Understanding the Role of Hunger in Dietary Cognition and the Food Insecurity-Obesity Paradox

Courtney-Jai Georgie Neal

A thesis submitted for the degree of Doctor of Philosophy

Population Health Sciences Institute, Faculty of Medical Sciences

Newcastle University

June 2024

Abstract

Food insecurity (FI) can be defined as limited or uncertain access to food of adequate nutritional quantity or quality and is often assumed to cause increased hunger levels. FI is associated with an increased risk of obesity in women in high-income countries. In recent years, the prevalence of FI has increased in higher-income countries where diet-related disease burden remains a significant public health challenge. Therefore, it is important to characterise the experience of hunger in FI and explore the potential effects that hunger has on the psychological processing of food to better predict the downstream effects on eating behaviour. I first report the results of a descriptive ecological momentary assessment study (EMA) that investigated the experiences of hunger across the day in two groups of women: one experiencing food security (FS) and one FI (Chapter 2). There was no difference between groups in average hunger and hunger variability within a day. However, in the FI group, there was greater variation in daily average hunger and within-day hunger variability between days. I then present two pieces of research focusing on how acute hunger impacts food-related cognition, as these mechanisms may underpin decision-making processes during food selection. In Chapter 3, I report two conceptual replications of a well-cited paper, which found that participants experienced greater attentional capture by food cues when hungry than when they were sated. However, the effect was not replicated. In Chapter 4, I explore the impact of hunger on the memory of food-related information in two studies, which comprised an image recognition and a price recall task. Hunger did not impact the memory of food-related stimuli in either study. Finally, I discuss the broader implications of my findings and consider future research directions (Chapter 5).

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor, Gillian Pepper, for believing in me and taking the time to help me grow. Without your wisdom, mentorship and convenient connections to Daniel Nettle, I never would have made it past the application stage for this doctoral programme, let alone to the end of it. I would also like to thank Daniel Nettle for trusting Gillian's judgement and accepting your role as my primary supervisor. I know I am an infinitely better researcher because of your patience, guidance, and "critical" feedback on my written work. I am very grateful to Caroline Allen for being exactly the supervisor I needed exactly when I needed her, especially in the early days of my PhD. As it takes 5-a-side team to get me through a PhD, I would also like to express my deep gratitude to my supervisors who arrived late to the party, Oliver Shannon and Melissa Bateson; your insights, enthusiasm and willingness to support me have made me a more confident researcher and challenged me to be a better scientist. I would also like to thank Quoc Vuong and Peter Gallagher for the encouragement and advice they generously provided as progress panel advisors.

To my adopted lab-mates past and present, thank you for providing such a Wellcoming, accepting and fun environment to work in. Ben Slater, Cat Pattie, Beth Little, Vera Vinken, Farhana Williamson, Chris Miller, Beshoy Agayby, Jen Nacef – when the going got tough, you kept me going.

I am forever grateful for the support of my parents, family and friends, particularly for the words of encouragement and acts of kindness that have helped see me over the finish line. I am especially grateful to Cassie Bakshani for showing me how it is done, being my voice of reason, and for motivating me to keep going in the face of adversity. *We just don't stop.*

Isaac Parker, thank you for being my champion and for your care, patience, and support over the last decade to help me achieve my goals. None of this would have been possible without you and the fluffy Pigg that lives rent-free in our house.

Lastly, I would like to express my gratitude to the hundreds of kind people who have participated in my research. This thesis would not exist if not for your enthusiasm, generosity and willingness to get hangry in the name of science.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AB	Attentional blink
BF	Bayes factor
CR	Correct rejection
EBA	Emotional blink of attention
EMA	Ecological momentary assessment
E1	Experiment 1 (*Chapter 3*)
E2	Experiment 2 (*Chapter 3*)
FA	False alarm
FI	Food insecurity/Food insecure
FS	Food security/Food secure
H	Hit
M	Miss
pd	Probability of direction
PPZ	Piech, Pastorino & Zald (2010)
ROPE	Region of practical equivalence
RSVP	Rapid serial visual presentation
SDT	Signal detection theory
S1	Study 1 (*Chapter 4*)
S2	Study 2 (*Chapter 4*)
USDA	United States Department of Agriculture
VAS	Visual analogue scale

Chapter 1 Introduction

1.1 Thesis Context

1.1.1 Obesity and Public Health

Obesity is a growing public health problem (Blüher, 2019). It is a disease that results from a complex web of relationships between biological, psychosocial and behavioural factors, such as genetics and socioeconomic status (Apovian, 2016). Many behavioural determinants of obesity are preventable, including diet quality and levels of physical activity (Hruby et al., 2016). These health behaviours are influenced by physical and digital food environments, as well as the built environment (Atanasova et al., 2022; Bennett et al., 2024; Swinburn et al., 2011). Framing obesity as an individual problem rather than a consequence of the environment is problematic, as resulting obesity policies and interventions are stigmatising to those living with obesity and are more likely to perpetuate existing health inequalities (Hill et al., 2021). For example, dietary behaviour change interventions requiring a high level of individual agency are less effective for people with lower socioeconomic status, while interventions which target "upstream" dietary influences (e.g., price) minimise inequalities (McGill et al., 2015).

A recent analysis found that, between 1992-2020, obesity policies in England placed high demands on an individual's agency (Theis & White, 2021). Such policies rely on individuals to change their behaviour instead of addressing systemic influences on behaviours of interest. However, there are also examples of existing policies targeting the obesogenic environment to improve dietary behaviour, both in the UK and abroad. These include but are not limited to the taxation of high-sugar products, the introduction of front-of-package nutrition labelling, and adding calorie information to menus (Malik et al., 2020; Rincón-Gallardo Patiño et al., 2020). However, these approaches have had limited success in changing the direction of current trends in population-level weight status in the UK and globally (Blüher, 2019; Finucane et al., 2011; Theis & White, 2021). Almost 60% of adults in Europe have overweight or obesity, but none of the Member States in the World Health Organisation European Region are on track to hit the 2025 target of halting increasing rates of obesity (WHO, 2022).

Recent developments in drug treatments, such as GLP-1 receptor agonists (Wang et al., 2023), are promising and may significantly impact the prevalence of obesity. However, patients require ongoing treatment to prevent weight regain following successful weight loss using GLP-1 receptor agonists, but the long-term effects of treatment are currently unknown (Wilding et al., 2022). Furthermore, even after controlling for cost, inequalities in access to treatment persist, with minoritised ethnic groups and patients with lower incomes less able to access treatment (Eberly et al., 2021). Recent increases in demand have also led to ongoing drug shortages, introducing additional barriers to patients wanting to initiate treatment and disrupting the care of those already being treated (Whitley et al., 2023). Given these challenges, pharmacological advances are unlikely to be the sole solution to the obesity epidemic. Complementary strategies of treatment and primary prevention (i.e., preventing the development of obesity) are required. To effectively target preventable determinants of obesity, we must first deepen our understanding of the psychological, biological, and social factors contributing to weight gain and limiting weight loss. Doing so will allow us to design successful solutions to reduce population weight status and increase healthy life expectancy.

1.1.2 Food Insecurity and Obesity

One factor that could be relevant to the prevalence of obesity is food insecurity (FI). FI can be defined as “limited or uncertain availability of nutritionally adequate and safe food or limited or uncertain ability to acquire acceptable foods in socially acceptable ways” (USDA Economic Research Services, 2023a). In recent years, the prevalence of FI has significantly increased in wealthy Western nations (e.g., UK, USA, Canada), likely due to cost-of-living increases linked to the COVID-19 pandemic, food shortages and high energy prices (FAO, 2023; The Food Foundation, 2024).

For women in developed nations, FI is robustly associated with obesity (Laraia, 2013; Nettle et al., 2017). In these populations, the odds of high body weight are 50% higher for women with FI than those with FS (Nettle et al., 2017). It is worth noting that although the literature often refers to differences in weight status under conditions of FI between *men* and *women* (genders), it is more true to theoretical accounts to report differences between *males* and *females* (sexes).

Primarily, gender refers to socially constructed norms and roles, and biological sex refers to an individual's biological traits; neither are binary, as gender is a spectrum and biological sex includes those who are intersex (Kaufman et al., 2023). Differential responsiveness in adiposity to FI is considered a difference between sexes rather than genders because of differences in biological traits, such as endocrinology and energetic requirements for reproduction. In humans, around 27% of body weight in females can be attributed to adipose tissue, compared to around 15% in males, a difference typically ascribed to the greater energetic requirements of reproduction for females (Norgan, 1997; Zafon, 2007). However, the greater average adiposity in females does not explain their increased sensitivity in adiposity under conditions of FI. Ultimate explanations underpinning sex differences in weight sensitivity to FI remain unclear (Nettle et al., 2017). In this thesis, I will use the terms *females* and *males* instead of the more commonly used *women* and *men* when discussing differences in weight status in response to FI, as the terms more accurately reflect theoretical accounts. Additionally, doing so supports the calls for researchers to pay closer attention to distinctions between sex and gender in health research to improve gender equity in healthcare (Kaufman et al., 2023).

Because of the link between FI and obesity, the rise of FI in wealthy countries is widening health inequalities and contributing to the growing global obesity crisis, both of which are major public health challenges (Case & Kraftman, 2022; Lucero-Prisno III et al., 2023; Williams et al., 2015). The association between FI and obesity is often referred to as the "food insecurity-obesity paradox" due to the apparent incongruity between its two components (Dietz, 1995; Dinour et al., 2007); how can those experiencing FI have a higher weight status than those who have food security (FS) while also having insufficient access to adequate nutrition? When individuals habitually consume more energy than they use, they gain weight. Early reports of a relationship between FI and obesity suggested that FI leads to the increased consumption of energy-dense foods and subsequently heightens the risk of obesity (Dietz, 1995). Evidence supports this; compared to individuals with FS, individuals experiencing FI consume more energy-dense foods that are typically cost-effective, highly palatable, and of poor nutritional quality (Eskandari et al., 2022; Leung et al., 2014). Despite increased consumption of energy-dense foods, several studies on different samples have reported no difference in daily energy intake between groups experiencing FI and FS (Leung et al., 2014; Morselli et al., 2024; Shinwell et al., 2022). Shinwell et al. (2022) explored dietary correlates of FI in the UK and found that their participants with FI reported a less diverse diet and greater variability in meal timing. The authors concluded that there may be metabolic and health consequences associated with the temporal irregularity in food consumption that they found to be characteristic of FI.

Evidence supporting changes in food choice, consumption and eating behaviour under conditions of FI only provides a possible proximate explanation for the link between FI and obesity. It does not offer an ultimate explanation, i.e., it tells us *how* FI may lead to obesity, but it does not tell us *why* individuals experiencing FI gain weight. Researchers who investigate the FI-obesity paradox from an evolutionary perspective propose that increasing adiposity may be an adaptive response to the uncertainty of food availability under conditions of FI (the insurance hypothesis; Bateson & Pepper, 2023; Nettle et al., 2017). Increasing body fat stores may improve short-term survival by acting as an energetic buffer in an environment in which food acquisition is unpredictable. Early experimental evidence supporting the insurance hypothesis came from the study of birds, which showed plasticity in fat storage within individual birds in response to unpredictable access to food (or FI, as described in the present thesis; Ekman & Hake, 1990; Witter & Swaddle, 1997).

To my knowledge, there is no published literature which attempts to experimentally create FI in humans – by manipulating the predictability of access to food – to study its effects on fat storage, behaviour and health. Manipulating FI in human research is ethically challenging and difficult to conduct safely. It is worth noting that I am involved in an ongoing pilot study which mimics the uncertainty of FI by manipulating the timing and predictability of eating events (based on the methods of Farshchi et al., 2004a, 2004b), but findings have not yet been published. Previously, researchers have tried to overcome the challenges associated with studying FI experimentally by controlling the study environment rather than the experience of FI itself. For example, researchers found that participants who had experienced FI prior to a four-day admission to a highly controlled metabolic ward had a higher respiratory quotient and lower levels of glucagon-like peptide 1 than participants who had not (Booker et al., 2022). These differences would likely ultimately contribute to increased energy balance (by promoting higher energy intake) and increased fat storage, highlighting that the metabolic effects of FI persist even when an individual is not living under conditions of FI at present. Another study from the same group supports increased energy intake in participants with FI (Stinson et al., 2018), contrary to findings from studies in free-living populations, as previously mentioned (Leung et al., 2014; Morselli et al., 2024; Shinwell et al., 2022). In this study, researchers recruited FI and FS participants to an eight-day highly controlled laboratory study and observed differences in how participants self-selected food from a vending machine for a three-day period. Participants who had previously experienced FI in their day-to-day lives consumed more energy, fat and carbohydrates during the vending machine paradigm, and consumed more energy across the three days than participants who had FS. Again, these findings could explain *how* FI leads to obesity but not *why.*

In place of FI, other researchers have studied the effect of experimentally induced subjective social status on energy intake and preference for energy-dense foods. Evidence showed that when individuals perceived themselves to be of lower socioeconomic status, they automatically exhibited a greater preference for more energy-dense foods and increased their energy intake (Cheon & Hong, 2017). Consequently, the authors proposed that – in social species such as humans – the psychological and physiological systems which regulate hunger levels may have evolved sensitivity to perceptions of deprivation in survival critical resources such as food. If true, the perception of food scarcity alone may be enough to alter behavioural and metabolic responses to the environment under conditions of FI, even without experiencing an energy deficit.

To summarise, as it stands, there is a wealth of observational evidence linking FI and obesity (e.g., Dinour et al., 2007; Franklin et al., 2012; Laraia, 2013; Larson & Story, 2011; Nettle et al., 2017), and limited experimental evidence linking FI to changes in dietary behaviour (Booker et al., 2022; Stinson et al., 2018), but there are no human experimental studies which identify causal pathways between specific components of FI and increased adiposity. Consequently, the biological causal mechanisms underpinning the association between food insecurity and obesity in females remain poorly described by scientific evidence.

1.1.3 Food Insecurity and Hunger

Exploring differences in psychological and physiological experiences in those with and without FI is a reasonable starting point to unpick the mechanistic link between FI and obesity without experimentally creating conditions of FI. For example, FI is often assumed to be associated with increased levels of hunger (e.g., Nettle, 2017). The United States Department of Agriculture (USDA) used to use the classification "Food insecurity with hunger" to describe households with very low food security, despite not asking specific questions about how hunger was experienced (USDA Economic Research Services, 2023a). To my knowledge, the hypothesis that FI is associated with increased hunger levels has not been tested.

The USDA describe hunger in FI as a physiological condition occurring due to resource constraints, leading to "discomfort, illness, weakness, or pain that goes beyond the usual uneasy sensation" (USDA Economic Research Services, 2023a). More broadly, hunger is defined as a universally experienced interoceptive sensation and psychological state associated with a desire or need to eat (Beaulieu & Blundell, 2021; Stevenson, 2023). It informs and reminds organisms that they need to find or eat food. From an evolutionary psychological perspective, hunger can be thought of as a coordinating mechanism that regulates psychological and physiological processes to source food to meet energetic demands (Al-Shawaf, 2016). As hunger is a drive that is predominantly biologically instigated, it can be considered a quantifiable biological trait in humans (Beaulieu & Blundell, 2021), as well as a predictor of willingness to eat and the quantity of food to be consumed (Stubbs et al., 2000). Distinctions are made between *homeostatic* and *hedonic* hunger. In a psychobiological framework of hunger, homeostatic hunger is associated with the processes linking an organism's energetic requirements to the behaviours which will satisfy those demands (Hopkins et al., 2017). In the same framework, hedonic (or reward-driven) hunger describes the drive to eat particular foods because of their palatability in periods of relative energy abundance (Lutter & Nestler, 2009). Both influence dietary decision-making, eating behaviour and diet-related health outcomes (Emilien & Hollis, 2017).

As nutritional insufficiency and uncertainty in meeting energetic requirements are the foundations of FI, I will be focussing on homeostatic hunger (henceforth, hunger) in this thesis. Because of the links between hunger, eating behaviour and diet-related health outcomes, quantitatively describing the experience of hunger in females with and without FI is an essential step towards understanding the link between FI and obesity. While it is established that hunger plays an important role in dietary decision-making, how it does so is less clearly defined. Understanding how hunger impacts dietary decision-making may elucidate possible causal links between FI, hunger, and obesity.

Many existing studies focus on general cognitive deficits resulting from hunger (Benau et al., 2014). Perhaps the most well-known example is the literature evidencing reduced cognitive function in children who do not eat breakfast (Hoyland et al., 2009). However, our understanding of the influence of hunger on cognition may benefit from incorporating a strengths-based approach, as suggested by Frankenhuis and Nettle (2020). In a strengths-based model, we would predict that hunger (a motivational state) will improve the processing of goal-relevant stimuli (i.e., food and information related to food acquisition, such as price). It is also possible that both strengths- and deficits-based models may be supported and simultaneously observable. For example, hunger may hinder cognition globally, but such an effect may be attenuated for stimuli relevant to food acquisition (Orquin & Kurzban, 2016). Cognitive enhancements or attenuations for food stimuli may be problematic for dietary decision-making in a modern food environment where exposure to and availability of food and food marketing is ubiquitous (Paquet, 2019). If people with FI differ in their experiences of hunger, they may be disproportionately impacted by such a mismatch between the environment and the cognitive processing of food. Therefore, we need to establish whether hunger improves the processing of stimuli relevant to food acquisition.

