jog that you said you were going to do, and you could just click yes or no. That would actually probably make me do it more, I think.		
Lack of clarity / guidance	Negative	Capability (psych) Motivation (auto)	Memory. ..., Knowledge	6	"There's no sort of introduction thing really. So it could've possibly done with an introduction page to start with."	Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
Technical difficulties	Negative	Opportunity (phys)	Envr.	2	And also with that was, if you accidentally click on the wrong date, you weren't able to unclick it. They ask you, you get a notification in the morning telling you to do yesterday's activity	Behavioural	Context → Subsequent Behavioural (micro)
				11	Yeah, it was fairly easy. I did come across a problem. I don't know. I guess I'd been using it a couple of months and then it just froze on me.		
Unable to use	Negative	Capability (phys) Opportunity (phys)	Envr. Skills	1	My kids are a bit young to do it for themselves because they're four and seven.	Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
				11	They tried a few with the little kind of game things that run their [inaudible 00:07:14] or something where you pull down that thing. My eldest son he's 18 fairly understood it, but the younger one really didn't.		
Perceived impact							

Accountability (to HCP)	Positive	Motivation (auto) Motivation (ref) Opportunity (soc)	Social Beliefs consq.	1	And also someone to be accountable for. So, someone to say, why didn't you have fruit on that day then? And because it's hard to explain it sometimes isn't it? If the truth is because I was a bit lazy. And if you know you've got to say that to someone next week, you're more likely to go and have an Apple love.	Affective Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Affective → Subsequent Behavioural (micro) Behavioural (micro) → Cognitive → Behavioural (macro)
Comfort / trust with HCP	Positive	Motivation (auto) Opportunity (soc)	Emotion Beliefs consq.	6 5	But if you didn't trust each other to start with, then I could see you might be thinking, well they're checking up on me or all this. But, because I just found the app really tedious at times, I knew that I had to update it and I just couldn't be bothered, honestly, to go and update it, I don't know, I feel it would make me out to be somebody that's not interested in my children's nutrition, and not worried, not interested in how I'm feeding them and stuff like that. It portrayed me to be somebody that's not taking an interest in their food and wanting to keep the healthcare professional updated. And so, that wouldn't be the truth because I would be doing my part. I just wouldn't be updating it on the app, but they wouldn't know that	Affective Behavioural	Context → Affective → Behavioural (micro)
Privacy / security concerns	Negative	Motivation (ref)	Beliefs consq.	13	I would think initially I would probably be quite skeptical about how secure it was, because anything to do with that, I'm really into the privacy. ... You need to know the information behind it and who's made it and where the information is stored and how it's used as well, definitely.	Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)

Useful feedback (from HCP)	Positive	Capability (psych) Motivation (auto) Motivation (ref) Opportunity (soc)	Reinforcement Beliefs cap. Knowledge Social	7	[If] she had a healthcare checkup, and actually, had they then interacted with us on the app saying, "Here's some really good things to do. Oh, I see you haven't been moving very much, do you need some more ideas of what to do?" Particularly during the times that we've had recently, actually, I think it may have been quite useful.	Cognitive Behavioural	Context → Cognitive → Subsequent Behavioural (micro) Context → Cognitive → Behavioural (macro)
				2	Whereas if they had access to the information, it would be a lot easier, I feel. They could be, "We need you to be getting on with this. We need you to be doing that. I see you're doing good with this." Also it's like they can see that you're doing stuff and there's no progress or the results they're looking for, but yet you're doing the work but there's no progress, then they can be seeing something's not working here, we have to take a different approach.		
Already doing behaviour	Negative	Motivation (ref)	Intentions / goals	4	We live very close to school, so it's easiest to walk. So that thing that I put in as our first goal about walking to school every day. We were going to use it every day, anyway. Then I suppose it's getting a bit more creative about what we could put on as another goal above that because I suppose I could put on his PE lessons, but he was already doing that every day.	Behavioural	Context → Behavioural (macro)
Bad weather	Negative	Opportunity (phys)	Envr. Emotion Beliefs consq.	12	And, there was definitely more motivation to do it when the weather was a bit warmer, but now it's dropped so cold, there is less.	Affective Cognitive	Context → Affective → Behavioural (macro)
				15	So once we've had a family meal it's dark, the children aren't wanting to go back out. I'm not really wanting to go back out. It's cold.	Behavioural	Context → Cognitive → Behavioural (macro)
Covid-19	Negative	Motivation (auto) Opportunity (phys)	Envr.	10	I guess it's been harder because we now don't go to school, so we don't walk to school or anything like that.	Affective Behavioural	Context → Behavioural (macro)
				11	It's not always easy, especially with lockdown when you don't feel great anyway, trying to look at healthiest snacks as well.		Affective → Behavioural (macro)
				3	And, just as well with the lockdown, a lot of the stuff you couldn't have done like swimming and that kind of thing.		

Family dynamics	Either	Opportunity (phys) Opportunity (soc)	Social Skills	13	I think for him, because he was at the age where quite a lot of groups and things we would have joined you had to be six or seven, and some of it is a safety, like riding his bike and things like that, he had more practice. He is not exactly road safe, so it's things like that.	Behavioural	Context → Behavioural (macro)
				14	He would eat loads of fruit, and generally most food, but he was so hard to get to eat vegetables. And he just refused to eat anything.		
				15	We have kiddie teasing in our house, I suppose older brother to younger brother teasing the other one. And the fat word does get used occasionally.		
Lack of accessibility / affordability	Negative	Opportunity (phys)	Envr.	1	My goal was to get outside more, be outside every day. We don't have outside space for us. It would be easier to let the kids play in the garden, but we don't have a garden. So, it's an effort to do it every day.	Cognitive Behavioural	Context → Behavioural (macro) Cognitive → Behavioural (macro)
				8	Because I know from doing the supermarket shop, it's cheaper and easier to get frozen food and unhealthy snacks than it is to get healthier ones.		
				10	Because if some of the suggestions, say we live in xxxxxx, it's not got the widest variety of activities that you can do around here. And therefore some of the ideas were good, but there's just nowhere around here that we'd be able to do them.		
Lack of motivation	Negative	Motivation (auto) Motivation (ref)	Emotion Intention s, Beliefs conseq.	4	Because we all know if we're tired, you're more likely to want to eat more because you feel you need something to give you a boost. It's like, "Okay, well I'm going to go and have something like a bar of chocolate because it's going to make me feel a bit better.	Affective Cognitive Behavioural	Context → Behavioural (macro) Affective → Behavioural (macro) Cognitive → Behavioural (macro)
				6	You know when you don't really want to do something yourself, you can't really find that enthusiasm to assist your child with it as well.		
				11	Just because it can be so easy just to grab something quick rather than, okay, I'm going to prepare a piece of fruit for my children and they will eat the fruits.		