1.2 Thesis Aims and Overview

This thesis has two primary aims:

Aim 1: To quantitatively describe and compare the experience of hunger in females with and without FI.

Given the plethora of evidence supporting the links between FI and negative health outcomes, there is a public health policy need for research into the drivers of these associations. Of particular importance is the seemingly causal relationship between FI and obesity; examining how experiencing FI may lead to changes in energy balance is critical to understanding this relationship. Moreover, if one is to understand how energy balance changes under conditions of FI, with the ultimate aim of identifying a solution, one must also investigate how potential drivers of these changes are impacted by FI. As previously discussed, hunger is a well-evidenced regulator of dietary behaviour and energy intake. While FI has long been assumed to be associated with increased levels of hunger, there is little evidence to support the claim. Therefore, it is necessary to study how hunger is experienced under conditions of FI before one can hypothesise about how changes in hunger may drive observed changes in energy balance under conditions of FI.

Aim 2: To investigate the effect of hunger on food-related cognition.

The insurance hypothesis proposes that, under FI, increased adiposity in response to uncertain food availability is likely an adaptive response to cues of food scarcity. Thus, the effects of FI on adiposity may be strategic to improve the odds of survival by increasing the body's energy buffer. If so, one may also expect to see strategic shifts in sensitivity to food cues to improve the likelihood of food acquisition and aid positive energy balance. Given that hunger is assumed to be experienced differently under FI and that hunger is known to influence dietary decision-making, it is reasonable to hypothesise that experiencing FI may increase an individual's sensitivity to food cues. Such increased sensitivity may, in turn, impact dietary behaviour. However, there are only a handful of studies which show associations between hunger and food cue sensitivity (via several cognitive mechanisms). Therefore, before investigating the effect of FI on food cue sensitivity, it is important to establish whether hunger itself is what drives this sensitivity.

My thesis comprises three data chapters, Chapters 2-4. All data chapters and their associated data and analysis scripts are publicly available. Chapter 3 is published as a peer-reviewed article in *Appetite*. Chapters 2 and 4 are available as pre-prints on the Open Science Framework and have been submitted for peer review to *Appetite* and *Adaptive Human Behaviour and Physiology*, respectively. The versions included in my thesis have been peer-reviewed or are under review. I have not included a standalone literature review chapter in my thesis to avoid repetition, as relevant literature is reviewed in detail in the introduction of each data chapter.

1.3 Thesis Outline

In Chapter 2, I use ecological momentary assessment to investigate the differences in the experience of hunger in females with and without FI. While this was chronologically the last study I conducted during my doctoral programme, I have chosen to present it in my thesis first as its results highlight the importance and relevance of studying hunger to improve our understanding of socioeconomic differences in human health. Thus, the findings of Chapter 2 provide a strong rationale for the data chapters that follow it.

In Chapters 3 and 4, I adapt cognitive tasks to test the hypothesis that hunger is associated with improved cognitive processing of food stimuli. I begin by investigating the effect of hunger on attention in Chapter 3 before moving on to memory in Chapter 4, as attentional capture of a stimulus precedes the encoding of a memory of the stimulus. Simply put, one cannot memorise a presented stimulus without first attending to it. Thus, if hunger influences food decision-making via cognitive improvements for food stimuli, I would need to address whether this was due to enhancements in attention, memory, or both. Therefore, in Chapter 3, I present two conceptual replications of a study by Piech et al. (2010) whose findings suggest hunger increases the perception of food cues via an attentional mechanism. Contrary to the original findings, both replications produced null results. I then present two studies in Chapter 4 investigating whether hunger improves memory of food and food-related information. Study 1 of Chapter 4 uses old data (previously collected for my MSc Psychology thesis) reanalysed using different methods and statistical approaches (Bayesian, in place of frequentist). Study 2 uses new data collected during my PhD. As the development of Study 2 was directly informed by Study 1, the reanalysis of Study 1 is included in the present thesis. Finally, in Chapter 5, I summarise my learning, consider the implications of my thesis and reflect on my development as a scientist throughout my PhD. I end by discussing future research directions.

1.3.1 Impact of COVID-19

To support my doctoral training, I was awarded a studentship by the Northern Ireland and North East Doctoral Training Partnership, which is funded by the Economic and Social Research Council. My studentship was awarded in the 2020-2021 round, meaning that I submitted my research proposal for the studentship before the beginning of the COVID-19 pandemic. The research question in my original research proposal was "Does food insecurity affect food-related cognition?". However, when I started my doctoral studies in September 2020, students were not permitted on campus, and in-person data collection was prohibited due to the ongoing pandemic. Consequently, I had to redesign my research programme to work within the restrictions enforced by the university and the government.

As we did not know when restrictions might be lifted, I started designing and conducting exclusively online research. Additionally, I switched my focus to hunger instead of FI for the cognitive research. I made the switch for two main reasons: (i) my supervisory team advised me that I may struggle to recruit FI participants, as they had previously found recruiting FI participants difficult even without the additional complication of the COVID-19 pandemic, and (ii) the body of literature supporting a hunger-driven improvement in food-related cognition was relatively small and patchy, meaning my research could add value there. However, during my second year in 2022, FI was on the rise once more in the UK (The Food Foundation, 2024), and I was frustrated that I had not been able to conduct FI-related research. After speaking with my supervisors about my frustrations, we designed a novel study to quantitatively describe the experience of hunger in FI; this study became Chapter 2. While Chapter 2 deviates from the narrative that I originally proposed for my PhD and was not a study I initially planned to conduct, I believe its results have the greatest potential for impact and provide an excellent rationale for the cognitive research in the chapters that follow it.

Chapter 2 The Daily Experience of Hunger in UK Females With and Without Food Insecurity

Preface

Chapter 2 addresses the first aim of the present thesis, “To quantitatively describe and compare the experience of hunger in females with and without FI”.

To unpick the complex causal relationships that lead to increased adiposity in females with FI, we must develop a better understanding of the experience of FI. In particular, we must examine the psychological and physiological components of FI which may be linked to dietary behaviour, such as hunger. Despite the assumption that increased levels of hunger are characteristic of living with FI, to my knowledge, the assumption has never been tested using quantitative methods.

In Chapter 2, I investigate the day-to-day experience of hunger in females with FS and females who report having experienced FI in the recent past. I chose to do this with free-living participants to develop a dataset reflective of real-world experiences. Furthermore, I studied hunger across a week so that I could assess whether there were within-day differences and/or between-day differences in average hunger levels and hunger variation between the groups. The rationale for examining within- and between-day differences was developed in light of the evidence showing that females with FI eat in more temporally variable ways than females with FS (Shinwell et al., 2022). If eating is more temporally variable in females with FI, then one might predict hunger would be, too. Thus, studying temporal variability in hunger at different temporal scales (e.g., a day or a week) may offer a better opportunity to elucidate the experience of hunger in females living with FI, in line with the first aim of the present thesis.

Publication

This chapter is published as a pre-print and has been submitted to *Appetite* for peer review and publication in the upcoming special issue “*Food Insecurity, Obesity, and the Cost-of-Living Crisis*”. I have not changed it except to refer to other material in this thesis. Thus, I retain the use of “we” in place of “I” as I use elsewhere in this thesis. The citation information for the pre-print is as follows:

Neal, C., Pepper, G. V., Shannon, O. M., Allen, C., Bateson, M., & Nettle, D. (2024). *Pre-print. The daily experience of hunger in UK females with and without food insecurity.* Open Science Framework. https://doi.org/10.31219/osf.io/24zxe

Author contributions

I prepared each section of the text, including figures, with guidance on narrative and structure from DN. DN also provided critical feedback on all sections, as did GVP and OMS. MB and CA reviewed and provided feedback on the Introduction, Methods and Results sections. I led the development of the rationale and study design, with support and guidance on key decisions from DN. GVP and CA offered iterative peer review during weekly supervisory meetings throughout the study development period. I was solely responsible for participant recruitment, data collection and data capture. I designed and conducted the data analysis with analysis design support from DN, predominantly regarding the calculation of derived variables. MB also contributed to the analysis protocol by sense-checking the approach and commenting on data visualisations.

2.1 Abstract

Food insecurity (FI) is associated with increased mortality risk, depression, and obesity in females in high-income countries, but causal mechanisms remain unclear. FI is often assumed to lead to increased levels of hunger. However, quantitative evidence describing daily experiences of hunger in FI is lacking. Our pre-registered study used ecological momentary assessment to capture experiences of hunger in two groups of UK-based females: those experiencing FI (N = 143) and those experiencing food security (FS; N = 149). Participants self-reported hunger hourly (0900-2100) for one week (Monday-Sunday). There was no difference between groups in mean hunger ($t(290) = 0.17$, $p = .866$, $d = 0.02$) nor within-day standard deviation in hunger ($t(290) = 1.31$, $p = .193$, $d = 0.15$). However, both quantities fluctuated more from day to day in the FI group. Compared to the FS group, participants in the FI group had a larger day-to-day variation in mean hunger ($t(284) = 2.43$, $p = .016$, $d = 0.29$) and a larger day-to-day variation in the within-day standard deviation of hunger ($t(284) = 2.90$, $p = .004$, $d = 0.34$). In exploratory analyses, we found that the hunger of the two groups patterned differently across the day. Our findings suggest that experiences of hunger are less stable in those experiencing FI, likely reflecting the environmental uncertainty in access to food associated with FI and more variable meal timing.

2.2 Introduction

Food insecurity (FI) can be defined as "limited or uncertain availability of nutritionally adequate and safe food or limited or uncertain ability to acquire acceptable foods in socially acceptable ways" (USDA Economic Research Services, 2023a) and is often described using a continuum from mild to severe (Piperata et al., 2023). In recent years, there has been a global increase in the prevalence of moderate to severe FI (FAO, 2023). While this has predominantly impacted low- and middle-income countries, FI has also risen in the world's wealthiest nations. In the UK, 14.8% of all adults and 20% of all households with children are experiencing moderate to severe FI (The Food Foundation, 2024).

In the UK, there are regional inequalities in the risk of FI, with the North East having one of the highest rates of FI in England (Department for Work & Pensions, 2023). The increasing prevalence of FI is likely a driver of widening health inequality in the UK as FI is associated with a range of adverse health outcomes, including increased risk of mortality (Banerjee et al., 2021; Office for Health Improvement & Disparities, 2023), developing chronic disease (Laraia, 2013) and poorer mental health outcomes (Fang et al., 2021; Reeder et al., 2022). Furthermore, there may be sex differences in vulnerability to FI-associated health risks. Females are more likely to experience depression under conditions of FI than males, and females experiencing FI in high-income countries have higher odds of overweight and obesity than females who have food security (FS; Nettle et al., 2017, Reeder et al., 2022). To develop targeted interventions that tackle this growing public health challenge in affluent countries, it is imperative that we better describe the psychological and physiological experiences of those living with FI in this context.

One such experience is hunger. Hunger is an important consideration for health outcomes as it is closely associated with food consumption and overconsumption (Emilien & Hollis, 2017). Hunger caused by FI has been described as a physiological condition occurring due to resource constraints, leading to "discomfort, illness, weakness, or pain that goes beyond the usual uneasy sensation" (USDA Economic Research Services, 2023a). While FI is often assumed to be associated with increased levels of hunger, to our knowledge, this hypothesis has not been tested. FI may also be associated with greater variation in hunger, as FI is associated with more variable patterns of eating and longer gaps between eating (Nettle & Bateson, 2019; Shinwell et al., 2022). Therefore, we may expect higher average hunger and higher hunger variation in individuals with FI than FS. Furthermore, variation in hunger could operate on two scales: within a day, with hunger peaks being more extreme with FI, and between days, with days being more different from one another with FI.

To understand daily hunger patterns in individuals with FS, we can draw upon a wealth of literature concerning the fluctuations of self-reported hunger throughout the day. Hunger usually peaks twice daily, at approximately midday and 6 p.m. (Mattes, 1990; McKiernan et al., 2008). The timing and magnitude of these peaks are associated with physiological factors such as blood glucose levels (Ciampolini & Bianchi, 2006), circadian rhythm (Scheer et al., 2013), and energetic demands (Blundell et al., 2012). Peaks are also linked to key psychological drivers, such as memory and learning (for reviews, see Stevenson, 2023; Stevenson et al., 2023). However, to our knowledge, descriptive daily hunger data are not available in the context of FI. Therefore, little is known about differences in daily hunger patterns between those experiencing FI or FS and, if there are any, how these impact individuals' behaviour and health outcomes. Thus, describing the experience of hunger in FI is a critical step to untangling potential drivers of its associated negative health outcomes. For such data to be meaningful, they should be collected as individuals go about their day-to-day lives. Ecological momentary assessment (EMA) is designed for such scenarios, allowing repeated, frequent data collection remotely across a given timeframe. This has several benefits, including reduced reporting errors due to less reliance on retrospective memory (Lucas et al., 2021; Monk et al., 2015).

In this paper, we present a pre-registered EMA study which aimed to compare the experiences of daily hunger in two UK groups of females – a group experiencing FI and a group experiencing FS – over a week. We use sex and the term female(s), and not gender, in this study as sex differences in health outcomes associated with FI are of interest. The study aimed to compare within-day and between-day experiences of hunger between the groups. Doing this will allow for the development of testable hypotheses to inform policy decisions. Given that the definition of FI incorporates uncertain food availability, and previous assumptions and evidence suggest that those with FI experience higher levels of hunger and greater variation in meal timing, we predicted that participants experiencing FI would have:

P1. A higher mean hunger rating,

P2. Higher within-day variation in hunger rating,

P3. Greater between-day variation in average hunger rating,

P4. Greater between-day variation in within-day variation in hunger rating.

2.3 Methods

2.3.1 Study Overview

The Newcastle University Faculty of Medical Science Research Ethics Committee (reference 2287/18715) granted ethical approval for this study. We pre-registered our protocol (https://osf.io/zx5bj) and analysis plan (https://osf.io/ancx9; see Appendix A.1 for a summary of deviations from our pre-registrations). This study used EMA to assess and compare daily experiences of hunger in females experiencing FI and females experiencing FS. Individuals were identified as experiencing FI or FS after completing an online expression of interest form. Both groups of participants were blind to the FI component of the study and completed the same study procedure throughout.

Participants took part remotely as they went about their everyday lives. They downloaded and used a mobile app, Ethica (2023), on their smartphones to respond to hourly momentary assessments between 9 a.m. and 9 p.m. each day for seven days (Monday to Sunday).

2.3.2 Participants

We recruited 305 participants who were female or assigned female at birth (we aimed to recruit 150 participants per group, a similar sample size to that used by Dzubur et al., 2022). After removing participants with low-quality or insufficient data (see ***2.3.5 Data Analysis***), 292 participants (143 FI, 149 FS, ages 20-64 years, M = 38.2, SD = 7.2) were included in the final analyses; 118 were recruited from parent populations of schools in the North East of England with above-average percentage of pupils eligible for free school meals, 133 from social media using targeted (by geographical location) advertising to recruit parents in the same region, and 41 by word of mouth. Furthermore, of the participants included in the final analyses, 69 were classified as having “low food security” (moderate FI; mFI), and 74 as having “very low food security” (high FI; hFI; USDA Economic Research Services, 2023b).

2.3.3 Procedure

Recruitment and data collection took place between June 5th and November 26th 2023. Individuals completed an online form to express their interest in participating. First, we presented participants with study information and a consent form. We informed participants that, to take part, they should be over 18 years old, female (or assigned female at birth), have a smartphone, be able to download an app and commit to regularly responding to push notifications for seven days. To minimise potential confounding effects on hunger ratings, we also informed participants that they should not be following a weight management program or engaging in excessive physical activity or overnight shift work during the seven days of the study.

Next, participants provided their contact details, age, gender, height, and weight. They then completed the Household Food Security Six-Item Short Form module (USDA Economic Research Services, 2023b), answered questions about their employment status and daily eating and working routines, and completed the MacArthur Ladder (Adler et al., 2000) and Material Needs Scale (Conger et al., 1994; as used in Allen & Nettle, 2021). We then emailed participants to invite them to the study, reminded them of the study requirements, and provided instructions on downloading the mobile app. We instructed participants to set up the app on Sunday so that the notifications could begin at the appropriate time on Monday.

During the EMA period, participants received push notifications from the mobile app that prompted them to open the app and complete a survey (see Figure 2.1 for a timeline). There were three types of notifications that participants received:

Momentary assessments. These notifications prompted participants to complete a five-question survey (see ***2.3.4 Measures***). The mobile app presented the questions in a random order within the survey. If the participant did not respond, the notification and survey disappeared after 30 minutes. Otherwise, the survey disappeared once the participant completed it.

Daily waking time. These notifications prompted participants to complete a survey asking when they woke up that day. The mobile app delivered this at 9 a.m. each day, and it disappeared after the participant responded or at midnight. We do not report these data in the current paper as they were not used in the presented analyses.

Debrief questionnaire. This notification prompted participants to complete a debrief survey. This notification and survey remained until the participant responded or the study period (as programmed on the app) ended the following day. We do not report these data in the current paper as they are beyond its scope.

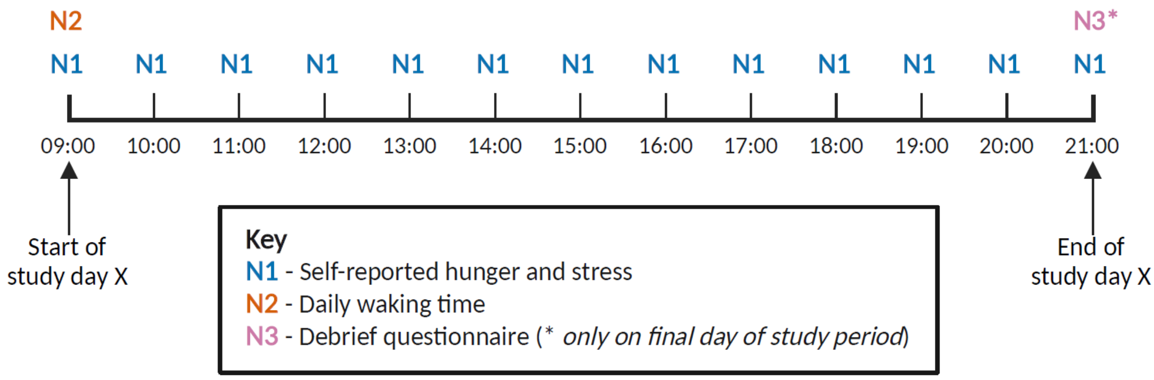


Figure 2.1. Timeline of push notifications for each day of the study period.

Furthermore, an “Additional Information” survey on the mobile app home screen was visible and accessible at all times. It allowed participants to submit additional information that they thought was relevant to the study or important for the researchers. We gave participants examples of how they might use this, including reporting why they had missed notifications or if they were feeling unwell.

At the end of the EMA period, we emailed participants debrief information and sent them a participation reward. Participants received a 10 GBP e-voucher for completing the study period and an additional 10 GBP e-voucher if they responded to over 70% of the momentary prompts. Participants with the top 10% of response rates were also entered into a lottery to win a 50 GBP e-voucher. To minimise missing data, we informed participants of these incentives during recruitment and contacted them during the EMA period if their response rate dropped below 70%.

2.3.4 Measures

Food insecurity status

In the online form at recruitment, participants completed the Household Food Security Survey Six-Item Short Form module (USDA Economic Research

Services, 2023b). We assigned participants to two groups in line with the module's scoring guidelines, as in Shinwell et al. (2022): FS (high or marginal food security, score 0-1) and FI (low/very low food security, score 2-6).

Momentary assessments

During the EMA period (N1), we asked participants, "*How hungry are you right now?*". They responded using a visual analogue scale (VAS) anchored at 0 ("*Not at all hungry*") and 100 ("*Extremely hungry*"). Blundell et al. (2010) recommend using this question in appetite research to assess hunger. We added "right now" at the end of the question to remind participants to report how they felt when they responded.

We also asked participants, "*Do you have the desire to eat something tasty right now?*". They responded using a VAS anchored at 0 ("*Not at all*") and 100 ("*Very much*"). Reichenberger et al. (2020) used this question to measure the intensity of the 'desire to eat' subcomponent of food craving (as in Cepeda-Benito et al., 2000).

Additionally, we asked participants three questions about their momentary perceived stress. We do not report these data in this paper (see Appendix A.2 for details of the measures).

2.3.5 Data Analysis

We analysed and visualised data in R (R Core Team, 2023). Our data and code are available at https://doi.org/10.17605/OSF.IO/6BP2Z.

We removed two momentary assessments from one participant completed before the sampling period at 8 a.m. In EMA research, a 75% response rate is often thought to be adequate, while a 90% response rate is considered excellent (Stone et al., 2023). To maintain sufficient coverage of the study period and sample in our data, we excluded days in which participants completed fewer than nine out of 13 momentary assessments (i.e., a daily response rate of less than 69.2%). This left a total of 1749 days across 292 participants. For the between-day analyses, we excluded six participants with only one day with a response rate higher than our inclusion threshold.

As the responses to the hunger and craving questions were strongly correlated with one another (Pearson's r = 0.80; $t(22,177) = 197.54$, $p < .001$, 95% CI[0.79, 0.80]), we calculated the average of the two scores as our hunger measure at each assessment. For our pre-registered analyses, we reduced each participant's momentary assessments to four summary variables. First, we calculated the mean and SD of the hunger measure for each participant on each day (we corrected the SD for the mean by using residuals from the regression of the SD on the mean). We then took each participant's average of these two values across the week. We refer to these variables as **within-day mean** (**wM**) and **within-day (corrected) SD** (**wCSD**). We calculated two further variables: **between-day variation in the mean** (**bvM**) and **between-day variation in the (corrected) standard deviation** (**bvCSD**). Each participant's **bvM** is the standard deviation of their **wM** from each day across the week, and their **bvCSD** is the standard deviation of their **wCSD** across the week.

Pilot data (described in the preregistered protocol at https://osf.io/xntdf) suggested that other possible descriptive measures (daily median, range, minimum, maximum and area under the curve (AUC)) were very highly correlated with the mean and SD and therefore did not add any additional information. This was true for the present dataset (see Appendix A.3). Thus, we do not use these alternate descriptive measures in the present paper.

For our pre-registered analyses, we conducted a MANOVA to assess differences between FI and FS participants in the four measures of hunger defined above. These were followed with univariate models to determine which variables differed between the FI and FS groups.

In addition to pre-registered analyses, we ran exploratory analyses to investigate whether our findings differed depending on the severity of FI by rerunning the pre-registered analyses with three levels of FI rather than two (using scorings defined by (USDA Economic Research Services, 2023b): FS, moderate FI (mFI) and high FI (hFI). Finally, we conducted exploratory analyses with the momentary assessment as the unit of analysis to investigate whether any differences in hunger between FI and FS were dependent on the time an assessment took place (i.e., time of day) or on the day type of the day an assessment occurred (i.e., weekend vs weekday). Here, our outcome variable was the momentary hunger measure, with FI group, time of day and day type as predictors. We included interactions between the predictors and added a random effect of participant to allow for repeated measures.

2.4 Results

2.4.1 Patterns of Hunger

To illustrate the structure of our data, we plotted the average hunger rating at each time point for all participants (Figure 2.2). On average, hunger peaked at 1200 and 1700 in our sample. The lowest hunger ratings were reported at the last assessment at 2100.

2.4.2 Pre-registered Analyses

A one-way MANOVA revealed a significant effect of FI on our within- and between-day hunger variables ($F(4, 281) = 2.84$, $p = 0.025$; Table 2.1). Follow-up univariate analyses showed that, between the FI and FS groups, there was no difference in within-day mean hunger (**wM**; $t(290) = 0.17$, $p = .866$, $d = 0.02$; Figure 2.3A) or within-in-day standard deviation of hunger (**wCSD**; $t(290) = 1.31$, $p = .193$, $d = 0.15$; Figure 2.3B). However, the FI group had higher variation in between-day mean hunger (**bvM**; $t(284) = 2.43$, $p = .016$, $d = 0.29$; Figure 2.3C) and between-day SD of hunger (**bvCSD**; $t(284) = 2.90$, $p = .004$, $d = 0.34$; Figure 2.3D) than the FS group.

To visualise the finding that the FI group had greater variation in the experience of hunger between days than the FS group, we plotted the SD of hunger at each momentary assessment. Figure 2.4 shows that the FI group had higher variation than the FS group in their hunger ratings at most of the assessment times.

Independent samples *t*-tests showed that, on average, the FI group were younger than the FS group (Table 2.1; $t(290) = -3.86$, $p < .001$, $d = -0.45$), but the results of the MANOVA did not change when we controlled for age (no effect of age on dependent variables, $F(4, 280) = 1.30$, $p = .27$; significant effect of FI on dependent variables, $F(4, 280) = 2.83$, $p = .025$). The FI group also had a higher BMI (Table 2.1; $t(290) = 4.07$, $p < .001$, $d = 0.48$) than the FS group. However, there was no significant difference between the groups in the percentage of momentary assessments participants responded to across the week (Table 2.1; $t(290) = -0.72$, $p = .475$, $d = -0.08$). There was also no significant difference between groups in how many days were included for analysis in the MANOVA ($t(290) = -1.02$, $p = .310$, $d = -0.12$).

2.4.3 Exploratory Analyses

Of the participants included in the analyses, 69 participants could be classified as having "low food security" (scores of 2-4) according to the USDA Household Food Security Survey Six-Item Short Form module (USDA Economic Research Services, 2023b) and 74 participants could be classified as having "very low food security" (scores of 5-6). Here, we will refer to these groups of participants as moderate FI (mFI) and high FI (hFI), respectively.

Compared to the FS group, there was no difference in mean hunger (**wM**) in the mFI group ($t(289) = -0.08$, $p = .934$) or the hFI group ($t(289) = 0.35$, $p = .728$; Appendix A.4), nor was there a difference in the within-day standard deviation of hunger (**wCSD**) in the mFI group ($t(289) = 1.59$, $p = .114$) or the hFI group ($t(289) = 0.56$, $p = .574$; Appendix A.4). The hFI group had greater variation in between-day mean hunger (**bvM**; $t(283) = 2.81$, $p = .005$) than the FS group, but the mFI group did not ($t(283) = 1.12$, $p = .264$; Appendix A.4). Similarly, the hFI group had greater variation in between-day SD of hunger (**bvCSD**; $t(283) = 3.21$, $p = .002$) than the FS group, but the mFI group did not ($t(283) = 1.49$, $p = .139$; Appendix A.4).

In our exploratory model using the momentary assessment as the unit of analysis, there were significant differences in hunger by time of day ($F(12, 21840) = 56.91$, $p < .001$) and by day type (i.e., weekend vs. weekday; $F(1, 21854) = 4.80$, $p = .029$). There was also a significant interaction between time of day and day type ($F(12, 21840) = 4.41$, $p < .001$), suggesting that hunger varies over the course of the day differently at the weekend. The main effect of FI was not significant ($F(1, 300) = 0.05$, $p = .830$). However, there was a significant interaction between FI and time of day ($F(12, 21840) = 1.89$, $p = .031$). The interaction between FI and day type was not significant ($F(1, 21854) = 0.21$, $p = .646$), and neither was the three-way interaction ($F(12, 21840) = 1.35$, $p = .181$). Thus, the exploratory analysis suggests that although members of the FI group were not more hungry overall, their hunger patterned differently over the day than that of the members of the FS group (Appendix A.5). FI participants started the study period each day with a higher hunger level than FS participants. In both groups, hunger peaked at 1200 and 1700, with the FS group reporting higher hunger levels at these times. At 1400, there was a trough in hunger rating that was lower in FS participants; FI participants reported being hungrier than FS participants at 1400. The lowest hunger rating was at 2100, with little apparent difference between the FS and FI groups.

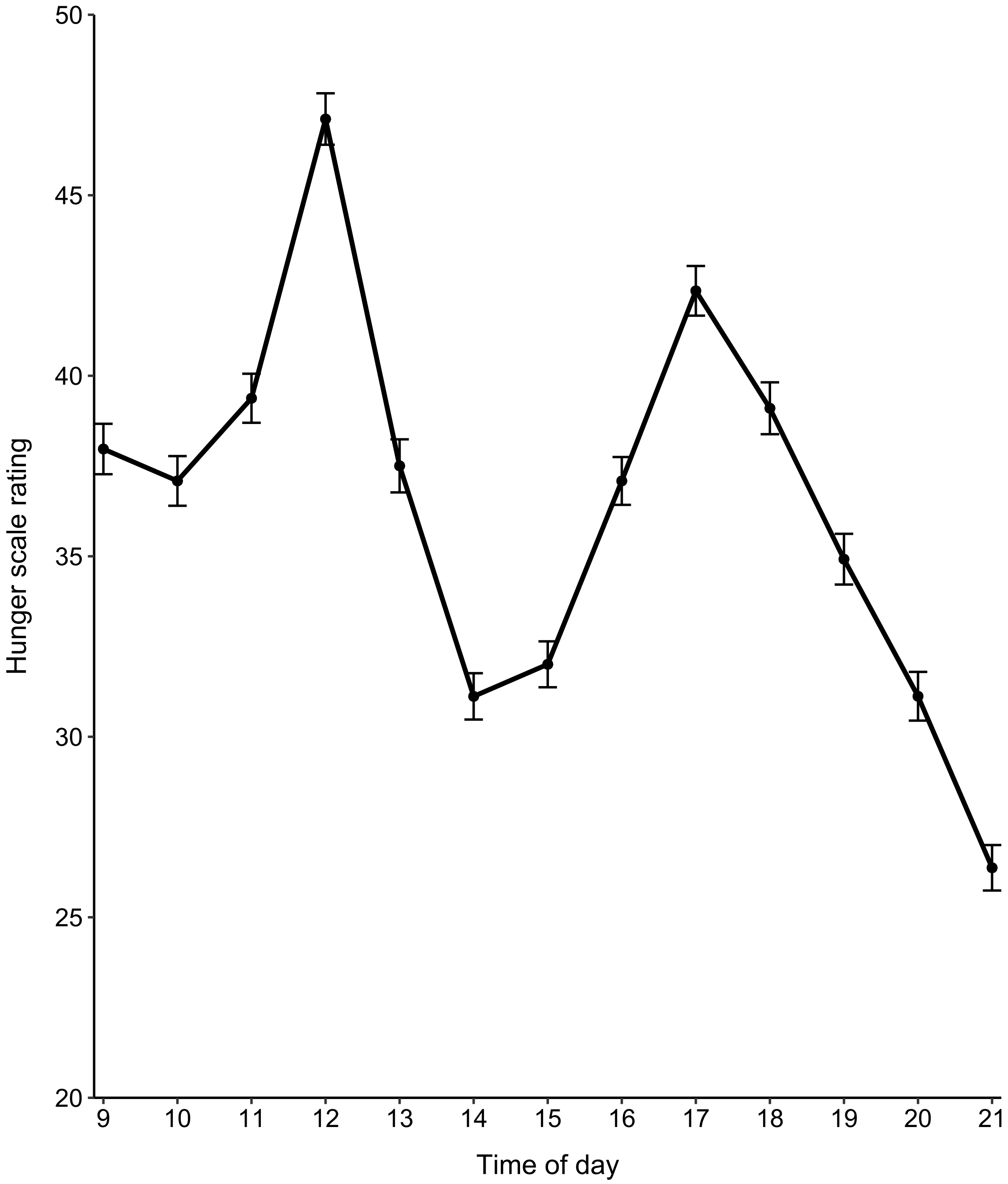


Figure 2.2. Mean hunger scale rating at each time point across all days for each participant.

Note. Error bars indicate one standard error of the mean for momentary assessments. N = 22,179.