Lack of skills / ability	Negative	Capability (psych) Capability (phys)	Skills	8	But I think they eating side of it is harder because it's just easier to buy your junk food, it's cheaper and it's easy to put... I mean, I do it myself sometimes, to put something in the oven to cook rather than cook from scratch. And some people might not necessarily have the skills.	Cognitive Behavioural	Context → Behavioural (macro) Cognitive → Behavioural (macro)
Lack of time	Negative	Opportunity (phys)	Envr.	5 15	Like, I'm a parent and finding the time to cook a healthy, nutritious meal, it's limited at times. We've literally been stuck at home with the children. I work in the afternoon, so once we've done schoolwork I find our time is very limited. I go out to work and I'm not back till half five/six	Behavioural	Context → Behavioural (macro)
Awareness of behaviour	Positive	Motivation (auto) Motivation (ref) Capability (psych)	Goals, Beliefs cap. Memory. ..., Behav reg.	2	We've got habits and this kind of brings it forward for you to actually acknowledge what you're doing and what you're not doing, and to help you to correct it.	Cognitive Behavioural	Behavioural (micro) → Cognitive → Behavioural (macro)
Doing as a family	Positive	Motivation (auto) Opportunity (soc)	Social Emotion	3	It just felt like something that we were all doing together. Soon as you downloaded it, it did feel sort of like a whole family thing that you were going to work, and do challenges together, and that's what I liked about it.	Affective Behavioural	Behavioural (micro) → Affective → Behavioural (macro)
App suggestions (for behaviours)	Positive	Capability (psych)	Knowledge	8 1	I thought the useful links were good that led you to other activities and different websites or places that you could go and visit because I think if you're not a family that's necessarily an outdoorsy family, you might not even know where to start with, where to go. But actually, stuff like choose a bit of fruit instead of a biscuit. It's such a sensible and obvious thing. And I couldn't see the wood for the trees for a bit if you know what I mean? And it was helpful in that way.	Cognitive Behavioural	Behavioural (micro) → Cognitive → Behavioural (macro)
Goal setting support	Positive	Capability (psych)	Goals Knowledge Beliefs	1 8	I liked that they were simple, and I liked the way that they were suggested. Otherwise, I would have been sitting there wondering, I don't know what a goal should be for this. I think the barrier thing was quite good as well	Cognitive Behavioural	Behavioural (micro) → Cognitive → Behavioural (macro)

Appendix 5-D. Case Study 2: Thematic framework

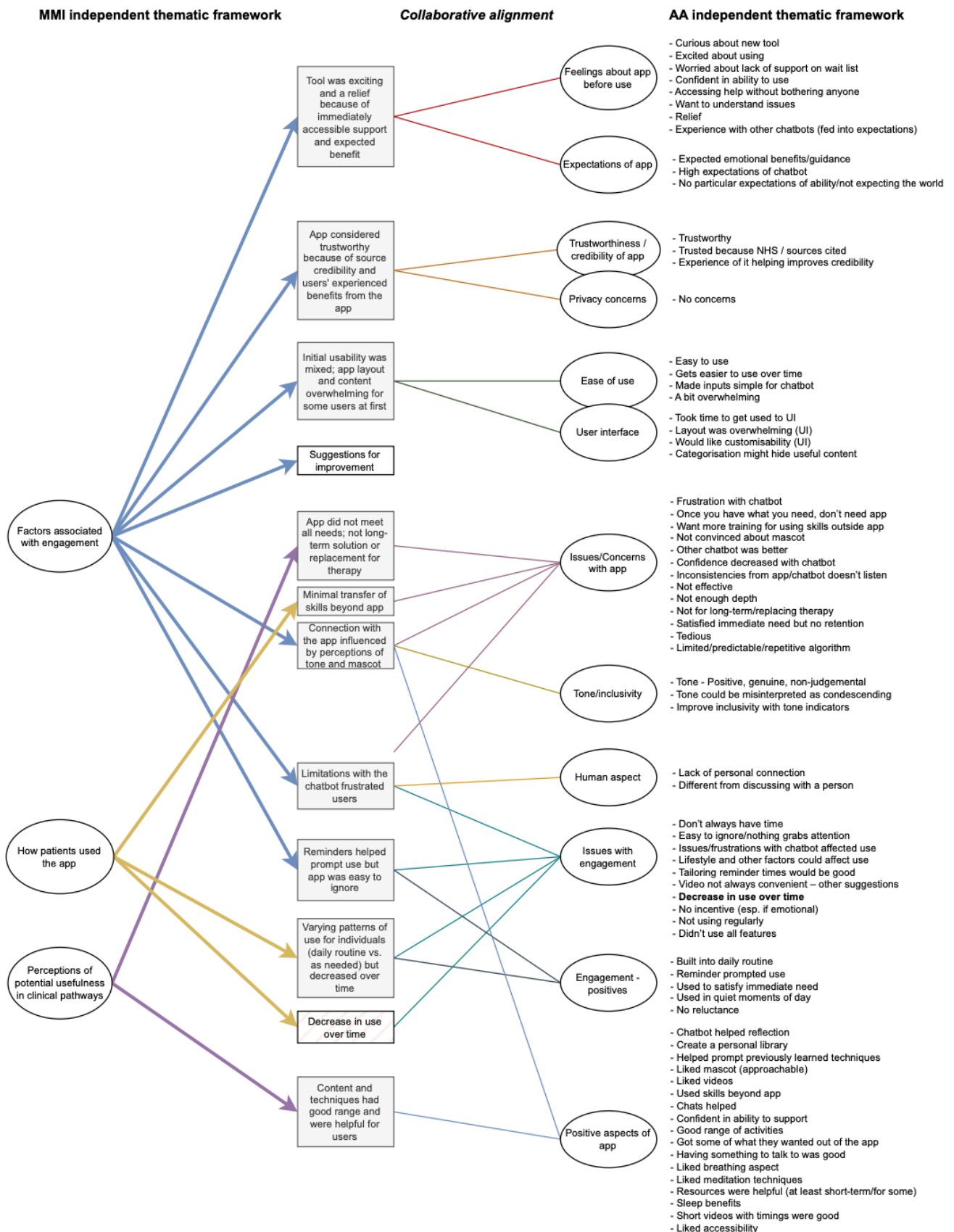
Engagement component	TFA construct	Code (sub-codes)	Subtheme	Organising theme
Affective	Affective attitudes	Affective attitudes - anticipated	Tool was exciting and a relief because of accessible support and expected benefit	Factors associated with engagement
Affective	Affective attitudes	<i>Curious about trial / new tool</i>		
Affective	Affective attitudes	<i>Excited about using</i>		
Affective	Affective attitudes	<i>Worried about lack of support on wait list</i>		
Affective	Affective attitudes	Relief		
Affective	Affective attitudes	Liked accessibility		
Affective	Affective attitudes, Ethicality	Accessing help without bothering anyone		
Affective	Affective attitudes	No reluctance		
Cognitive	Perceived effectiveness	Perceived effectiveness - anticipated		
Cognitive	Perceived effectiveness	<i>Expected emotional benefits / guidance</i>		
Cognitive	Perceived effectiveness	<i>High expectations of chatbot</i>		
Cognitive	Perceived effectiveness	<i>No particular expectations of ability - not expecting the world</i>		
Affective	Affective attitudes	Tone	Tone affected user connection with app	
Affective	Affective attitudes	<i>Positive, genuine, non-judgmental</i>		
Affective	Affective attitudes	<i>Tone could be misinterpreted as condescending</i>		
Affective	Affective attitudes	Liked mascot (approachable)		
Affective	Affective attitudes	Not convinced about mascot		
Affective	Affective attitudes	Frustration with chatbot	Chatbot limitations frustrated users	
Affective	Affective attitudes	Lack of personal connection		
Affective	Affective attitudes	Tedious		
Cognitive	Burden	Made inputs simple for chatbot		
Cognitive	Perceived effectiveness	Perceived effectiveness - experienced		