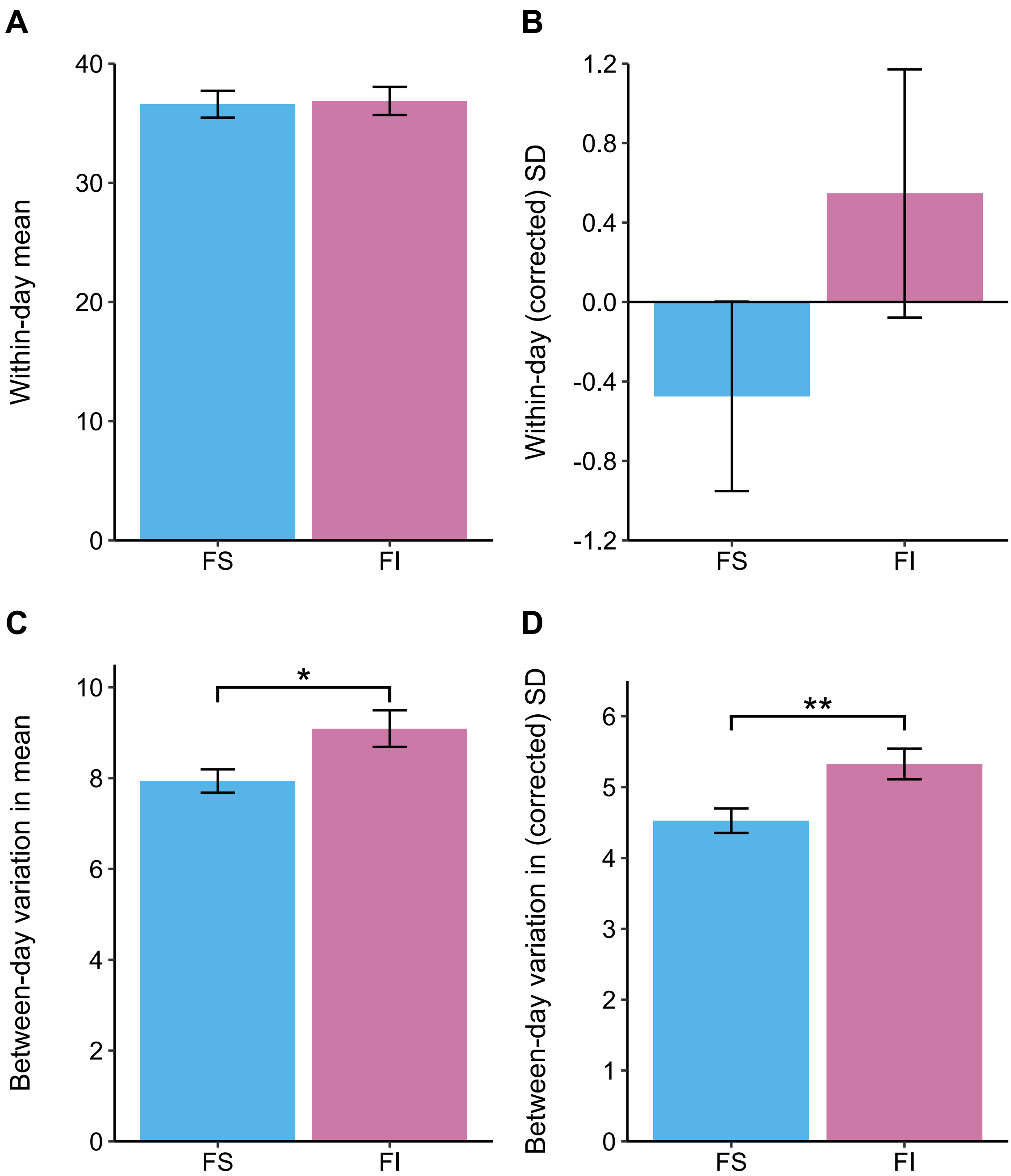


Figure 2.3. Within-day (A, B) and between-day (C, D) hunger measures in FI and FS groups.

Note. Error bars indicate one standard error of the mean. *significance of *p* <.05. **significance of *p* <.01.

Table 2.1. Descriptive statistics of hunger measures, participants' age, BMI, total compliance (across the study period), and number of days were included for analysis in the MANOVA in FI and FS groups.

		FS			FI		
		Mean	*SD*	Range	*Mean*	*SD*	Range
Hunger measures	Within-day mean (**wM**)	36.6	13.7	5.9-82.1	36.9	14.12	2.62-74.9
	Within-day (corrected) SD (**wCSD**)	-0.48	5.82	-16.40-17.12	0.55	7.46	-16.81-19.56
	Between-day variation in the mean (**bvM**)	7.95	3.10	1.84-17.12	9.10	4.78	0.73-33.17
	Between-day variation in the (corrected) standard deviation (**bvCSD**)	4.52	2.07	0.34-15.25	5.32	2.57	0.53-13.52
	Age	39.7	1.2	20-64	36.5	6.7	20-57
	BMI	27.4	6.0	18.3-47.6	31.0	8.9	16.4-71.0
	Total compliance (%)	83.8	14.2	14.3-100.0	82.6	14.7	28.6-100.0
	Number of days included in MANOVA	6.1	1.5	1-7	5.9	1.6	1-7

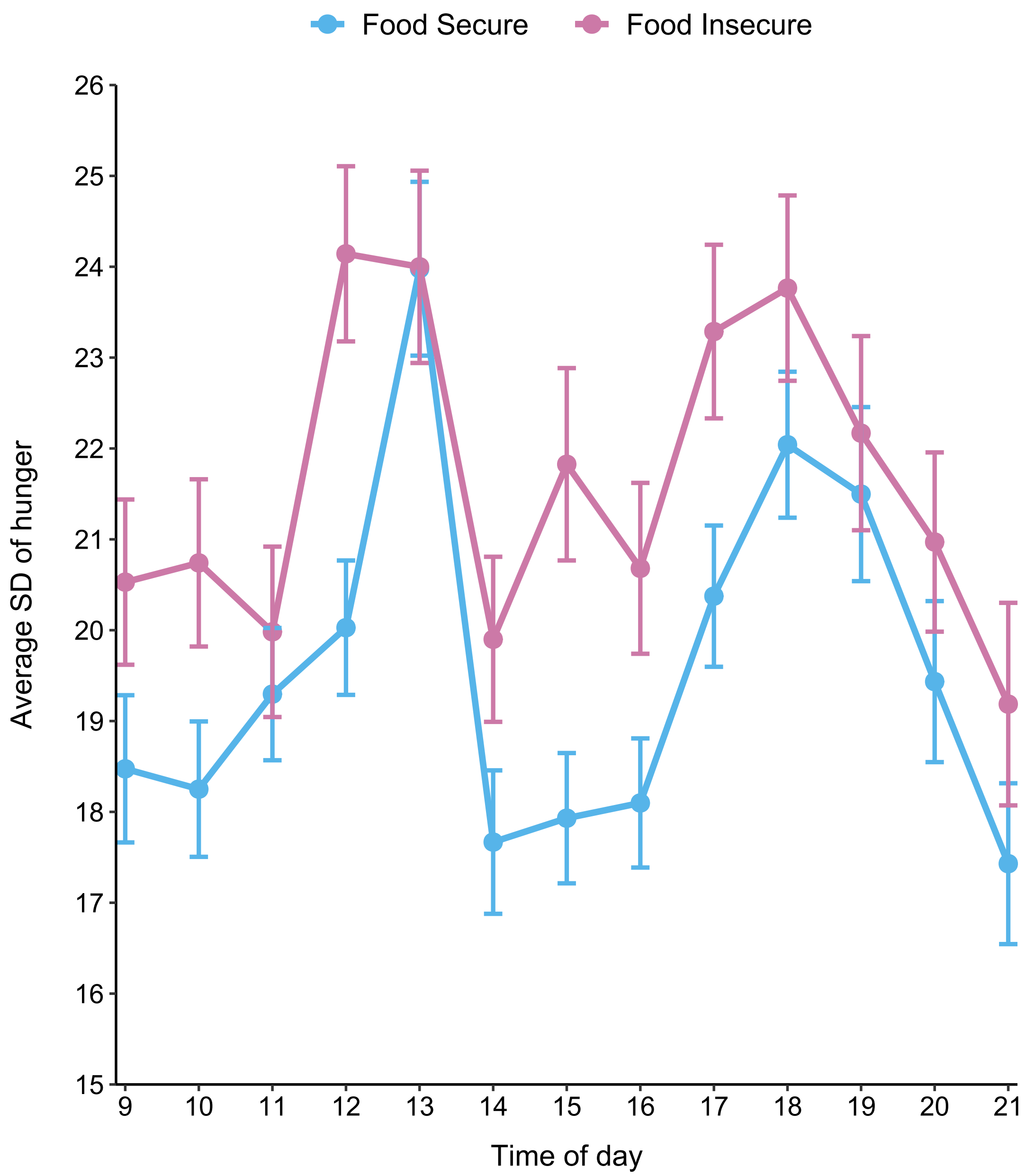


Figure 2.4. Average SD at each assessment across the week for FI and FS groups.

Note. For each participant, we calculated the SD of their hunger ratings for each assessment in the study period. We then averaged the SDs for each assessment for each group. Error bars indicate one standard error of the mean.

2.5 Discussion

In the present study, we measured hunger throughout the day for a week in two groups of UK females, one with and one without FI. We found no differences between groups in mean hunger or within-day hunger variation. However, we found that the FI group had greater variation in mean levels of daily hunger and variation of hunger across days than the FS group. Our results suggest that the experience of hunger under conditions of FI differs more from day to day than is true under FS. To our knowledge, this is the first dataset that quantitatively describes the daily experience of hunger in conditions of FI.

2.5.1 Interpretation of Results

Our within-day results contradicted our predictions that the FI group would have higher average hunger (P1) and within-day variation (P2), suggesting that FI is not associated with globally higher levels of hunger and hunger variation. However, we observed higher between-day variations in daily mean hunger and hunger variation in the FI group, which supported predictions P3 and P4, respectively. We also found that the severity of FI is relevant to our findings; only the hFI group had greater variation in mean levels of daily hunger and variation of hunger across days than the FS group. There was no significant difference between mFI and FS (Appendix A.4), which suggests that the severity of FI impacts how the experience of hunger differs between days from that experienced under conditions of FS.

Greater between-day variation in daily mean hunger and hunger variation may reflect particular experiences which are known to be associated with FI. For example, these results may reflect greater uncertainty and instability in access to food, higher variation in meal timings, or more demanding or chaotic lives at home and work (Eicher-Miller et al., 2023; Nettle & Bateson, 2019; Shinwell et al., 2022). These characteristics could also help to explain our exploratory results which suggested that, on average, hunger was distributed differently across the day in the FI group. It is worth noting that these features of FI could also lead to more hunger variation within an individual within a day; however, our results suggest that they do not do so overall.

2.5.2 Use of EMA Methods

Using EMA methods in this study allowed us to sample participants repeatedly during their daily lives. EMA study design minimises recall bias, improves ecological validity, and allows for detailed data collection over longer timeframes than would be reasonable in a laboratory setting (Shiffman et al., 2008). Of course, collecting data outside of a controlled laboratory setting means that participant responses may not be as consistent, reliable, or considered due to distractions from their surroundings. However, in our study, potential noise in the data resulting from *in situ* data collection was largely overcome by more frequent repeated sampling, longer study periods, and increased sample sizes than what would be achievable in the laboratory. Additionally, reporting how hungry one feels is a relatively simple task, and responses are unlikely to be significantly affected by environmental distractions. Existing evidence also highlights the improvements in validity and adherence that app-based EMA may offer over traditional pen-and-paper VAS methods for free-living assessments of hunger (Holliday et al., 2021).

Furthermore, because participants completed the study as they went about their day-to-day lives, we had to consider how to maximise study adherence and minimise how many momentary assessments participants missed. To encourage participants to respond as often as possible, we offered an additional voucher to participants who responded to more than 70% of the momentary assessments across the week and prompted participants during the study period if their response rate fell below this. In addition to incentives (Wrzus & Neubauer, 2022), study design can also minimise how many momentary assessments are missed in EMA studies; minimising momentary assessment length improves study compliance, but sampling frequency has no impact (Eisele et al., 2022). Therefore, we used a high-frequency sampling schedule to ensure we obtained high-fidelity data that accurately captured fluctuating hunger levels throughout the day and minimised the time it took a participant to respond to each momentary assessment. However, by only including five questions with VAS responses at each assessment, we were limited in the amount of data we could collect and consequently could not gather additional information about participants' eating episodes and dietary habits. While we recognise dietary data could have been a valuable addition to our dataset, we believe the quality of our existing dataset benefitted from limiting the scope of our study. Compliance would likely have been lower had we increased the participant burden through additional questioning. We achieved 83% compliance across the study, above the 79% average found in a recent meta-analysis of EMA research (Wrzus & Neubauer, 2022), and there was no difference in compliance between groups. Therefore, we conclude that we were justified in limiting the scope of our study, and we do not believe study compliance impacted the conclusions drawn from the between-group comparisons of the present study.

2.5.3 Assessing Hunger

At each momentary assessment, we used two questions about different subcomponents of hunger (subjective hunger (Blundell et al., 2010) and craving (Reichenberger et al., 2020)) to investigate whether the experience of different aspects of hunger differed between FI and FS. However, we did not use these two questions as independent measures of hunger as their responses were strongly correlated. Instead, we created a hunger measure that was the average of the two responses. To our knowledge, this exact measure has not been used in prior research. Nonetheless, we find similar overall patterns of daily hunger using our hunger measure as in other research (e.g., McKiernan et al., 2008), indicating our approach produces outcomes similar to previous work despite its differences.

Another consideration for EMA research is whether repeated sampling of the same questions influences participants' responses over time (referred to as reactivity; Shiffman et al., 2008). However, there is little evidence for reactivity in EMA designs which do not aim to change a target behaviour as in the present study (Shiffman et al., 2008). Furthermore, the hunger questions that we used at each momentary assessment have been successfully used in other studies which do not report issues with reactivity (Dzubur et al., 2022; Reichenberger et al., 2020). We also ensured that participants in both groups were blinded to the FI component of the study (FI questions were masked in a “Cost of Living” section of the initial survey) so that knowledge of the study topic did not impact responses. In the study debrief, we informed participants about the FI aspect of the study, but they did not know which group they were in.

2.5.4 Study Sample

Our study focused on recruiting female parents of school-aged children living in the North East of England. Therefore, we cannot claim that our sample broadly represented UK females. However, there was a strong rationale for our targeted sample. For example, we initially planned to recruit exclusively from parent populations of schools with above-average proportions of pupils eligible for free school meals and match FI and FS participants based on BMI, age, and school (to minimise the impact of unknown confounds of the local area). We focussed on schools with high free school meal eligibility to maximise our odds of recruiting females experiencing FI. Furthermore, we concentrated on recruiting parents of school-age children as rates of FI are higher in households with children than without (The Food Foundation, 2024), which further improved our chances of recruiting our target sample.

However, we could only recruit half of our sample from school parent populations due to recruitment challenges. Consequently, we had to remove our matching criteria and use targeted Facebook advertising to recruit the remainder of our sample. Before we shifted our approach, all participants joined the study via schools in the NE of England, which had some of the highest regional rates of FI when the study began (Department for Work & Pensions, 2023). Subsequently, we retained our recruitment focus on the NE, as sampling here improved our chances of recruiting enough participants to our FI group to meet our planned sample size.

2.5.5 Barriers to Participation

The intensity of the sampling schedule and week-long time commitment likely created a barrier to study participation; during recruitment, many individuals indicated they would like to participate but could not because working arrangements meant they would be unable to respond to momentary assessments during working hours. Such limitations may have affected certain groups more than others (e.g., females working in education or healthcare) and led to a less representative sample, particularly as many females in the UK work in these settings (e.g., almost 1 million women in the UK work in the NHS (NHS, 2019)). Future studies may consider how EMA methods, such as those used in the present study, may be adapted to allow for the inclusion of these participants, for instance, using wearable devices to respond to momentary assessments. Furthermore, our study relied on participants owning a smartphone and having a level of digital literacy that allowed them to download and use an unfamiliar app. While these factors may have further limited who was able to participate in our study, they are unlikely to have had a significant impact on the representativeness of our sample, as 92% of adults in the UK have a smartphone (Ofcom, 2024). Digital poverty was another potential barrier to study participation, especially given that half of our sample was likely experiencing financial instability and 28% of UK households struggle to afford communications services (Ofcom, 2024). However, continuous access to mobile data or a Wi-Fi connection throughout the study period was not required as the app downloads the notification schedule to a participant's phone when they sign up and uploads recorded responses to the server when the participant's device next has an internet connection. Therefore, only those experiencing the most severe levels of digital data poverty would not have been able to participate in our study.

2.5.6 Future Research and Implications

We recognise that the sample in the present study was limited and did not include males or older females, and, as such, our findings cannot be assumed to reflect their experiences. Future research should explore whether observed between-day hunger differences are also present in males with FI, which may provide additional insights into sex differences in the association between weight status and FI exposure (Nettle et al., 2017). Furthermore, future research should consider older females. Although rates of FI in UK adults over 65 are lower than in younger adults (Department for Work & Pensions, 2024), older females may be especially vulnerable to the health-harming effects of FI because of the compounding effect of their age (Gundersen & Ziliak, 2015). Moreover, we did not collect data regarding ethnicity in the present study. Given that the North East of England has particularly low ethnic diversity relative to the rest of the UK, our sample is likely not representative of the ethnic diversity of the UK population (UK Government, 2022). Minoritised ethnic groups often have different dietary habits and higher rates of FI than the White British population in the UK (Department for Work & Pensions, 2023; UK Government, 2024). Thus, future research should ensure these people are meaningfully represented in the sample to ensure that the results reflect their experiences.

The causal mechanisms linking FI to poorer health outcomes in females, particularly increased weight status (Bateson & Pepper, 2023), remain poorly understood. Stress is frequently discussed as a possible mediator (Bateson & Pepper, 2023; Franklin et al., 2012; Kowaleski-Jones et al., 2019; Laraia, 2013). Our next steps with this dataset are to use longitudinal analysis methods to investigate the temporal relationship between momentary hunger and perceived stress and explore differences in this relationship between the FI and FS samples. An EMA study that investigated hunger and stress in a population of vulnerable young adults found that when participants experienced above-average hunger, they reported greater stress variation at the next assessment (Dzubur et al. 2022). Furthermore, when individuals who reported higher average stress levels became hungry, they became significantly more stressed than individuals with lower average stress. Given that FI is associated with higher levels of stress (Martin et al., 2016), greater variation in hunger across days may exacerbate already high stress levels in the FI sample compared to FS. Research into the links between these psychological components of FI, as well as behavioural (e.g., diet and exercise measures) and physiological (e.g., blood glucose, energy expenditure) components, is critical to establish the causal pathways that lead to the adverse health outcomes associated with FI and evidence the need for policy change to overcome the growing public health challenge that FI has become in the UK.