Cognitive	Perceived effectiveness	<i>Different from discussing with a person</i>		
Cognitive	Perceived effectiveness	<i>Confidence decreased with chatbot</i>		
Cognitive	Perceived effectiveness	<i>Inconsistencies from app / chatbot doesn't listen</i>		
Cognitive	Perceived effectiveness	<i>Limited / predictable / repetitive algorithm</i>		
Behavioural	-	Micro engagement		
Behavioural	-	<i>Issues/frustrations with chatbot affected use</i>		
Affective	Ethicality	Privacy concerns associated with what's shared with app	Trustworthiness shaped by source credibility and personal experience	
Affective	Ethicality	No concerns		
Affective, Cognitive	-	Trustworthy		
Affective, Cognitive	-	Trusted because NHS / sources cited		
Cognitive	Perceived effectiveness	<i>Experience of it helping improves credibility</i>		
Cognitive	Burden	A bit overwhelming	Easy to use but a bit overwhelming	
Cognitive	Burden	Easy to use		
Cognitive	Self-efficacy	Confident in ability to use		
Cognitive	Opportunity costs	No costs		
Cognitive	Opportunity costs	Short videos with timings were good		
Cognitive	-	Categorisation might hide useful content		
Behavioural	-	Micro engagement	Reminders helped prompt use but app was easy to ignore	
Behavioural, Cognitive	-	<i>Easy to ignore / nothing grabs attention</i>		
Behavioural, Cognitive	-	<i>Reminder prompted use</i>		
Behavioural, Cognitive	-	<i>Tailoring reminder times would be good</i>		
Affective	Affective attitudes	No incentive (esp. if emotional)		
Cognitive	-	App suggestions helpful - keep fresh	Suggestions for improvement	
Cognitive	-	Create a personal library		

Cognitive	Burden	Video not always convenient - other suggestions		
Behavioural	-	Macro engagement		
Behavioural, Cognitive	-	Want more training for using skills outside app		
Cognitive	Perceived effectiveness	Perceived effectiveness - experienced	Range of resources were useful for users	Perceptions of potential usefulness in clinical pathways
Cognitive	Perceived effectiveness	Confident in ability to support		
Cognitive	Perceived effectiveness	Having something to talk to was good		
Cognitive	Perceived effectiveness	Good range of activities		
Cognitive	Perceived effectiveness	Liked breathing aspect		
Cognitive	Perceived effectiveness	Resources were helpful (at least short-term/for some)		
Cognitive	Perceived effectiveness	Sleep benefits		
Cognitive	Perceived effectiveness	Chats helped		
Cognitive	-	Chatbot helped reflection		
Cognitive, Behavioural	-	Helped prompt previously learned techniques		
Affective	Affective attitudes	Liked videos		
Cognitive	Perceived effectiveness	Perceived effectiveness - experienced	App did not meet all needs; not long-term solution or replacement for therapy	
Cognitive	Perceived effectiveness	Benefits depend on user not just tool (e.g. motivation, experience with therapy, mental health issues)		
Cognitive	Perceived effectiveness	Not effective		
Cognitive	Perceived effectiveness	Not for long-term / replacing therapy		
Cognitive	Perceived effectiveness	Premium better than free?		
Cognitive	Perceived effectiveness	Got some of what they wanted out of the app		
Cognitive	-	Want to understand issues		
Behavioural	-	Micro engagement	Varying patterns of use for individuals	How patients used the app
Behavioural, Cognitive	Burden, Opportunity costs	Don't always have time		
Behavioural	-	Not using regularly		

Behavioural	-	<i>Used in quiet moments of day</i>		
Behavioural	-	<i>Used to satisfy immediate need</i>		
Behavioural	-	<i>Built into daily routine</i>		
Behavioural	-	<i>Lifestyle and other factors could affect use</i>		
Behavioural	-	<i>Didn't use all features</i>		
Behavioural	-	Micro engagement	Decrease in use over time	
Behavioural	-	<i>Decrease in use over time</i>		
Behavioural	-	Macro engagement	Minimal transfer of skills beyond app	
Behavioural	-	<i>Not much outside app</i>		
Behavioural	-	<i>Once you have what you need, don't need app</i>		
Behavioural	-	<i>Used skills beyond app</i>		

Appendix 5-E. Case Study 2: Initial independent thematic maps and collaborative alignment



Appendix 5-F. Case Study 2: Thematic framework mapping factors to engagement components and patterns of engagement

Organising theme	Theme	Ppt	Quotes	Code	Engagement component	Pattern of engagement
Factors associated with engagement	Tool was exciting and a relief because of accessible support and expected benefit	2	So I was just really curious on what kind of things they would be available to use.	Anticipated affective	Affective	Affective → Initial Behavioural (micro)
				<i>Curious about trial / new tool</i>	Affective	
		1	So the whole idea that it's an app and not another person is that... I'm not bothering anyone... why I don't talk about things is because I don't wanna burden people with my issues.	<i>Excited about using</i>	Affective	Cognitive → Initial Behavioural (micro)
				<i>Worried about lack of support on wait list</i>	Affective	
				Relief	Affective	Cognitive → Affective → Initial Behavioural (micro)
				Liked accessibility	Affective	
				Accessing help without bothering anyone	Affective	
		4	So I feel like it could be useful for many different people who are waiting for help.	No reluctance	Affective	
				Perceived effectiveness - anticipated	Cognitive	
				<i>Expected emotional benefits / guidance</i>	Cognitive	
				<i>High expectations of chatbot</i>	Cognitive	
				<i>No particular expectations of ability - not expecting the world</i>	Cognitive	
	Tone affected user connection with app	4	I like that it was very judgment free, like it didn't like point fingers.	Tone	Affective	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro)
				<i>Positive, genuine, non-judgmental</i>	Affective	
		3	It would ask you why you didn't check in in the morning. There was like 3 responses that it gave you and if you said any of them, it would come back with like quite a condescending perhaps response, but this could just be a unique issue because I have like difficulties understanding tone.	<i>Tone could be misinterpreted as condescending</i>	Affective	
				Liked mascot (approachable)	Affective	
	Limited chatbot frustrated users			Not convinced about mascot	Affective	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
		1	I know it's not the same as having a person there... but I found it becoming frustrating because like I say, I could predict it and if I said no to something it would ask me again or	Frustration with chatbot	Affective	
				Lack of personal connection	Affective	
				Tedious	Affective	
				Made inputs simple for chatbot	Cognitive	

		2	and... there were inconsistencies as well. It does affect me in terms of how much I access the app because unless I really want to use it, I probably just ignored it.	Perceived effectiveness - experienced	Cognitive	Initial Behavioural (micro) → Cognitive → Affective → Subsequent Behavioural (micro)
				<i>Different from discussing with a person</i>	Cognitive	
				<i>Confidence decreased with chatbot</i>	Cognitive	
				<i>Inconsistencies from app / chatbot doesn't listen</i>	Cognitive	
				<i>Limited / predictable / repetitive algorithm</i>	Cognitive	
				Micro engagement	Behavioural	
				<i>Issues/frustrations with chatbot affected use</i>	Behavioural	
	Trustworthiness shaped by source credibility and personal experience	1	I would have had deep confidentiality concerns because and I'm putting all my trust in the NHS there and hoping they've done due diligence, but you know they're a government entity, so they'd be foolish not to.	Privacy concerns associated with what's shared with app	Affective	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro) Initial Behavioural (micro) → Cognitive → Affective → Subsequent Behavioural (micro)
				No concerns	Affective	
				Trustworthy	Affective, Cognitive	
		4	<i>Did you find the program and the resources credible and trustworthy?</i> Yes, the previous exercises were really helpful actually, in moments and stress.	Trusted because NHS / sources cited	Affective, Cognitive	
				<i>Experience of it helping improves credibility</i>	Cognitive	
	Easy to use but a bit overwhelming	2	I think the way it was presented for us to follow through app is quite easy and just really straightforward. So there was no learning curve and trying to do the exercises.	A bit overwhelming	Cognitive	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
				Easy to use	Cognitive	
				Confident in ability to use	Cognitive	
				No costs	Cognitive	
		4	Not like really difficult, but like, and the overwhelming is the layout I think. Because it's not sectioned off, [the content is] kind of just there.	Short videos with timings were good	Cognitive	
				Categorisation might hide useful content	Cognitive	
	Reminders helped prompt use but app was easy to ignore	2	I don't think anything cuts my attention as much.	Micro engagement	Behavioural	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
		3	To be honest, I just kind of it was just kind of one of the things where it was very easy to ignore, which I	<i>Easy to ignore / nothing grabs attention</i>	Behavioural, Cognitive	
				<i>Reminder prompted use</i>	Behavioural, Cognitive	

		3	think as I was saying earlier about people like there's the option to just not open the app at all.	<i>Tailoring reminder times would be good</i>	Behavioural, Cognitive	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro)
			If you're in the heat of anger, you don't really have the time to do that, or you just can't be bothered to do that.	No incentive (esp. if emotional)	Affective	Context → Affective → Behavioural (micro)
	Suggestions for improvement	3	I think like even having just a rotating list that changes every day of just why don't you try this and then at the end saying is that worked or if it didn't work you got like delivery like saying like ohh did that work or did it not work because the only really thing it asks you is right your experience I don't think anybody is going to do that because it's just gonna be reading the app and not the exercise itself.	App suggestions helpful - keep fresh	Cognitive	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
				Create a personal library	Cognitive	Behavioural (micro) → Cognitive → Behavioural (macro)
				<i>Video not always convenient - other suggestions</i>	Cognitive	
				Macro engagement	Behavioural	
				<i>Want more training for using skills outside app</i>	Behavioural, Cognitive	
	Perceptions of potential usefulness in clinical pathways	4	The chat function was really good because it would help you. It wouldn't just reflect on it. It would help you like reset your mindset and most like it would make you rethink how you feel or something.	Perceived effectiveness - experienced	Cognitive	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
				<i>Confident in ability to support</i>	Cognitive	
				<i>Having something to talk to was good</i>	Cognitive	
		1	Taking a look at the demo dealing with whatever depression, anxiety with whatever mental health issues it's, it's certainly it a really good introduction to that and lots of good	<i>Good range of activities</i>	Cognitive	Initial Behavioural (micro) → Cognitive → Affective → Subsequent Behavioural (micro)
				<i>Liked breathing aspect</i>	Cognitive	
				<i>Resources were helpful (at least short-term/for some)</i>	Cognitive	
				<i>Sleep benefits</i>	Cognitive	

			different resources and covering different topics to to give people the confidence without feeling they're being railroaded into something to to actually go out and and do a little bit more investigation themselves.	Chats helped	Cognitive	Behavioural (micro) → Cognitive → Behavioural (macro)
				Chatbot helped reflection	Cognitive	
				Helped prompt previously learned techniques	Behavioural, Cognitive	
				Liked videos	Affective	
	App did not meet all needs; not long-term solution or replacement for therapy	2	Uh, I think I would give it on a scale of 10, maybe it's six cause for me definitely the exercises help, but if I'm looking for uh uh, not journaling, but like that conversation to help me ease mood or something. When you're like having like an episode or something, it definitely not gonna help uh with the current limitations and how the chat works.	Perceived effectiveness - experienced	Cognitive	Context → Behavioural (macro) Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
		1	But for me, I needed something more. I needed 'Why is this happening?' kind of thing ... rather than a quick sometimes this happens, sometimes that happens and I'm thinking well, no, that's not the case here.	Benefits depend on user not just tool (e.g. motivation, experience with therapy, mental health issues)	Cognitive	
				Not effective	Cognitive	
				Not for long-term / replacing therapy	Cognitive	
				Premium better than free?	Cognitive	
				Got some of what they wanted out of the app	Cognitive	
				Want to understand issues	Cognitive	
How patients used the app	Varying patterns of use for individuals	3	And so I said that I'm frustrated and it gave me an 8 minute video to watch. And in my head, I was like, well, I didn't have time to do that.	Micro engagement	Behavioural	Context → Subsequent Behavioural (micro)
				Don't always have time	Behavioural, Cognitive	
				Not using regularly	Behavioural	
		2	Because I feel like if I have, I'm like forced to always check in like in the morning and it becomes like too tedious and I will not use it like genuinely. I would just like check in for sake of checking in, so I'm not really using it that it's beneficial for me and it's more like I'm being forced to do it and out of my own.	Used in quiet moments of day	Behavioural	Context → Behavioural (macro)
				Used to satisfy immediate need	Behavioural	
				Built into daily routine	Behavioural	
				Lifestyle and other factors could affect use	Behavioural	
		4	I did like the check-ins that it does. So it's like what will happen in the day or what might happen in the day and at the end of the day what has happened that day.	Didn't use all features	Behavioural	

	Decrease in use over time	3	I think I started to use it less. You you get the information you need and then you leave.	Micro engagement	Behavioural	
				<i>Decrease in use over time</i>	Behavioural	
	Minimal transfer of skills beyond app	1	But yeah, it's more, you know, I see something. Satisfies that need and then it's forgotten. I would have liked if I was able to to retain more of it rather than kind of being reliant on it.	Macro engagement	Behavioural	
				<i>Not much outside app</i>	Behavioural	
				<i>Once you have what you need, don't need app</i>	Behavioural	
				<i>Used skills beyond app</i>	Behavioural	

Appendix 5-G. Case Study 3: Number of participants involved over the study stages per site

Study stage	Timeline	Oxford	Imperial	Total
Recruited participants	Sep 2021 - Jan 2022	128	97	225
<i>Excluded pre Dora call</i>		7	6	13
<i>Patients excluded after call</i>		2	8	10
Eligible patients in study		119	83	202
Calls completed autonomously by Dora		114	81	195
Participants contacted to collect demographic and usability data	Oct 2021 - Mar 2022	104	76	180
Participants interviewed	Feb 2022 - Mar 2022	10	10	20