2.6 Conclusion

We have shown that females experiencing FI have greater variation in daily mean hunger and hunger variation across days than those experiencing FS. However, we did not find evidence that these groups differed in average hunger or hunger variation within a day. To our knowledge, our study is the first to gather hourly, quantitative measurements of hunger in an FI and FS group. The hourly hunger patterns in this study largely follow the trends expected from research monitoring hunger in the general population (McKiernan et al., 2008), but the FI group appear to be more variable in how hungry they feel at each assessment than the FS group. We suggest future research should measure hunger longitudinally with other psychological, behavioural and physiological factors in participants experiencing FI (with an FS comparison group) to investigate potential causal pathways of how FI negatively impacts health. Future research should also explore whether the greater variation in hunger under FI conditions that we have described can be explained by uncertainty in food access, greater variability in meal timing, and more demanding and chaotic lives in those experiencing FI.

Chapter 3 No Effect of Hunger on Attentional Capture by Food Cues: Two Replication Studies

Preface

Chapter 3 addresses the second aim of the present thesis, "To investigate the effect of hunger on food-related cognition".

In the introduction, I suggested that we may expect to see strategic shifts in sensitivity to food cues as an adaptive response to cues of food scarcity, such as those experienced by individuals with FI. In Chapter 2, I have shown that hunger is experienced differently in females with and without FI. In light of Chapter 2's findings and the understanding that hunger has a strong influence on dietary decision-making, I hypothesise that hunger may contribute to changes in dietary behaviour under conditions of FI, via increased sensitivity to food cues. If the hypothesis is supported, one might also expect to observe increased sensitivity to food cues in a state of acute hunger, which has been found in a small number of studies (discussed earlier and in the following two chapters).

In Chapter 3, I replicate a previously published study which found that visual food cues captured attention to a greater extent when individuals were hungry compared to sated. Here, I choose to focus on visual cues of food as food imagery is a ubiquitous feature of contemporary Western living, meaning results from such research would have a high degree of ecological relevance to everyday life. Furthermore, I opted to replicate Piech et al. (2010), which used an Emotional Blink of Attention paradigm. Compared to other cue reactivity paradigms, Attentional Blink paradigms are better at allowing us to examine how stimuli are prioritised during "cognitive bottlenecks" – i.e., situations in which multiple cues are simultaneously competing for limited cognitive resources. If a particular type of cue consistently captures attention, then one may hypothesise that the prioritisation of the cue is strategic. The Emotional Blink of Attention paradigm allows us to investigate how emotionally salient, biologically relevant stimuli are prioritised during cognitive bottlenecks relative to stimuli which are considered neutral (or irrelevant to survival). Such cognitive bottlenecks can be representative of real world experiences of managing the cognitive requirements of competing interests and goals.

Using the Emotional Blink of Attention paradigm with food cues can tell us about how food information is prioritised relative to other types of cues (biologically relevant or otherwise). It also allows us to study how food may be processed differently in response to changes in emotional or motivational state, particularly in situations where the change in state is related to the type of cue presented. In the present context, the state of interest is hunger. Therefore, the following chapter aims to investigate the effect of hunger on food-related cognition – the second aim of this thesis – by using an Emotional Blink of Attention paradigm to examine the relationship between hunger and the prioritisation of food cues when cognitive resources are constrained.

Publication

This chapter is a peer-reviewed publication in *Appetite.* I have not changed it except to refer to other material in this thesis. Thus, I retain the use of “we” in place of “I” as I use elsewhere in this thesis. The citation information for the pre-print is as follows:

Neal, C., Pepper, G. V., Allen, C., & Nettle, D. (2023). No effect of hunger on attentional capture by food cues: Two replication studies, *Appetite*, *191*, 107065. https://doi.org/10.1016/j.appet.2023.107065

Author contributions

I prepared each section of the text, including figures, with guidance and critical feedback on writing style, narrative and structure from DN, GVP and CA. I led the development of the rationale and DN suggested the previously published paper for replication. I devised the experimental protocol and coded the experimental task, with input on key decisions from DN, GVP and CA. I was solely responsible for participant recruitment, data collection and data capture. I designed and conducted the data analysis with analysis design support from DN, predominantly regarding the use of Bayesian statistical methods.

3.1 Abstract

Food cues potently capture human attention, and it has been suggested that hunger increases their propensity to do so. However, the evidence for such hunger-related attentional biases is weak. We focus on one recent study that did show significantly greater attentional capture by food cues when participants were hungry, using an Emotional Blink of Attention (EBA) task (Piech et al., 2010). We conducted online (N=29) and in-person (N=28) replications of this study with British participants and a Bayesian analytical approach. For the EBA task, participants tried to identify a rotated target image in a Rapid Serial Visual Presentation (RSVP). Targets were preceded by “neutral”, “romantic”, or “food” distractor images. Participants completed the task twice, 6-11 days apart, once hungry (overnight plus 6h fast) and once sated (after a self-selected lunch in the preceding 1h). We predicted that food images would create a greater attentional blink when participants were hungry than when they were sated, but romantic and neutral images would not. We found no evidence that hunger increased attentional capture by food cues, despite our experiments passing manipulation and quality assurance checks. Our sample and stimuli differed from the study we were replicating in several ways, but we were unable to identify any specific factor responsible for the difference in results. The original finding may not be generalisable. The EBA is more sensitive to the physical distinctiveness of distractors from filler and target images than their emotional valence, undermining the sensitivity of the EBA task for picking up subtle changes in motivational state. Moreover, hunger-related attentional bias shifts may not be substantial over the intensities and durations of hunger typically induced in laboratory experiments.

3.2 Introduction

Hunger is a coordinating mechanism of psychological and physiological processes to solve the adaptive problem of acquiring food (Cosmides & Tooby, 2000). When acquiring food is an organism's most dominant adaptive concern, attentional resources should be taken away from other adaptive problems and reallocated to stimuli likely to increase the odds of successfully sourcing food (Al-Shawaf, 2016). Therefore, when a person becomes hungry, we should expect food cues to capture their attention more readily. In modern food environments where energy-dense foods are ubiquitous, such attentional shifts may be maladaptive. If increasing attentional bias (AB) for food contributes to increased food intake (Werthmann et al., 2015), individuals may be at a heightened risk of developing obesity.

Many researchers have investigated food-related ABs and their associations with obesity, disordered eating and dietary restraint (see Field et al. (2016); Hardman et al. (2021); Werthmann et al. (2015) for reviews). There is consistent evidence that food readily captures human attention. However, the evidence that hunger increases AB for food is less consistent. A recent meta-analysis concluded that existing evidence does not support this hypothesis (Hardman et al., 2021).

A few studies have used the emotional blink of attention (EBA) task to investigate attentional capture by food (Arumäe et al., 2019; Davidson et al., 2018; Piech et al., 2010) rather than more commonly used attentional paradigms, such as modified Stroop, visual/dot probe, or eye-tracking. In the EBA task, the participant tries to detect a target image in a Rapid Serial Visual Presentation (RSVP). An attentional blink effect occurs if performance is poorer when a distractor image is placed two images before the target (lag2) than when it is placed eight images before the target (lag8). Such a difference is thought to occur because of the salience of the distractor image: the more salient the image is to the participant, the more likely it is to capture their attention and prevent them from attending to the target they are looking for in the immediate aftermath of the distractor's presentation. In longer-lag trials (e.g., lag8), this 'blink' is assumed to have resolved. It is often assumed that feelings of hunger increase the value of food and food cues (see Redlich et al. (2021) for a discussion). Thus, if a participant completes an EBA task with food image distractors when they are hungry, they should show a larger EBA (i.e., a larger difference between lag2 and lag8 performance) than when they are sated.

Piech et al. (2010) (henceforth PPZ) used an EBA task to investigate the effects of hunger on attention in a sample of US undergraduate students. In a within-subjects design, they found that attentional capture by food cues was greater when participants were hungry (following a six-hour fast) than sated (after eating as usual). PPZ used three different types of images as distractors, categorised as neutral, romantic, and food. Their key result was that participants had worse performance on lag2 trials with food distractors when hungry, which was not true for lag2 trials with neutral or romantic distractors. This was despite their participants receiving financial incentives for performing well on the task. Therefore, PPZ's results suggest that participants could not ignore task-irrelevant visual food cues when hungry, even when there was a financial incentive to do so.

Davidson et al. (2018) used an adapted version of PPZ's EBA paradigm to assess the relationship between the ability of food stimuli to create an emotional blink of attention and the motivation to eat. They found that task performance in trials with food distractors was worse than in trials with neutral distractors, consistent with the general attentional potency of food-related cues. Additionally, performance after food distractors became worse as appetite increased. These findings are consistent with those reported by PPZ. However, they do not represent a close replication because Davidson et al. (2018) were interested in a different research question concerning sensory-specific satiety, and their experimental design differed from PPZ. Consequently, they tested non-fasted participants at regular intervals before and after consuming a midday meal, they used only food and neutral distractors, and their food distractors were specifically chosen to be similar or dissimilar to the midday meal participants consumed.

More recent work on the EBA paradigm suggests that an image's physical distinctiveness, rather than its content, dictates its ability to capture attention and create an attentional blink (Santacroce et al., 2023). If this finding is correct, it makes it less plausible that PPZ's finding that hunger increases the EBA created by food images is a robust one. Even if hunger could increase the salience of food distractors in the EBA task, this would unlikely be enough to overcome the determining influence of the images' physical distinctiveness.

Accordingly, some studies fail to find an effect of hunger on EBA for food images. Arumäe et al. (2019) used an EBA task and hunger manipulation more in line with PPZ than Davidson et al. (2018). Hunger had no impact on performance in trials with food distractors. While their EBA task followed a procedure adapted from PPZ, there were key deviations in their methods. They used different filler, target, and distractor image sets to PPZ, used only neutral and food distractors, and presented distractors 2 or 4 images before the target, not 2 or 8 as in PPZ. Their fasting and sated conditions were similar to PPZ's. However, their participant sample and experimental procedure were not: only women were recruited to their study, and participants completed two additional tasks in each session in a counterbalanced order. Thus, although the findings of Arumäe et al. (2019) suggest that PPZ's claim that hunger increases EBA for food images may not be very robust or generalisable, a closer replication (in the sense of Brandt et al. (2014)) would be useful.

Considering the theoretical and clinical implications of PPZ's central finding and the low replication rates in psychological research (Open Science Collaboration, 2015), further replication attempts are necessary. In this paper, we report the results of two pre-registered experiments designed to replicate PPZ using British samples. Although we aimed for our experiments to be as close to PPZ's design and procedure as practicable, several differences could not be avoided (see ***3.3.1 Overview of Experiments***). Our central aim was to replicate the increased attentional capture by food (but not other) cues in the hungry (but not sated) state: that is a state by image category interaction effect in lag2, but not lag8, trials. In line with the findings of PPZ, we hypothesised that hunger would increase the attentional capture of food cues - but not other types of cues - and that this effect would be lost when participants were sated.

3.3 Methods

3.3.1 Overview of Experiments

We carried out two pre-registered experiments (henceforth E1 and E2). The pre-registered protocols and predictions are available online at https://osf.io/w2a8f and https://osf.io/v4wpt. The Newcastle University Faculty of Medical Science Research Ethics Committee (reference 8999/2020) granted ethical approval for both studies. All aspects of the experiments were presented on PsyToolKit (Stoet, 2010, 2017). Whilst the two experiments were designed to replicate the experiment reported in PPZ as closely as practicable, there were several mostly unavoidable differences, which we summarise here before giving fuller information in the sections that follow.

E1 was conducted online due to the COVID-19 pandemic, whereas E2, like PPZ, was an in-person experiment. We drew our stimulus images from the same image bank as PPZ but are unlikely to have used exactly the same subset of images. Our participant pool was different: as well as being from Britain rather than the USA, our two samples were not recruited by their student status, whereas PPZ used undergraduate students. Although PPZ did not provide descriptive statistics on the ages of their participants, we infer, given their recruitment strategy, that their participants would have had a lower mean age and a narrower range than ours. Our procedure for manipulating hunger was based on PPZ's. Further, in addition to the participant instructions that PPZ reported, we stipulated that our participants should abstain from satiating drinks in the hungry condition (not mentioned by PPZ); and that they should eat lunch within the hour prior to the session in the sated condition (PPZ assumed they would eat but did not instruct them to do so). Possibly for these reasons, our hunger manipulation was more effective than PPZ's, in both E1 and E2 (see ***3.4.1 Hunger Manipulation Check***).

We pre-registered and used a Bayesian approach to data analysis and hypothesis testing. This has several advantages (summarised in Wagenmakers et al. (2018)). Notably, it provides, through the Bayes factor (BF), a means of testing the support for the null hypothesis. That is, it allows researchers to distinguish the case of the null being likely to be true from the case of the results being inconclusive through insufficient statistical power. Relatedly, the Bayesian approach obviates the need to predetermine a target sample size through a priori power analysis. Instead, researchers can, without inflating the type-II error rate, continue sampling until the evidence either decisively supports the experimental hypothesis or decisively supports the null hypothesis. Since our Bayesian approach differs from PPZ's frequentist one, we also conducted frequentist analyses of our data using exactly the same strategy as PPZ. The conclusions were the same. Because it was what we pre-registered, and due to the advantages described above, we report the Bayesian analyses in the main paper and frequentist analyses in Appendix B.

3.3.2 Participants

For both experiments, we recruited 30 participants, the same number as PPZ (though PPZ analysed data from only 23 participants after exclusions). We pre-specified a flexible stopping rule for sample size, requiring a minimum sample size of 30 and a Bayes factor of < 1/10 or > 10 for the critical state by image category interaction in lag2 trials (see ***3.3.6 Data Analysis***) to stop participant recruitment. Both experiments met the Bayes factor criterion at the first point of inspection, after 30 participants.

For E1, we recruited 30 individuals using opportunity sampling, mainly from social media (ages 21–34 years, M = 28.4, SD = 3.7; women = 17, men = 13). For E2, we recruited 30 individuals from a research volunteer pool maintained by Newcastle University (ages 20–79 years, M = 42.9, SD = 20.2; women = 18, men = 11, non-binary = 1).

3.3.3 Experimental Design

Both experiments had a within-subjects design. Participants completed two sessions on different days, six to 11 days apart: one in the hungry condition and one in the sated condition. The order of hungry and sated sessions was counterbalanced.

3.3.4 Procedure

Recruitment and data collection for E1 took place from June 14th, 2021–July 23rd, 2021. We informed participants that they would need access to Google

Chrome on a PC or laptop with a physical keyboard and a quiet place where they would not be disturbed during the study. Recruitment and data collection for E2 took place from November 16th, 2021–February 16th, 2022. We informed participants that they would need to attend Newcastle University on two occasions, approximately one week apart. In both experiments, we informed participants that, to take part, they should have normal or corrected-to-normal vision and should not have a medical condition requiring them to eat regularly that would exclude them from safely completing a fast.

All sessions started six hours after participants had woken up. Waking time and session times were agreed upon with each participant during recruitment and were the same for both sessions. PPZ did not indicate what time of day their sessions took place. By personalising and standardising session timing for each participant, we minimised the potential impacts of circadian rhythm or fatigue on cognitive performance (Schmidt et al., 2007; Valdez et al., 2008) and other unidentified confounding factors related to timing.

In the hungry condition of both experiments, we instructed participants to refrain from eating from waking until after their session that day. This resulted in a minimum of six hours without eating before the experiment, the same as PPZ. We instructed participants to drink water and caffeinated drinks as usual in the hungry condition but to avoid satiating drinks (such as those with high milk, sugar, or calorie content). PPZ instructed participants to “continue drinking as usual” in the hungry condition. We excluded the consumption of potentially satiating drinks to create a robust hunger manipulation (an approach used in more recent research with hungry and sated conditions (Redlich et al., 2021)). We specified that participants could consume caffeinated but non-satiating drinks in the hungry condition to limit the potential impacts of caffeine withdrawal on the cognitive performance of habitual caffeine users (James & Rogers, 2005).

In the sated condition of both experiments, we instructed participants to eat and drink as usual from waking, as PPZ did. We asked participants to eat lunch in the hour before their session started, which PPZ did not specify. We implemented this requirement to minimise the level of hunger participants experienced in the sated condition and hence to maximise the difference between conditions.

In both E1 and E2, each session lasted approximately 35 minutes. Participants provided informed consent and then gave their age and gender. They then completed the EBA task and were asked if they had been interrupted during the task upon its completion. They provided a self-reported hunger rating, answered additional questions, and, only in their second session, completed a dietary restraint scale. At the end of their second session, we debriefed participants and reminded them how and when they would receive their rewards. Participants received a 10 GBP retail gift card as a show-up recompense for each completed session. In addition, like PPZ, we incentivised accuracy on the task. If their average accuracy across both sessions was over 80% or 90% (on trials with a target), participants received an additional 5 GBP or 10 GBP gift card, respectively. The participant with the highest average accuracy score in each experiment also received a prize of a 50 GBP gift card. We informed participants of these monetary incentives during recruitment.

3.3.5 Measures

Emotional Blink of Attention (EBA) Task

E1 participants were asked to complete the task in a quiet place where they would not be disturbed. E2 participants completed their sessions onsite in a controlled laboratory environment and were tested alone. The display in E2 was a 61.13 cm, 1920 x 1200 resolution monitor, which participants viewed from approximately 70 cm away.

The EBA task used in both experiments was as similar as practicable to the version used by PPZ (Figure 3.1). The task consisted of one block of 16 practice trials and six blocks of 32 real trials. There were one-minute breaks between blocks. The trial order was randomised within each block. Each trial was a rapid stream visual presentation (RSVP) of 17 images shown for 100ms each. Images were shown on a full-screen black background. Each trial started with a fixation cross, and participants pressed the spacebar to start. A target was present in 75% of the real trials. The target was an image that had been rotated by 90º, clockwise or anticlockwise. Both RSVP filler images and target images were photos of landscapes, some of which contained buildings.

A single distractor image was present in all trials with a target. Distractors were categorised as either food, romantic, or neutral images. They were only in position four, six, or eight in the RSVP sequence and either two positions (lag2) or eight positions (lag8) before a target.

Participants had to identify whether a target image was present in each trial by using key presses. They had five seconds to respond after every trial. If they correctly identified the presence of a target, they had to indicate the direction of its rotation by using the arrow keys within five seconds. PPZ did not report the response window that they used. We also instructed participants to respond as accurately and quickly as possible to each trial.

Accuracy was calculated by dividing the number of target trials in which the participant correctly identified the direction of the target rotation by the total number of trials with targets, then multiplying by 100. Participants were shown their cumulative accuracy for that session after each block and their total accuracy for that session at the end of the task. The displayed accuracies were based only on trials with targets. It is not clear whether or how PPZ provided accuracy feedback. In the paper that PPZ cite as the origin of their EBA procedure (Most et al., 2005), one of the two experiments did so on a per-trial basis.

We drew images from the same image sets as PPZ, as the original authors shared these with us. Most of these images had been acquired from the International Affective Picture System (IAPS) database (Lang et al., 1997), with additional supplementation from the internet for romantic and food distractors. As there were more images than required in each category, the subset we used may have differed slightly from the subset PPZ used. When selecting food images, we selected an even number of savoury and sweet food images (28 of each). In total, 168 different distractor images were used (56 from each category), alongside 84 landscape images as fillers. An additional 84 landscape images were used as target images; these were duplicated, with one copy rotated 90° clockwise and one copy rotated 90° anticlockwise.

Self-Reported Hunger Rating

After completing the EBA task in all sessions, in E1 and E2, participants answered the question ‘How hungry are you?’ using a scale anchored at 0 (not at all hungry) and 7 (extremely hungry). We used their responses as a manipulation check.

Additional Measures

For all sessions, in E1 and E2, we asked participants when they last had something to eat. This came after the self-reported hunger rating. It provided a condition compliance check and an alternative measure of hunger for exploratory analyses. We also asked if they regularly skipped breakfast for the purpose of exploratory analyses.

Dietary Restraint

Participants completed the dietary restraint scale (Herman et al., 1978; Herman & Polivy, 1975) after the EBA task in their second session. We scored participants using the methods of (Herman & Polivy, 1975). PPZ used this scale to explore the relationship between dietary restraint and attentional capture of food cues in lag2 trials in the hungry condition. While they did not find evidence of a significant relationship, we retain it here for comparability. We report the results relating to dietary restraint in Appendix B.1.

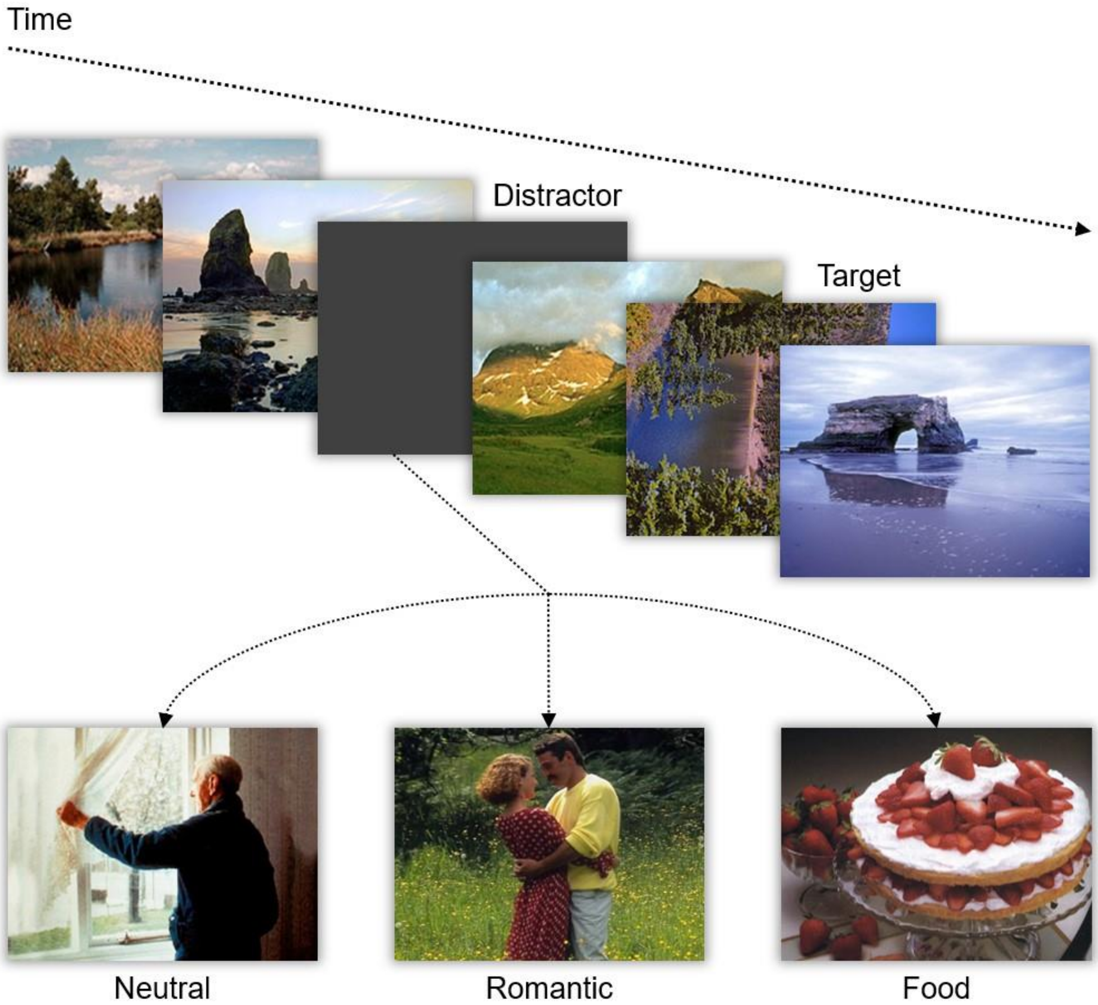


Figure 3.1. Representation of part of a single EBA task trial.

Note. In half of the trials with targets, the distractor image was shown eight places before the target (lag8) rather than two (lag2) as shown.

3.3.6 Data Analysis

Data were analysed and visualised in R (R Core Team, 2023). Our data and code are available at https://osf.io/w5en6/.

PPZ excluded participants who reported a lower hunger rating in the hungry condition than in the sated condition. We used the same criteria, leading to no exclusions in E1 and one in E2. PPZ also excluded participants with accuracy more than two standard deviations below the mean for the respective hunger condition. We excluded one participant from E1 and one from E2 based on these criteria.

We fitted Bayesian linear mixed models, which followed the structure of the repeated-measures ANOVAs used in PPZ (see Appendix B.2 for model specifications). Models included a random effect of participant to allow for the repeated measures. We used weakly informative priors of *N*(1, 10) for all parameters (McElreath, 2020). A variable called “sequence” was included in these models, as in PPZ, to account for a session order effect (that is, whether a participant completed their hungry or sated session first). This was included as PPZ found a practice effect across sessions and that this effect differed depending on which condition was completed first. We initially fitted models analysing all trials together before fitting separate models for lag2 and lag8 trials.

We used paired Bayesian *t*-tests to assess differences in hunger rating between states and to test accuracy differences between lags within each category, each category (in lag2 and lag8 trials, separately), and states for each category in lag2 trials.

Our additional frequentist analyses followed PPZ’s analysis strategy exactly. They involved repeated-measures ANOVAs followed up with paired *t*-tests. The results of the frequentist analyses are reported in Appendix B.3.

We pre-registered conditional requirements for successful replication of the main findings of PPZ. Our conditions were based on the strength of evidence for two key predictions:

P1. There will be a state by image category interaction effect on accuracy in lag2 trials – participants' accuracy will only be reduced after food distractors in their hungry session.

P2. There will not be a state by image category interaction effect on accuracy in lag8 trials.

Our statistical conditions required a Bayes factor of greater than 10 or less than 1/10 to support the prediction or the null, respectively. For successful replication, a Bayes factor greater than 10 was required for both P1 and P2. Alongside Bayes factors, we present posterior medians and their 89% credible intervals (CI). Although the 89% is arbitrary, it has become a convention in Bayesian data analysis (McElreath, 2020). We also report the probability of direction (pd). This indicates "the probability that a parameter is strictly positive or negative" (Makowski et al., 2019b).

3.4 Results

First, we report hunger manipulation, paradigm, and practice effect checks. We then present results related to the two predictions required for successful replication (P1 and P2) and additional exploratory analyses. We report E1 and E2 results together. Ancillary results, produced by conducting other analyses also reported by PPZ, can be found in Appendices B.4to B.9.

3.4.1 Hunger Manipulation Check

Participants had higher hunger ratings in their hungry session than in their sated session (BFs > 1000; Table 3.1; PPZ descriptives given for comparison), in E1 (median difference = 5.4, 89% CI [5.8, 5.1], pd = 100%) and E2 (median difference = 5.0, 89% CI [5.5, 4.5], pd = 100%).

Table 3.1 Descriptive statistics of self-reported hunger rating in each state and of accuracy on the EBA task in lag2, lag8, session 1 and session 2.

	Mean hunger rating (SD)		**Mean accuracy (SD)**			
	Sated	*Hungry*	*Lag2*	*Lag8*	*Session 1*	*Session 2*
PPZ	2.4 (1.2)	5.4 (1.4)	-	-	75.5 (7.3)	80.2 (6.7)
E1	0.5 (0.7)	6.0 (0.9)	84.0 (11.1)	91.0 (8.4)	86.2 (7.7)	88.8 (5.8)
E2	0.4 (0.9)	5.4 (1.4)	82.3 (12.3)	89.5 (9.8)	84.4 (8.7)	87.3 (8.0)

Note. PPZ values are missing for lag2 and lag8 columns as they were not reported.

3.4.2 Blink of Attention Check

Accuracy was higher in lag8 trials, compared to lag2 trials (Table 3.1; Figure 3.2), across all distractors in E1 (BF > 1000, median difference = 6.9, 89% CI [5.6, 8.1], pd = 100%) and E2 (BF > 1000, median difference = 7.2, 89% CI [5.8, 8.4], pd = 100%). This suggests that all distractors produced a blink of attention at lag2. Planned paired Bayesian *t*-tests (Appendix B.4) provided evidence for accuracy differences in lag2 and lag8 trials for each distractor category, in line with PPZ (Figure 3.2).

3.4.3 Practice Effect Check

There was evidence to support a practice effect in E1 and E2 (Table 3.1); participants had higher accuracy in their second session. While the Bayes factors did not reach our strict threshold (BF > 10) for supporting this prediction, the evidence for a practice effect was substantial as per the Bayes factor thresholds of Wetzels et al. (2011), in E1 (BF = 7.6, median difference = 2.6, 89% CI [1.2, 4.1], pd = 99.92%) and E2 (BF = 9.9, median difference = 2.9, 89% CI [1.3, 4.3], pd = 99.88%). The pd values also suggest a significant practice effect (Makowski et al., 2019b).

3.4.4 Replication of Main Findings (P1 and P2)

There was evidence to support the absence of a state by image category interaction effect in lag2 trials in E1 and E2 (Table 3.2; Figure 3.2). Thus, P1 was not supported. In E2, there was no evidence of a state by image category interaction effect in lag8 trials (Table 3.2). In E1, the Bayes factors did not reach our strict threshold (BF < 1/10) to support the absence of a state X image category interaction at lag8 (Table 3.2). However, the evidence was substantial as per the Bayes factor thresholds of Wetzels et al. (2011). Overall, given the lack of support for P1 in our data, we did not replicate the key finding of interest in PPZ.

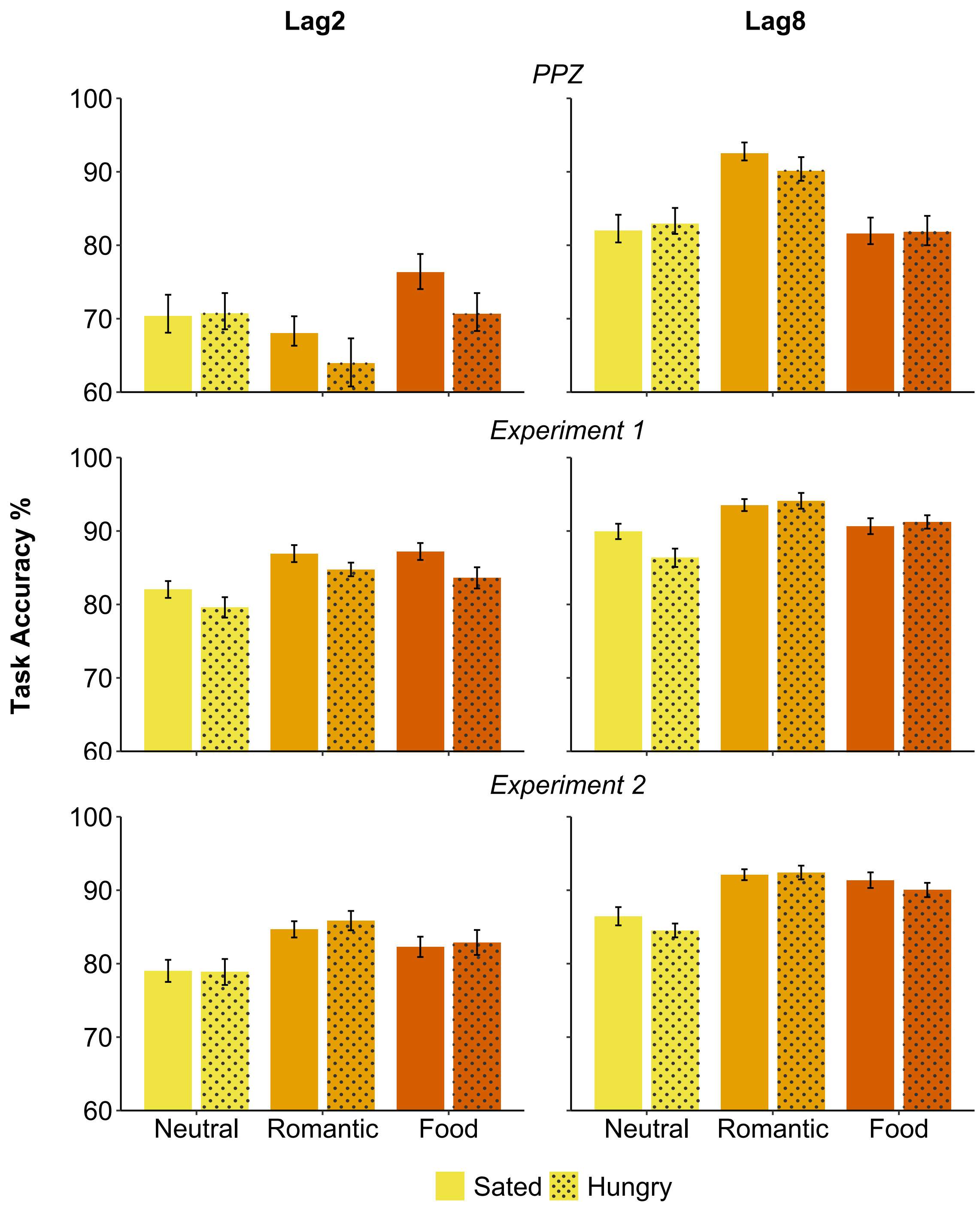


Figure 3.2. Accuracy in PPZ, E1 and E2 separated by lag, image category, and state.

Note. Trials are grouped by image distractor category, lag, and state. Lower task performance was hypothesised to indicate a greater attentional blink effect. The PPZ plots have been reproduced using *Graph Data Extractor* (2010) to extract data from the original published plots. Error bars indicate one standard error of the mean. For E1 and E2, error bars are within-subjects centred. This was not possible for PPZ because raw data were not available.

Table 3.2. Bayesian model output for P1 and P2 in E1 and E2, and corresponding findings of PPZ.