Appendix 5-H. Case Study 3: Independent thematic framework

Theme	Brief description	TFA constructs	Codes
Expectations prior to the call	Most participants did not have any expectations about the Dora call; of the few that did, opinions were a mix of positive and negative.	Affective attitude (anticipated)	Negative No expectations or unsure Positive or no concerns
		Perceived effectiveness (anticipated)	Confidence No expectations Uncertain about Dora's ability
		Burden	Anticipated - could be frustrating
Ease of use	Almost all of the participants found the Dora call easy to use, requiring little effort, despite not all participants understanding how the technology worked.	Burden	Easy to use
		Intervention coherence	Understand call procedure Understand tech (at least broadly) Didn't have clear understanding
Usability issues or concerns	Some participants identified issues around knowing how to respond in a way that Dora would understand, feeling pressure to respond quickly, and what to do if they missed the call.	Affective attitude (experienced)	Uncertainty on how to respond What happens if call missed Not much time to think
		Burden	A bit too long Give Dora simple answers
		Self-efficacy	Low - difficult
Accessibility	Potential accessibility issues were identified around the ability of people with hearing difficulties or first languages other than English to understand Dora and its appropriateness for different types of patients (e.g. patients with autism).	Perceived effectiveness (experienced / long-term)	Didn't feel understood - e.g. Language or accent barrier Difficult or inappropriate for certain types of patients - accessibility
		Burden	Hearing and Dora's accent is a potential difficulty
Positive attitudes towards and confidence in Dora	The attitudes of most participants towards Dora were neutral to positive and many had confidence in Dora as an effective and satisfactory tool.	Affective attitude (experienced)	No concerns Good experience Neutral experience (fine) - no particular likes or dislikes Reassuring to have 'safety net'
		Self-efficacy	High - confident, comfortable expressing self
		Perceived effectiveness (experienced)	Perceived effective, satisfactory (confidence in Dora) Potential for harm - not likely
Importance of human aspect	Dora was not necessarily seen as appropriate for patients with complications or concerns because of the	Perceived effectiveness (experienced)	Dependent on complications, concerns
		Self-efficacy	Depends

	lack of human element and ability to address emotional needs and participants felt that it was important to have a human element in the pathway as a backup to catch any issues or support patients with more complex situations.	Opportunity costs	Lack of human element (empathy, reassurance, etc)
		Perceived effectiveness (experienced)	Dora as first step - clinician as backup is important
		Affective attitude (experienced)	Perfunctory, brisk Impersonal, remote feeling Bad experience
		Ethicality	Can't treat people as objects - need human aspect
Concerns about Dora's effectiveness	Participants also raised concerns that Dora's structured delivery risked missing important clinical nuances that would be captured in a conversational setting with a clinician where patients can ask questions and clarify or qualify responses.	Affective attitude (experienced)	Cannot have two-way communication to discuss, question, clarify
		Affective attitude (experienced)	Covers current time point only Not taken seriously
		Burden	Had to interpret questions
		Perceived effectiveness (experienced / long-term)	Perceived NOT effective (no confidence in Dora) Potential for harm - possible Not good for long-term application
		Perceived effectiveness (experienced)	Concern that non-verbal info or nuances missed Hard to catch Dora's failures
Potential benefits more broadly	Many participants felt positively about or identified potential benefits of using more generally (time and cost savings primarily).	Perceived effectiveness (long-term)	Could have benefits if used more extensively (cost saving, data analysis, minimise human shortcomings, saves time, etc.)
Willingness to use Dora	Patient preferences for Dora or face-to-face appointments were mixed, but almost all of the participants would be willing to use Dora again (with varying degrees of enthusiasm or reluctance).	Affective attitude (experienced)	Preference Dora or F2F
		Affective attitude (experienced)	Willing to use again
		Opportunity costs	Lost time Nothing gained Nothing lost except short phone time
Suggestions	Participants made several suggestions for how Dora could be improved in the future.	Other	Ability to add more info or detail Ability to schedule and reschedule Accessible to all Don't use it Monitor and update (keep it learning) More clarity - objective, questions Paper version Questions in advance Reminders System set up to include Dora with safety net of human

Appendix 5-I. Case Study 3: Thematic framework mapping factors to engagement components and patterns of engagement

Theme	TFA constructs	Codes	Pp t	Quotes	Eng. component	Pattern of engagement
Expectations prior to the call	Affective attitude (anticipated)	Negative No expectations or unsure Positive or no concerns	3	Well, I was worried in case it was a bit like press one, press two, all this kind of relationship when actually there is nothing... It could be frustrating. I was worried in case it could be frustrating, but it wasn't.	Affective Cognitive	Cognitive → Initial Behavioural (micro)
	Burden	Anticipated - could be frustrating	16	Well, I thought the rationale for it was very good really, because it's quite a interventionist procedure, a kind of cataract removal. So to have a source of advice, guidance, information, was a good thing. So I was expecting to give feedback and get information and felt positive.		Cognitive → Affective → Initial Behavioural (micro)
	Perceived effectiveness (anticipated)	Confidence No expectations Uncertain about Dora's ability	7	So I'm coming from a position of a lot of confidence with anything that comes out of the NHS really. I'm taking the position that anybody from the NHS contacting me with a thing like this is starting from a high bar. So I would have no reason to be suspicious of it really.	Cognitive Behavioural	Context → Initial Behavioural (micro)
			1	Well, I think I have to be honest and say I didn't really know what to expect... I wondered if Dora would understand my answers.		
Ease of use	Burden	Easy to use	14	It didn't take a great deal of effort. It was quite straightforward.	Cognitive	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
			15	So the liked bit is it's very quick and efficient.		
	Intervention coherence	Understand call procedure Understand tech (at least broadly) Didn't have clear understanding	1	So I think what I realized that Dora would've been programmed by a clinician to ask certain questions, but if there was a query about something, then Dora would not have had the wherewithal to further probe what that problem might be.	-	
			13	I don't know how you interpret the information or how the information is used there.		

Usability issues or concerns	Affective attitude (experience d)	Uncertainty on how to respond What happens if call missed Not much time to think	13	I think it is quite precise, but like I say, I couldn't quite know like from what perspective is she answer the question for me. <i>So there was a bit of uncertainty than it sounds like about how best to respond, is that right?</i> Yeah, correct.	Affective Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Affective
			9	If you were told you only had two seconds to answer everything or two seconds to write something down, it's all instant, instant, instant, you wouldn't, I feel get the right picture, but that's the sort of thing that was happening with Dora, you were just... doing it, doing it, doing it, answering.		Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
	Burden	A bit too long Give Dora simple answers	5	Well, obviously just perhaps a bit slower, and simple words. Like talking to a toddler really, isn't it?	Cognitive Behavioural	
			3	Knowing it was a machine, you keep it simple and you go to the essential, which is actually a good thing.		
	Self-efficacy	Low - difficult	13	Quite difficult. You can't because I don't know if I'm talking to a person or a machine. A person, a machine, what is it, a computer?	Cognitive	
Accessibility	Perceived effectiveness (experience d / long-term)	Didn't feel understood - e.g. Language or accent barrier Difficult or inappropriate for certain types of patients - accessibility	18	She was very fluent in what she has to say, even though she had an accent. Because she had an accent, I was listening more deeply, just to get the t's and the l's. So yeah. Communication was okay.	Affective Cognitive	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
	Burden	Hearing and Dora's accent is a potential difficulty	11	I was extremely distressed by the whole experience. Now that may be because I'm autistic, and so I process things neurologically very differently from a neurotypical person. So if this Dora thing is asking questions, which from an AI point of view, it would be asking closed questions. They will only get, if they're on to an autistic person, they will only get a certain answer... and I think that's dangerous.		Initial Behavioural (micro) → Cognitive → Affective → Subsequent Behavioural (micro) Context → Behavioural (micro) Context → Cognitive → Behavioural (micro)

Positive attitudes towards and confidence in Dora	Affective attitude (experienced)	No concerns Good experience Neutral experience (fine) - no particular likes or dislikes Reassuring to have 'safety net'	20	Well, I probably would say no follow up, really. Oh, I don't know. Maybe Dora, just because that is reassuring, or would be reassuring, to some people I think.	Affective Cognitive	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro) Behavioural (micro) → Affective Behavioural (micro) → Cognitive → Affective
			8	Well, the questions she asked, I mean, I can't remember offhand, but it made me more comfortable after the operation.		
			6	I didn't really [have any concerns], because there's nothing to lose and a lot to gain if it actually works.		
	Self-efficacy	High - confident, comfortable expressing self	8	Well, it was quite easy. I felt comfortable with it, with her.		
	Perceived effectiveness (experienced)	Perceived effective, satisfactory (confidence in Dora) Potential for harm - not likely	10	She answered the questions that I asked to my satisfaction... It was good. It was clever.	Cognitive	
			17	I think it would work very well. It was clear. The question that he asked was clear and direct.		
			19	Harm? I don't think so... can't see any reason why it would harm anybody		
Importance of human aspect	Perceived effectiveness (experienced)	Dependent on complications, concerns	1	I don't know how good it would be if you had a very nervous patient or a very worried patient, or somebody who had not expected the outcome that had actually happened post surgery.	Affective	Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro) Behavioural (micro) → Cognitive → Affective Behavioural (micro) → Affective Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
			11	I cannot for the life of me, see how a robot can possibly be of any use in talking, well, I use the term loosely, in talking to patients. The situation is far too complex for artificial intelligence, and I use the word intelligence very loosely as well.		
			2	But my own view is that in probably 90% of cases, the end result is perfectly satisfactory for the people who've used it, but there may be 10% where you're not necessarily getting from it the full picture.		
	Self-efficacy	Depends	15	Partially. So as it happened, everything was fine. So I felt able to tell it that. But if it hadn't been or if there was... If I'd had a worry, I don't think... I'm not sure I would've been reassured.	Affective	

Opportunity costs	Lack of human element (empathy, reassurance, etc)	10	If Dora had rung me up, and I'd had a cataract operation, and I had lost the sight in my right eye, how would Dora have coped with that? You could talk about the facts, but have they got enough emotion into the artificial intelligence to be able to be sensitive to people's feelings.	Affective	Context → Cognitive → Affective → Behavioural (micro)
		9	I thought it was fine, but it felt abstracted. It wasn't personal. It really came through that way.		
		11	<i>What do you think you lost or gave up by having a call with Dora instead of an in-person follow-up assessment?</i> Humanity.		
Perceived effectiveness (experience d)	Dora as first step - clinician as backup is important	3	Rather than wasting time to have an appointment to go to a hospital just to be told that you don't have any problems. One, you already know that you don't have any problems. I think that can also be useful, provided that you know that in case you have problems or in case you want to talk to a human, being seen that's still available.	Cognitive	
		16	And for Dora to develop to that degree of sophistication might not be cost productive, but I mentioned [my complication] because I think it's quite a good example of why we still need readily available clinician backup.		
Affective attitude (experience d)	Perfunctory, brisk Impersonal, remote feeling Bad experience	10	I always feel a bit remote and lonely when I'm talking to a machine but yeah, I mean, I think that this is the way things are going.	Affective Behavioural	
Ethicality	Can't treat people as objects - need human aspect	7	Well, almost like talking to a person, but not quite. It didn't quite have the warmth of a conversation with a person.		
		11	Yes, I do [think it could harm]. I was extremely distressed by the whole experience... I think when you are dealing with health and patients, I don't think you can put them through a tick box exercise because it requires a level of being human, I suppose.		

Concerns about Dora's effectiveness	Affective attitude (experience d)	Cannot have two-way communication to discuss, question, clarify	14	Well, you're more likely to be able to develop a two-way exchange if you are speaking to a person.	Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
			2	I'll give you one minor example. When I had the first eye done... I could see a slightly brownish area in my vision. My thoughts were, I don't think this is anything to worry about... but face to face, I might have raised that. Whereas not all the questions really seemed to encourage me to include that in it.		Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
	Affective attitude (experience d)	Covers current time point only Not taken seriously	1	I personally had no concern about talking to Dora, but I could understand that if Dora was going to replace a face-to-face interview with the consultant or one of his team, then I might feel a bit concerned in case my worries or my fears weren't taken seriously.	Affective Cognitive Behavioural	Affective → Subsequent Behavioural (micro)
			15	Or something that isn't happening now, but did happen and you're worried it might happen again so that... So there isn't a yes or no answer, because it's like, again, back to the redness, "Well, no, [my eyes are] not red, but they were and I'm a bit worried about it. And do you think... Dora, is that going to be all right, Dora?" You know?		Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro)
	Burden	Had to interpret questions	4	And if I didn't understand it, and I interpreted what I thought the question might have meant.	Cognitive	
	Perceived effectiveness (experience d / long-term)	Perceived NOT effective (no confidence in Dora) Potential for harm - possible Not good for long-term application	2	When you've got a person in front of you, you can express a question in a particular way and the person can analyze that and give you an answer. AI relies on the programming of the AI responses. It's not as flexible as a human is.	Affective Cognitive Behavioural	
			4	I didn't think that I had all that much confidence. I wanted to ask the questioner who was a machine, not a person, what do you mean by that? I couldn't do that.		
			11	Personally, on a scale of one to 10, with 10 being excellent and zero being completely useless, I'd have to put it at zero. I think if you've been through a surgical procedure, you do NOT want to be contacted by a robot.		

	Perceived effectiveness (experienced)	Concern that non-verbal info or nuances missed Hard to catch Dora's failures	1	Because when you speak face-to-face, or when you're talking to somebody, fear or worry or whatever, might either be visible on your face or in your voice. And that obviously didn't apply to Dora.	Affective Cognitive	
			16	My main criticism, and I suppose it's inevitable in an automated system, it doesn't provide some nuances and shades of opinion. I was slightly nonplussed, upset, when it said about this irritation or, like I've told you, just to the eyes, yes or no. And I said I'd had it intermittently.		
Potential benefits more broadly	Perceived effectiveness (long-term)	Could have benefits if used more extensively (cost saving, data analysis, minimise human shortcomings, saves time, etc.)	2	I'm delighted that they're doing a proper follow-up...There are all sorts of operations that happen in hospitals and I wonder if eye surgery's the only one where Dora is useful as a followup. It probably could be useful in other situations.	-	-
Willingness to use Dora	Affective attitude (experienced)	Preference Dora or F2F	16	Well, if I felt things were going okay, and I didn't want the hassle of being in the hospital or going down there, I settle for Dora.	Affective Cognitive Behavioural	Initial Behavioural (micro) → Cognitive → Subsequent Behavioural (micro)
			17	I feel more comfortable talking to [Wysa] than a person telling the truth. When you're talking to a person, sometimes you have to answer the question over again. [Wysa'a] straight, clear and direct.		Context → Behavioural (micro)
			19	I think face to face would be better for me, I get lost in between, sometimes I don't understand the questions, if I'm face to face I'm able to ask to explain it to me better.		Initial Behavioural (micro) → Affective → Subsequent Behavioural (micro)
	Affective attitude (experienced)	Willing to use again	4	I would tolerate it, if that's the only way I could make a contribution after the operation.	Affective Behavioural	
			6	Well, I've got another [operation] this month. So if she wants to ring me up after having the op, fine.		
			13	Happy to use DORA again, yeah.		
	Opportunity costs	Lost time Nothing gained Nothing lost except short phone time	1	I didn't really lose anything except for 20 minutes or so of a telephone call.	Cognitive Behavioural	

Suggestions	Other	Ability to add more info or detail Ability to schedule and reschedule Accessible to all Don't use it Monitor and update (keep it learning) More clarity - objective, questions Paper version Questions in advance Reminders System set up to include Dora with safety net of human	3	So to make sure that if you miss the call, how to alert the system that you are available.	Affective Cognitive Behavioural	Context → Behavioural (micro)
			16	See if the advice responses, follow-up questions could be developed to make it even more comprehensive.		Context → Cognitive → Affective → Behavioural (micro)
			2	To be honest with you, I wonder if a paper version of Dora would be perhaps slightly more effective because people would have more time to actually think about the answers they want to give.		