PPZ	E1	E2
P1: Interaction effect of image category and state in lag2 trials		
Significant *p* = .03 *F*(2, 21) = 3.80	**Evidence for null** BF = 0.06	**Evidence for null** BF = 0.09
	sated:romantic: median diff. = 0.02, 89% CI [-3.6, 4.1], pd = 50.28%	sated:romantic: median diff. = -0.8, 89% CI [-5.3, 4.2], pd = 60.27%
	sated:food: median diff. = 1.3, 89% CI [-2.4, 5.2], pd = 71.33%	sated:food: median diff. = -0.3, 89% CI [-4.7, 4.8], pd = 53.73%
P2: Interaction effect of image category and state in lag8 trials		
Not significant Not reported	**Inconclusive** BF = 0.26	**Evidence for null** BF = 0.06
	sated:romantic: median diff. = -3.6, 89% CI [-6.9, -0.1], pd = 94.97%	sated:romantic: median diff. = -1.9, 89% CI [-5.1, 1.4], pd = 82.10%
	sated:food: median diff. = -3.6, 89% CI [-6.7, 0.07], pd = 95.50%	sated:food: median diff. = -0.4, 89% CI [-3.8, 2.9], pd = 56.33%

Note. The dependent variable is accuracy (%). "*median diff.*" is the median difference.

3.4.5 Exploratory Analyses: Effects of Food Type, Gender and Age

In exploratory analyses, there was no evidence that food type (sweet or savoury) affected task performance in food trials in E1 (Appendix B.10). The evidence was inconclusive in E2. Furthermore, there was no evidence of a state by food type interaction on accuracy in either experiment.

In E1 and E2, evidence for a main effect of gender or an interaction effect of gender by state by category was inconclusive (Appendix B.11). Likely, E1 and E2 were not sufficiently statistically powered to detect these effects.

There was inconclusive evidence of a main effect of age or an interaction effect of age by state by image category in both E1 and E2 (Appendix B.12). However, supplementary frequentist analyses suggested a main effect of age and an interaction effect of age by state by image category in E2 (Appendix B.12). Given this, we used median-split age groups to establish whether the difference in performance on food trials in hungry and sated conditions differed in younger and older participants. Paired Bayesian *t*-tests were inconclusive but tended towards supporting the null, suggesting there was likely no difference in performance between food trials in the hungry and sated condition in younger participants (BF = 0.35, median difference = -0.05, 89% CI [-1.48, 1.36], pd = 52.35%) nor in older participants (BF = 0.40, median difference = 0.33, 89% CI [-1.21, 2.04], pd = 63.15%). Equivalent frequentist *t*-tests (Appendix B.12) were not significant, suggesting that state did not impact performance on food trials regardless of age.

3.5 Discussion

We have reported two attempts to replicate the main finding of PPZ – an interaction effect of hunger condition and image category on accuracy in lag2 trials of an EBA paradigm (P1), but not in lag8 trials (P2). Evidence to support both P1 and P2 was required to deem PPZ's main finding to have been replicated. We found no evidence for P1 in E1 or E2, instead finding evidence supporting the null. We found evidence to support P2 in both experiments, but without support for P1, this does not constitute even partial support for PPZ's effect. The differences in findings were despite our efforts to ensure our pre-registered replication studies were as close to PPZ as practicable. We liaised with the original authors (Piech et al., 2010), who supplied the original image sets for our use and additional details about their experiment and procedure.

We ran several experimental checks that agreed with PPZ, and thus, such differences are unlikely to explain our differences in results. In E1 and E2, as in PPZ, there was substantial evidence for a practice effect across sessions and no correlation between dietary restraint and accuracy in lag2 trials with food distractors in the hungry condition (Appendix B.1). Furthermore, evidence from E1 and E2 indicated that the paradigm successfully created a blink of attention at lag2 for all distractors, as accuracy in lag2 trials was lower than in lag8 trials. It is also unlikely that our unsuccessful replications were due to insufficient sample sizes. In both experiments, we employed a pre-registered Bayesian stopping rule during data collection. This ensured that we continued sampling until the evidence for the null hypothesis was conclusive. Our use of Bayesian analyses strengthened our conclusions by allowing us to evaluate the strength of evidence in favour of the null hypothesis rather than just rejecting it or failing to reject it (Wagenmakers et al., 2018).

Below, we consider our null result in the context of each of our studies' key limitations and differences from PPZ. We then discuss recent criticisms of the EBA paradigm and their implications for our findings.

3.5.1 Sample Demographics

The most apparent difference between E1 and E2 and PPZ is in the study samples. They differed in geographical location, gender balance, and age of the participants. E1 and E2 used samples from the British population, whereas PPZ used a sample of US undergraduates. E1 and E2 both had a more even balance of genders than in PPZ. While our experiments were insufficiently powered to study gender differences or moderation of an experimental effect by gender, exploratory analyses uncovered no evidence of them. Therefore, it is unlikely that the gender profile of our samples was responsible for our difference in results to PPZ.

PPZ did not report descriptive statistics of the age of their sample. Hence, we could not definitively determine how the ages of the E1 and E2 samples differed from theirs. However, given that their participants were undergraduate students, we infer that the mean age would have been lower and the age range narrower. We ran additional analyses to determine whether our samples' age distributions explained our difference in results. While frequentist tests suggested an age by state by image category interaction in E2 lag2 trials, follow-up *t*-tests confirmed that this was not driven by a difference in how hunger conditions impacted accuracy in food trials between younger and older participants. Thus, we conclude that there are no grounds for thinking that differences in age distributions were responsible for the difference in results on the key effect of interest (that is, hunger condition impacting performance on lag2 food trials).

We did not record participants' BMI in E1 or E2, as PPZ did not report doing so. Arumäe et al. (2019) proposed that attentional biases for food cues may only be present in specific subpopulations, such as individuals with obesity (Castellanos et al., 2009). They suggested that strict control over such moderators may be required to produce the expected effect. However, in their meta-analysis, Hardman et al. (2021) found there to be no overall relationship between an individual's weight status and their attentional bias for food cues. Another meta-analysis found no difference in attentional bias to food stimuli across several tasks between people with obesity or overweight and people with healthy weight (Hagan et al., 2020). Therefore, even if there were a BMI distribution difference between our samples and that of PPZ, we have no reason to believe this could explain the difference in result.

3.5.2 Hunger and Hunger Manipulation

It is unlikely that differences in participant hunger can explain our null results. In both E1 and E2, our manipulation of hunger was successful and produced a larger difference in mean hunger rating between sessions than in PPZ. In particular, our participants in the sated condition were less hungry than theirs. This may have been due to our additional explicit instruction to eat lunch in the hour prior to the session in the sated condition. If anything, a greater difference in hunger rating between sessions would be more likely to produce the predicted outcome – a greater attentional blink in lag2 food trials in the hungry condition than in the sated condition – because of a higher level of motivation (hunger) and a consequent increase in the value of food distractors.

3.5.3 Experimental Setting

E1 was hosted online due to COVID-19 restrictions, and hence, we could not ensure display conditions during the experiment were consistent between participants. Yet, given our within-subjects design and as participants likely completed both sessions in the same setting on the same device, display conditions are unlikely to have significantly impacted data quality. We were aware of this limitation before we conducted E1. Hence, we pre-registered our commitment to run a second experiment (E2) in a controlled laboratory setting if the first experiment produced null results. As the outcomes of E2 supported those of E1, experimental settings are unlikely to be a significant cause of the unsuccessful replications.

3.5.4 Overall Accuracy

Overall accuracy was higher in E1 and E2 than in PPZ, but it is unclear why. It is possible that our distractors were less effective at capturing the attention of our participants than in PPZ or that our participants were more able to suppress stimulus-driven attentional capture. However, we do not have sufficient data to test these speculations.

Another possibility is that our monetary compensation and accuracy feedback differed from PPZ. Participants in PPZ received monetary compensation based on their task performance; they received 10 USD if they >=80% and 20 USD if they scored >=90%. These incentives were applied to each session. PPZ also reported awarding “The best participant from each group of 20” an additional 50 USD prize. We had a similar but not identical incentive strategy, with a 5 GBP or 10 GBP gift card for participants who scored >= 80% or >= 90%, respectively, on average across both sessions. The highest-scoring participant also received a 50 GBP gift card. In addition, though, we also included a 10 GBP gift card as a show-up fee. In E1 and E2, participants were shown their average score after each block, which meant they were aware of their performance. It is unclear whether PPZ also showed participants this information, but we chose to so that participant motivation remained high throughout the task. Thus, to the extent there were differences in incentives and feedback, these were subtle.

3.5.5 Differences in Accuracy Across Image Categories and Image Sets

We found notable differences in the main effects of image category on accuracy (Appendix B.4 to B.8) between PPZ and our replications. For example, PPZ found that romantic trials had the lowest accuracy at lag2, but at lag2 in E1 and E2, romantic trials had the highest accuracy. This means that of all the distractors at lag2, romantic distractors were the most likely to create an attentional blink in PPZ but the least likely to create an attentional blink in E1 and E2.

These differences (at least in part) may be because of the image sets used in our replications; the exact image sets used in E1 and E2 were probably different to the original study, as the original authors were unable to identify the exact image subsets used from the larger image sets shared with us. This could account for the lack of state by image category interactions in our replications and those of Arumäe et al. (2019), as even subtle differences in image sets may alter whether an EBA occurs (Santacroce et al., 2023).

In these particular experiments, the characteristics of the food images used – such as calorie content or whether a food is sweet or savoury – could be an important consideration for replication. Calorie data were not available for the images we were provided. We ensured that a balanced sample of sweet and savoury food images was selected for use in food trials. Furthermore, we ran additional analyses on lag2 food trials to assess whether task performance on these trials was affected by whether a food image was sweet or savoury (Appendix B.10). Our results suggest this was unlikely. It is worth noting that this outcome is supported elsewhere in the literature. Arumäe et al. (2019) categorised their food distractor images based on their fat content (high or low) and whether they were sweet or savoury; they found that food type did not impact task performance. Furthermore, in a meta-analysis, Hardman et al. (2021) found no relationship between hunger and attentional bias for high- or low-calorie food stimuli.

We also note that if an increased EBA for food images when hungry is restricted to only a specific, narrow set of images that PPZ happen to have used, it seems unlikely to be of any broad practical or clinical importance.

3.5.6 The EBA Paradigm and Hunger as a Motivational State

Santacroce et al. (2023) recently published a comprehensive study of the EBA task itself. They showed that it is not the emotional valence of the distractor that leads to an attentional blink in an EBA task but its physical distinctiveness from filler and target images. They surmised that such distinctiveness creates a 'pop-out' effect so the distractor can capture attention, which is not achieved by the emotional content of the distractor alone. This 'pop-out' effect then results in a blink that may subsequently be magnified by the emotional content of the distractor. Santacroce et al. (2023) also found that even when an EBA occurs, it is weaker than the attentional blink produced in a conventional attentional blink paradigm, in which participants must identify two targets that appear in close succession in an RSVP.

To summarise, the EBA effect appears less reliant on the emotional valence of an image than previously thought. Hence, a change in the emotional salience of a distractor following a change in motivational state may not impact the attentional blink to any observable extent. Consequently, there appear to be significant limitations in using the EBA paradigm to study changes in attentional blinks following a shift in motivational state.

In the experiments presented here and in Arumäe et al. (2019), the motivational state of interest is hunger. As PPZ did, we assume that the emotional valence, and consequently attentional bias, of food cues will increase with increasing hunger, resulting in a more pronounced EBA (for discussion of this hypothesis within an evolutionary psychological framework, see Al-Shawaf (2016)). However, the results of Redlich et al. (2021) suggest that hunger may not be an appropriate manipulation for increasing the value of food stimuli, and a meta-analysis of 98 effect sizes found only a very weak positive correlation between hunger and attentional bias to food cues (Hardman et al., 2021). Given this, it is unlikely that hunger alone is capable of dramatically increasing the emotional valence of food cues to increase the strength of an emotional blink of attention. These considerations tend to support the possibility that PPZ's main finding may have been a false positive.

3.6 Conclusion

Methodological, demographic or cultural differences across the studies may account for our failure to replicate the original finding of interest from PPZ. However, the failure may also result from limitations of the EBA paradigm and/or a weak relationship between hunger and attentional bias for food cues. At the very least, this suggests that the key findings of PPZ have limited generalisability, and, at most, it may suggest their finding was a false positive. Maxwell et al. (2015) suggested that adopting a Bayesian approach in parallel with multiple replication attempts can help to elucidate the likelihood of the null hypothesis given the results of the replication data. We used both strategies in this present study in an attempt to conduct a rigorous replication and quantify the strength of evidence in favour of the findings of PPZ or the null hypothesis.

We did not find a relationship between hunger and the attentional capture of food cues in the present study. Our findings agree with those of Arumäe et al. (2019) but contest those of PPZ (Piech et al., 2010). Therefore, the evidence that hunger affects attentional allocation to food stimuli may be weaker than previously thought. We suggest that further replication attempts are required and that the role of hunger as a motivational driver for shifting cognitive resources towards food stimuli needs better characterisation. One such avenue could be to assess whether hunger needs to be experienced with greater intensity, over longer periods or more frequently (e.g., in populations experiencing food insecurity) to have measurable effects on food-related cognition rather than the acute hunger manipulation used here.

Chapter 4 No Effect of Hunger on the Memory of Food Images and Prices

Preface

Chapter 4 addresses the second aim of the present thesis, “To investigate the effect of hunger on food-related cognition”.

In Chapter 3, I found no evidence that hunger increases attentional capture of food cues. However, it is possible that other cognitive processes, such as memory, prioritise the retention of food cues when individuals are in a state of hunger. Doing so may provide an adaptive advantage for finding and acquiring food when a psychological and physiological drive to eat is present, as individuals may have an improved ability to remember information they have previously seen about a food source.

In Chapter 4, I report two studies, each comprising an image recognition and a price recall task, which examines whether hunger influences the memory of food images and food-related information (i.e., prices). As in Chapter 3, I focus on visual food cues due to their pervasiveness in modern life. Moreover, here, I also choose to examine the relationship between hunger and information directly related to food acquisition in day-to-day life – price – as such information is pertinent to food decision-making in a real-world context. Therefore, one might expect that participants with higher levels of hunger may be better at remembering food prices, as they must cognitively prioritise retaining more information relevant to food acquisition to solve their hunger. If food cues and food-related information are more memorable to individuals when they are hungry, and individuals with FI experience hunger differently (as shown in Chapter 2), cognitive shifts resulting from hunger may be an important driver of dietary behaviour and energy balance in people with FI. Therefore, the following chapter aims to investigate the effect of hunger on food-related cognition – the second aim of this thesis – by using recognition and recall tasks to explore the relationship between hunger and the memory of food-related information.

Publication

This chapter is published as a pre-print and has been submitted to *Adaptive Human Behaviour and Physiology* for peer review and publication. I have not changed it except to refer to other material in this thesis. Thus, I retain the use of "we" in place of "I" as I use elsewhere in this thesis. The citation information for the pre-print is as follows:

Neal, C., Pepper, G. V., Allen, C., Shannon, O. M., & Nettle, D. (2024). *Pre-print. No effect of hunger on the memory of food images and prices.* Open Science Framework. https://doi.org/10.31219/osf.io/mra7u

Author contributions

I prepared each section of the text, including figures, with guidance and critical feedback on narrative and structure from DN, GVP, CA and OMS. I led the development of the rationale and study design, with support and guidance on key decisions from GVP in Study 1 and DN, GVP and CA in Study 2. I was solely responsible for participant recruitment, data collection and data capture. I designed and conducted the data analysis with analysis design support from DN (specifically regarding derived variables and model structure). GVP, CA and OMS also contributed to the analysis protocol by sense-checking the approach and commenting on data visualisations.

Statement on the inclusion of Study 1 in the present thesis

In the following chapter, "Study 1" refers to an experiment which was previously conducted and submitted for assessment as part of my MSc Psychology thesis. However, the work presented below is entirely new for the present thesis. All text has been rewritten, thus there is no overlap in written work between the present PhD thesis and my MSc thesis. Furthermore, all analyses and data visualisations are novel as I previously used different tests, software and statistical approaches to what is presented in this chapter. In my MSc thesis, I conducted frequentist ANCOVAs using SPSS with outcome variables of (i) the difference between the number of food and non-food items correctly selected in the image recognition task, (ii) the number of food items incorrectly selected in the image recognition task, and (iii) the difference between the number of food and non-food prices. In contrast, below, I conduct Bayesian linear mixed models using R (which include item type – food or non-food – as a fixed effect) with outcome variables of (i) *d*' (a measure from Signal Detection Theory) and (ii) bias (frequency of positive responses to a food item) for the image recognition task and, for the price recall task, (iii) the number of prices correctly recalled and (iv) the mean sum of squares of the differences between the actual item price and the price recalled. I have simply used the dataset collected during my MSc to create the improved revision that constitutes Study 1 in the present chapter. Consequently, the work presented below is a fundamental revision of work previously submitted for my MSc thesis. I have completely overhauled how I operationalised and derived appropriate measures from the dataset and the statistical approach I have taken to analyse these measures to test my hypotheses.