Appendix 6-A. Thematic synthesis framework

*Yellow = Case 1: NoObesity, blue = Case 2: Wysa, green = Case 3: Dora

Analytical theme	Descriptive theme	Subtheme / Code / TFA construct	Codes / Sub-codes / Codes	Factor summary
1) Connection with the intervention	Tone	Liked NoObesity name	Medical name is good Opinion of name? (2) Prevent from downloading? (0)	Negative / positive connotations
		Issues with name	Can see why others wouldn't like Make name more positive Negative connotations Opinion of name (0,1) Prevent from downloading (1,2)	
		Make app more positive	Positive feedback / encouragement Focus on feelings not measurements	
		Liked family aspect	Liked family involvement / doing together	Doing as a family
		Tone	Positive, genuine, non-judgmental Tone could be misinterpreted as condescending	Tone
		Liked mascot (approachable)		Tone
		Not convinced about mascot		Tone
	Sense of connection	Perceived effectiveness - exp.	Different from discussing with a person	Capacity of chatbot for communication
		Lack of personal connection		Connection
		Perceived effectiveness - exp.	Confidence decreased with chatbot Inconsistencies from app / chatbot doesn't listen Limited / predictable / repetitive algorithm	Capacity of chatbot for communication
		Opportunity costs	Lack of human element (empathy, reassurance, etc)	Human element
		Perceived effectiveness (exp.)	Dora as first step - clinician as backup is important	Human element
		Affective attitude (experienced)	Perfunctory, brisk Impersonal, remote feeling Bad experience	Connection with intervention
		Ethicality	Can't treat people as objects - need human aspect	Human element

		Burden	Anticipated - could be frustrating	
	Concerns (broader implications)	Weight / health measurements	Discomfort tracking kids weight Discomfort tracking own weight	Discomfort tracking weight
		Accessing help without bothering anyone		Burden on others
		Anticipated affective	Worried about lack of support on wait list	Need/desire for mental health support
		Relief		Expectations of benefit
		Affective attitude (experienced)	Cannot have two-way communication to discuss, question, clarify	Quality of communication / conversation
		Affective attitude (experienced)	Covers current time point only	
		Affective attitude (experienced)	Not taken seriously	
		Perceived effectiveness (exp.)	Concern that non-verbal info or nuances missed	Confidence in Dora
2) Usability of the intervention for the target population	Usability	Good usability	Easy to use (1 - goal setting) User friendly / easy to use (2) Kid friendly (2)	Ease of use
		Poor usability	User friendly / easy to use (0,1) Kid friendly (0,1) Differences within family	
		Improve usability	Interface redesign Date of birth rather than age Back button functionality Intro / guide to use	
		Barriers to app use	Child's age Covid-19 Device	Inability to use
		Lack of clarity	Points not clear / need more meaning Unclear how it would work Confused about who it was for	Lack of clarity / guidance

		Lack of guidance	Lack of guidance (usability) Easy to use (0 - goal setting) Guidance (0 - goal setting) No help function	
		Specific complaints	Specific complaints Technical difficulties Specific complaints (games) Can't change inputted data Inaccuracies	Technical difficulties
		Made inputs simple for chatbot		Capacity of chatbot for communication
		Easy to use		Ease of use
		A bit overwhelming		Amount / presentation of content
		Categorisation might hide useful content		
		Confident in ability to use		Ease of use
		Burden	Easy to use	Easy to use
		Burden	A bit too long Give Dora simple answers	Ease of use
		Burden	Had to interpret questions	Ease of use
		Other	More clarity - objective, questions	Suggestions
		Self-efficacy	Low - difficult	Confidence in ability to use
		Self-efficacy	High - confident, comfortable expressing self	
		Self-efficacy	Depends	
	Accessibility	Barriers to behaviour	Accessibility Other health conditions	Lack of accessibility / affordability
		Liked accessibility		Accessibility of support
		Perceived effectiveness (experienced / long-term)	Didn't feel understood - e.g. Language or accent barrier Difficult or inappropriate for certain types of patients - accessibility	Feeling understood, accessibility

		Burden	Hearing and Dora's accent is a potential difficulty	Feeling understood, accessibility
		Other	Accessible to all	Suggestions
		Other	Questions in advance	Suggestions
		Affective attitude (experienced)	No concerns Good experience Neutral experience (fine) - no particular likes or dislikes Reassuring to have 'safety net'	Experience on call
		Affective attitude (experienced)	Uncertainty on how to respond What happens if call missed Not much time to think	Concerns during call
		Perceived effectiveness (exp.)	Hard to catch Dora's failures	Confidence in Dora
	Personalis-ability	Elements liked in Family Survey	Personalisation (1 - family survey) Inclusion of other healthy behaviours Liked oral health addition	Personalisation
		Increased personalisation	Personalisation Separate goals	
		Create a personal library	Video not always convenient - other suggestions	Personalisation
		Other	Ability to add more info or detail Ability to schedule and reschedule Paper version	Suggestions
3) Attention and incentive to use	Rememberin g to use	Issues with notifications	Daily notifications too much Didn't take much notice of notifications Noise of notification frustrating Notifications came at bad times	Notifications
		Improve notifications	Prompts / reminders Targeted notifications	
		Not memorable	Liked recording progress (0) Not memorable (family survey) Forgot to keep up	Novelty / variety
		Micro engagement	Easy to ignore / nothing grabs attention Reminder prompted use Tailoring reminder times would be good	Reminders

		Other	Reminders	Suggestions
	Ability to sustain interest / enjoyment	Not enough variety - repetitive	Repetitive (games) Repetitive (goal setting) Simple goals (0)	Novelty / variety
		Not enough variety - limited content	Limited things to do Liked recording progress (0) Boring Nothing new Mostly used to record progress Not realistic (healthy swaps) Looked for suggestions / ideas outside of app Liked healthy eating swaps (0) Flexibility (0 - goal setting)	Novelty / variety
		Anticipated affective	Curious about trial / new tool Excited about using	Curiosity and interest
		<i>Perceived effectiveness - exp.</i>	<i>Confidence decreased with chatbot</i> <i>Inconsistencies from app / chatbot doesn't listen</i> <i>Limited / predictable / repetitive algorithm</i>	<i>Capacity of chatbot for communication</i>
		Tedious		Interest
		Frustration with chatbot		Frustration
		Micro engagement	Issues/frustrations with chatbot affected use	Frustration
	Context of use	No costs		Time required
		Short videos with timings were good		Time required
		Micro engagement	Don't always have time Not using regularly Used in quiet moments of day Used to satisfy immediate need Built into daily routine Lifestyle and other factors could affect use	Time required Context of use
		Macro engagement	Not much outside app Once you have what you need, don't need app Used skills beyond app	
		Opportunity costs	Lost time Nothing gained Nothing lost except short phone time	Time required

4) Confidence in the intervention	Perceived effectiveness and credibility	Concerns with linking with HCP	Communicating via app (0) Privacy / security concerns	Privacy / security concerns
		Privacy concerns associated with what's shared with app		Concerns about app
		No concerns		Concerns about app
		Trustworthy		Credibility / trustworthiness
		Trusted because NHS / sources cited	Experience of it helping improves credibility	Credibility / trustworthiness
		Perceived effectiveness - anticipated	Expected emotional benefits / guidance High expectations of chatbot No particular expectations of ability - not expecting the world	Expectations of benefit
		Perceived effectiveness (exp.)	Perceived effective, satisfactory (confidence in Dora) Potential for harm - not likely	Confidence in Dora
		Perceived effectiveness (exp.)	Dependent on complications, concerns	Confidence in Dora
		Perceived effectiveness (experienced / long-term)	Perceived NOT effective (no confidence in Dora) Potential for harm - possible Not good for long-term application	Confidence in Dora
Other	Features supporting engagement and self-efficacy	Poor acceptability	Kids liked games (0) Participants liked games (0) Liked goal setting (0,1) Aesthetic (0) Poor acceptability	Aesthetic appeal
		Feedback was lacking	Feedback useful (0) Lack of feedback Liked points / rewards (0) Notifications useful/lacking? (0)	Feedback / visualisations
		Improve self-monitoring feature	Trends (incl. retrospective review and recording, visualise progress) Simplify recording progress Add notes	

	Liked games	Kids liked games? (1) Participant like games? (1) Educational	Interactivity / gamification
	More engagement for kids	More family / kid engagement Competition would motivate kids Points / rewards could be motivating (esp for kids)	
	Liked goal setting	Flexibility (1) Goals helped with planning Liked goal setting? (2) Simple goals Thinking about barriers was useful	Goal setting support
	Liked app feedback	Notifications useful / lacking? (1) Feedback useful? (1) Liked points trophies? (1)	Useful feedback
	Liked app feedback	Notifications useful / lacking? (1) Feedback useful? (1) Liked points trophies? (1)	Notifications
	More content and more variety of content	More variety More information / education Home workouts Local activities Ways to overcome personal barriers Recipes / meal planning / shopping support Vegetarian / vegan options	Novelty / variety
	More content and more variety of content	More variety More information / education Home workouts Local activities Ways to overcome personal barriers Recipes / meal planning / shopping support Vegetarian / vegan options	App suggestions (for behaviours)
	Suggestions were helpful	App supported self-efficacy? (1) Overcoming barriers helps with self-efficacy Suggestions supported self-efficacy (1) Liked activity suggestions? (1) Liked healthy eating swaps? (1) Liked parent survival guide? (1) Links to other resources were useful Goals helped with planning (1) Liked goal suggestions / found helpful Guidance (1 - goal setting)	App suggestions (for behaviours)
	App suggestions helpful - keep fresh		Suggestions
	No incentive (esp. if emotional)		Incentive

		Macro engagement	Want more training for using skills outside app	Suggestions
		Intervention coherence	Understand call procedure Understand tech (at least broadly) Didn't have clear understanding	Understanding of how call works
		Other	Don't use it Monitor and update (keep it learning) System set up to include Dora with safety net of human	Suggestions

Appendix 7-A. Positionality and reflexivity statement

It is important to acknowledge how my personal characteristics and perspectives have shaped this research. I am a doctoral candidate and research fellow in digital health. I have been working in the digital health field for nearly 6 years, but my academic background is primarily in psychology, with an emphasis on interdisciplinary research. During my previous university education, I conducted research around emotion, decision-making, and behaviour, while my current work as a research fellow exposed me to a variety of means by which digital technology is employed with the aim of improving healthcare. This background drove my interest in the psychology behind digital health engagement and my desire to understand how affective and cognitive aspects could influence what we observe as behavioural engagement with interventions or target health behaviours.

Over the three case studies I conducted, I interacted with a variety of participants with different characteristics and backgrounds who shared the key features of being UK-based (1) parents concerned about childhood obesity, (2) people seeking mental health support, or (3) older adults who had undergone cataract surgery. For the first case study, I was aware that - as an active young woman without children, who has never struggled with food insecurity or other structural barriers to healthy behaviours, and who is knowledgeable about the psychology around behaviour change - my experiences and perspectives on childhood obesity and healthy eating and physical activity behaviours would likely be different from many of the study participants. As my background has focused on psychology, which inevitably includes a focus on the individual, I was careful to ensure that structural factors and barriers identified by participants were clear in my analysis. This was supported through the use of the Capability, Opportunity, Motivation - Behaviour model to structure the analysis, which ensured that the 'opportunity' or contextual aspect received equal consideration.

Both the first and second case studies focused on sensitive topics that can experience social stigma - childhood obesity and mental health conditions. In both cases, the interviews focused on engagement and experiences with the app, but through this, some participants shared reflections on stigma and challenges related to healthy behaviours and experiences with mental health. I was aware during the interviews of the importance of not being or seeming judgmental about anything participants might share.

The third case study also highlighted another characteristic which shaped my interactions with participants and possibly also my interpretation of the data: I am Canadian. This

influenced the third case study particularly as my interactions were over the phone with older participants, many of whom had minor hearing difficulties. Several participants found my accent difficult to understand, which might have influenced their level of comfort in the interview. More broadly across the three case studies, although I have been living in the UK for nearly 6 years, participant responses may have included cultural phrasing or meanings that I may have missed or misinterpreted.

Another factor that I considered during the interpretation of the data was the influence of the context of the study. In cases 2 and 3, funding for the studies was won in collaboration with digital health companies. Although we conducted the evaluations independently and I was clear in my role of producing an accurate assessment, I was aware that the companies would desire a positive conclusion about their product. To mitigate this as a potential influence on my interpretation, I made a particular effort to identify negative feedback, which also provides valuable insights for improvement.

Another influence on my interpretation of the data is my background knowledge about affective and cognitive processes and my own experiences engaging with digital health apps. In my reflection on why I observed a difference between the hypothesised and empirical patterns of engagement (affective → cognitive vs. cognitive → affective), I suggested that one reason could be how participants find it most logical or clear to report their experience (identifying what caused particular emotional experiences); this could be equally true of my interpretation. This was one of the main reasons why I conducted a re-analysis of the individual within-case mappings and the cross-case synthesis several months after I initially conducted them. On review with fresh eyes, I realised that I had not been sufficiently transparent about how I had outlined the patterns of engagement reported in Chapter 6, which were aligned with this potential bias. I held this potential bias in mind when re-analysing the data and aimed to clarify how I had arrived at the patterns by including more detailed thematic frameworks with participant quotes, so that readers could determine for themselves whether they agreed with my interpretation. In selecting participant quotes, I aimed to include a diverse set from various participants with different perspectives to counter any unintentional focus on quotes that may have resonated more with me personally.

In this statement, I have made an effort to highlight key areas where my personal characteristics and perspectives may have influenced my research and how I have reflected on these implications.