4.1 Abstract

Food acquisition is an adaptive problem resolved via both physiological and psychological processes. Hunger could serve as a coordinating mechanism for these processes. When hunger increases, it may be beneficial to shift cognitive resources away from other adaptive problems and towards functions that increase the chances of acquiring food, such as memory for food information. However, there is limited research exploring the impacts of hunger on food-related memory, and the results are mixed. We conducted two studies investigating whether increased hunger levels improve memory for food images and prices – but not non-food images and prices – in image recognition and price recall tasks, respectively. Study 1 was an online, observational study (N = 91) using self-reported hunger as a continuous measure. Study 2 was an in-person, between-subjects interventional study (N = 102) where participants were randomly allocated to a hungry or sated condition. We predicted that higher levels of hunger would improve participants' ability to discriminate between food images they have and have not seen before and correctly recall food prices. We found no evidence of a hunger-related memory enhancement for food stimuli in either study. This contrasts with older research but supports more recent work, suggesting that this effect of hunger on food memory may be sensitive to study design and not as broadly generalisable as first thought.

4.2 Introduction

Food acquisition, intake, and choice have been driven by adaptive biological and psychological mechanisms throughout human evolutionary history to optimise nutrient availability and meet energetic demands for survival. It is important to understand the cognitive processes underpinning food-related behaviour in a modern context because of mismatches between the ancestral and modern food environment and rising levels of weight-related adverse health outcomes globally (Murray et al., 2020). Hunger is widely acknowledged as a factor significantly influencing food consumption and over-consumption (Emilien & Hollis, 2017). Hunger may act as a coordinating mediator between the psychological and physiological mechanisms required to find and consume food of sufficient quality and quantity (Al-Shawaf, 2016). If hunger occupies such a role, it is likely accompanied by cognitive adaptations that increase the likelihood of food acquisition and survival. Loewenstein (1996) described hunger as a "visceral influence" and proposed that when an individual is hungry, goal-congruent goods and behaviours (e.g., food-related stimuli or tasks) will be more desirable and therefore prioritised over their goal-incongruent counterparts (e.g., non-food-related stimuli or tasks). In the context of memory, a similar concept has been termed "the relevance hypothesis of memory facilitation" (Montagrin et al., 2013; Montagrin et al., 2021). This hypothesis suggests that goal-relevant stimuli (e.g., food) may become more arousing with increasing goal relevance (e.g., when hungry), which may mediate improved memory facilitation of goal-congruent stimuli.

Morris and Dolan (2001) found evidence to support these concepts using a yes/no image recognition task comprising food and non-food images. After viewing an initial image sequence (i.e., encoding), participants were shown sequences of "old" images (seen during encoding) and "new" images (never seen before); they used key presses throughout the sequence to indicate whether they thought it was an old or new image. The researchers experimentally manipulated the motivational significance of the food images by having the same participants complete the task in both hungry and sated states. There was an interaction between participants' motivational state and their recognition memory; in the hungry state, participants had a memory advantage for food images (relative to non-food images) that was lost once they became sated. However, the sample size was small (N = 10) as the study was centred around costly fMRI measurements. Talmi et al. (2013) expanded on their work with a larger sample size (N = 43), using a between-subjects manipulation of hunger. Using an adapted version of Morris and Dolan's (2001) image recognition task, they asked participants to encode the images under divided attention conditions to additionally establish whether hungry participants preferentially attended to food images (as supported in previous research; Channon & Hayward, 1990; Mogg et al., 1998; Piech et al., 2010; Placanica et al., 2002; Stockburger et al., 2009). Response modality differed between the studies, as Talmi et al. (2013) asked participants to recall as many images as possible in writing. Despite these differences, Talmi et al. (2013) also found that hunger enhanced memory of food images relative to non-food images.

Montagrin et al. (2021) also found memory enhancements for food stimuli associated with hunger. In their between-subjects study (N = 74), participants completed an encoding session, followed by a test session 48h later. In the test session, participants completed a written free recall of images shown in the encoding session and then a forced-choice image recognition task with paired old and new images. Participants in the hungry condition had improved memory performance for food but poorer memory performance for non-food (in both free recall and forced choice tasks) compared to the sated condition. The study differed methodologically from those previously mentioned in that only the encoding session – not the testing session –was completed in a state of hunger by those in the hungry condition. Thus, the findings suggest that feeling hungry when food items are encoded is enough to improve later recall of food items, even when testing is completed in a sated state.

Most recently, Thieleking et al. (2023) found no evidence of a relationship between subjective hunger rating and memory of food images in a yes/no image recognition task; to our knowledge, this is the only published null result of hunger's effect on food memory. In their study (N = 60), memory encoding occurred during a wanting task, assessing participants' subjective desire to eat food images and to have non-food (art) images. Despite the lack of a hunger effect and in line with other food memory research (e.g., Seitz et al., 2021; de

Vries et al., 2020), their trial-level results showed that increasing the calorie content of a food image improved recognition memory independent of their desire to eat.

In this paper, we present two studies (S1 and S2), which aim to replicate previous findings that hunger enhances image recognition memory of food stimuli but not non-food stimuli. We also extend previous research by investigating the effect that the calorie content of an image has on memory performance (as in Thieleking et al. (2023)) and exploring how hunger affects memory for food-related information relevant to food acquisition in a modern food environment. If hunger supports the successful acquisition of food via improved memory of food-related information, we may expect to see memory improvements for cues that would help obtain food (e.g., the location or price of food). In a modern environment where food is a primary good (directly relevant to the motivational state of hunger), money may be considered a secondary good that can be exchanged for food (Orquin & Kurzban, 2016). We, therefore, assume that the price of food is indirectly relevant to the motivational state of hunger and expect to see enhancements in recall of food prices in the hungry relative to the sated state.

The central aim of S1 and S2 was to establish whether hunger improves memory of food images and prices. We suggest that hunger may improve memory for food information by generally reducing cognitive performance but reducing the performance of food-related processes to a lesser extent. Alternatively, we propose that hunger may improve memory for food information by shifting cognitive resources towards improving performance for food-related processes. Three possible scenarios would support these hunger-driven effects on memory. Task performance could be:

1. Worse for non-food stimuli when hungry compared to sated, but the same for food stimuli.
2. Worse for both types of stimuli when hungry compared to sated, but lower performance reduction in food stimuli compared to non-food stimuli.
3. Better for food stimuli when hungry than sated, but the same for non-food stimuli.

If any of these scenarios are supported in our data, we expect to find a significant interaction between hunger and item type when predicting task performance. Therefore, there were two main hypotheses for each study:

H1. Hunger enhances recognition memory of food images, but not non-food images.

If H1 is true, we expect increased accuracy in food trials but not non-food trials when participants are hungry. We may also see increased response bias for food items when hungry (i.e., increased propensity to say “yes” to them, regardless of whether they are old or new).

H2. Hunger increases recall memory of food prices, but not non-food prices.

We propose two observable outcomes that would support H2: (a) higher number of food items than non-food items exactly recalled correctly when hungry, or (b) higher recall accuracy of prices of food items than non-food items when hungry.

4.3 Methods

4.3.1 Overview of Studies

We conducted two studies (S1 and S2), each comprising a food-related image recognition and price recall task. S1 was exploratory and was not pre-registered. S2 was pre-registered (pre-registration at https://osf.io/7euza). Ethical approval was granted for S1 by the Northumbria University Department of Psychology Ethics Committee (reference #26002) and for S2 by the Newcastle University Faculty of Medical Science Research Ethics Committee (reference 1281_1/14850/2021).

In S2, we improved upon the protocol of S1 to increase confidence in our results, their replicability, and their causal significance. Here, we will summarise each study, its variables, and our analytical approach. For the complete study-specific methods of S1 and S2, see Appendices C.1 and C.2, respectively.

S1 was a correlational study using self-reported hunger rating as a continuous predictor. It was conducted remotely online (due to the COVID-19 pandemic) using Qualtrics (software version August 2020). To capture a broad range of self-reported hunger ratings, we recruited participants in two waves on the same day, either side of a typical lunchtime (Wave 1 at 11:10 and Wave 2 at 14:00). For analysis, 91 participants (women = 50, men = 41; age 18-59 years old, M = 33.4, SD = 10.3) were included.

S2 was conducted in person using PsyToolKit (Stoet, 2010, 2017). It employed a between-subjects experimental manipulation of hunger (later called "state") as a categorical independent variable (IV). For analysis, there were 49 participants in the hungry condition (age 18-66 years old, *M* = 28.1, *SD* = 11.1; women = 35, men = 14) and 53 participants in the sated condition (age 18-67 years old, *M* = 31.1, *SD* = 12.0; women = 36, men = 17).

Food and non-food images (and the associated calorie content of food images) for both studies were from the food-pics database (Blechert et al., 2014). Since an international team produced the database and it contains items from various cuisines, we chose images likely to be familiar to our British study participants. We determined prices for both studies by matching items selected from food-pics to a similar item on the Tesco website (2020) to gauge an ecologically valid price. We chose Tesco as it has the largest market share of all UK supermarkets (Kantar, 2024) and is regularly reported as having mid-range prices compared to other UK supermarkets (Walsh & Simmonds, 2024).

4.3.2 Overview of Variables

Where possible, we used the same outcome variables for the image recognition and price recall task in S1 and S2 to maximise the comparability of results across both studies. For clarity, we describe image recognition outcome variables below.

In the image recognition tasks used in S1 and S2, participants had to identify whether they had seen an item before by responding "yes" they believed they had seen that item before (they believe it is an old item), or "no" they believed they had not seen it before (they believe it is a new item). Consequently, there are four possible trial outcomes: hit (H), miss (M), false alarm (FA), and correct rejection (CR) (Table 4.1).

Table 4.1. Terms used to define types of participant responses to the image recognition task.

Response	Participant has seen the item before "***Old item***"	Participant has not seen the item before "***New item***"
Yes	Hit	False alarm
No	Miss	Correct rejection

However, the hit rate alone does not tell us how well participants remember food images; participants could achieve a perfect hit rate by indiscriminately saying "yes" to every image. Signal Detection Theory (SDT) provides a means to overcome this difficulty in interpretation by using a mathematical framework to understand the strategies participants use when faced with uncertainty in decision-making. SDT allows for the differentiation between discrimination sensitivity (**d'**; i.e., the ability to differentiate "old" from "new" stimuli) and response bias (**c**; i.e., the overall tendency to respond "yes" or "no") (Abdi, 2009).

For the image recognition tasks presented in both S1 and S2, we calculated d', (see Table 4.2) to assess whether hunger increases participants' ability to discriminate between "old" and "new" food and non-food items.

In S1, c was not an appropriate measure of bias as study design limitations meant that c values for food and non-food would be directly associated. This is because participants responded to the S1 image recognition task by selecting an exact number of food and non-food words in a grid, meaning the number of selections participants made would always be constant. Thus, to measure image recognition bias in S1, we used the total number of times a participant said "yes" to food items.

However, in S2, we overcame the design challenges of S1 by using a yes/no paradigm in the response section of the task; participants were shown a sequence of images and used key presses to indicate whether they believed the image was "old" or "new", but they were not limited in how many times they could respond "yes" or "no". Consequently, we calculated c (see Table 4.2) to investigate whether hunger influences participants' tendency to respond "yes" or "no" to food and non-food items.

To assess the effect of image calorie content on task performance in the image recognition tasks, we used the binary outcome "hit" or "miss" (see Table 4.3) in our trial-level analyses. This was essential in S1, as there were no "new" images to have an associated calorie content because participants responded to the task using a grid of words. We maintained this measure for S2 for comparability to S1 and ease of interpretation. The image calorie content data were scaled (converted to z-scores) prior to analysis for both S1 and S2 due to substantial skews in the distribution of calorie content across images.

All outcome variables used in both the image recognition and price recall tasks for S1 and S2 are described in Table 4.2.

Table 4.2. Derived outcome variables for recognition and recall tasks in S1 and S2.

Task	**Measure**	**Description**	**Interpretation**	**Calculation**
Recognition	d'*	SDT measure that describes the ability to discriminate between "new" and "old" images.	A **higher** value indicates better discrimination between "old" and "new" items.	$d' = z(H) - z(FA)$
	Bias *(S1)*	Measure of positive response bias for food items (i.e., the total number of times a participant responded "yes" to a food item).	A **higher** value indicates a participant said "yes" to more food items.	-
	c* *(S2)*	SDT measure of bias that describes the tendency to say "yes" or "no".	**0** = balanced observer who "minimises conjointly the risk of a Miss and a False Alarm" (Abdi, 2009). **>0** = more "conservative" observer (i.e., more likely to say "no" than the balanced observer, thus reducing the risk of FAs) **<0** = more "liberal" observer (i.e., more likely to say "yes" than the balanced observer, thus reducing the risk of misses).	$c = -0.5(z(H) + z(FA))$
Price recall	Score	Total number of prices recalled exactly.	A **higher** value indicates better task performance.	-
	Accuracy	Mean sum of squares of the differences between the actual item price and the price recalled.	A **higher** value indicates a lower accuracy in price recall.	-

***Note*.** *We calculated these measures using the R package "psycho" (Makowski 2018).

4.3.3 Overview of Statistical Approach

Data were analysed and visualised in R (R Core Team, 2023). Our data and code are available online (doi.org/10.17605/OSF.IO/9HSDX). We fitted Bayesian linear mixed models, the structures of which are presented in Table 4.3. We used weakly informative priors of *N*(1, 10) for all parameters (McElreath, 2020). Our models included a random effect of participant (to allow for the repeated measures in our within-subjects variable (e.g., item type)), as well as gender and age as covariates. Gender was included as some evidence suggests that responses to visual food stimuli are more robust in women, particularly when hungry (Chao et al., 2017) and age was included as memory performance is known to decline with age (e.g., in recognition memory, (Bender et al., 2010)). We also fitted Bayesian logistic models to investigate the impact of calorie content on the trial outcome (Table 4.3). In these trial-level models, we included an additional random effect of food item. Furthermore, we fitted a Bayesian linear mixed model to test whether whole number pricing improved price recall in S1, as some of the prices used in S1 were whole numbers (e.g., 2.00 GBP) rather than non-whole numbers (e.g., 2.30 GBP; Table 4.3). For S2, the models we fitted (Table 4.3) had different structures to those that we pre-registered; we modified our analysis strategy as our pre-registered analysis protocol was not best suited to our data (for full details, see Appendix C.3).

We present posterior medians (or odds ratios) and their 89% credible intervals (CI) for all analyses. Although the 89% is arbitrary, it has become a convention in Bayesian data analysis (Makowski et al., 2019; McElreath, 2020). If the 89% CI does not include 0 (or 1 in the case of the trial-level Bayesian logistic regression models), we infer that the effect is likely present in our sample. To further support our inferences, we also report the probability of direction (pd) and the full region of practical equivalence (ROPE; for descriptions and interpretation of both, see Appendix C.4). In S2, we deviated from our pre-registered analysis strategy by using the 89% CI to establish the presence or absence of an effect instead of a Bayes Factor (see Appendix C.3).

In Figure 4.1, we plot alternative image recognition measures, as these are closer to the raw data and more intuitive to interpret than the SDT measures used in our models. We use the number of hits and correct rejections in place of d' for S1 and S2 and the number of hits and false alarms in place of c for S2.

Table 4.3. Structure of models used in S1 and S2.

Task	Outcome	Fixed predictors	Random effects
Recognition	Discrimination (d')	Hunger variable, Item type, Age, Gender	Participant
	Hit or miss*	Hunger variable, Calorie content, Age, Gender	Participant, Food item
	Total number of food "yes" responses (*S1 only*)	Hunger variable, Age, Gender	Participant
	Response bias (c) (*S2 only*)	Hunger variable, Item type, Age, Gender	Participant
Recall	Score	Hunger variable, Item type, Age, Gender	Participant
	Correct or not*	Hunger variable, Calorie content, Age, Gender	Participant, Food item
	Proportion of prices recalled (*S1 only*)	Whole number, Item type, Age, Gender	Participant
	Accuracy	Hunger variable, Item type, Age, Gender	Participant
	Accuracy	Hunger variable, Calorie content, Age, Gender	Participant, Food item

Note. "**Hunger variable**" in S1 is "hunger rating" (the self-reported hunger rating a participant gave on the hunger scale), and in S2 is "state" (the condition the participant was in, i.e., hungry or sated). "**Item type**" indicates whether the outcome variable is for food or non-food trials. "**Calorie content**" is the scaled number of kcal in the image of that trial. "**Hit or miss**" is a dummy coded variable stating whether the participant correctly responded to that trial in which they had seen the image previously (a hit; 1) or not (a miss; 0). "**Score**" is the total number of prices participants correctly recalled exactly. "**Whole number**" indicates whether the items had whole number pricing (see ***4.4.3 Price Recall*** for more). "**Accuracy**" is the average sum of squares of the difference between the actual price and the price the participant recalled. *Trial-level logistic mixed models; all others are linear mixed models.

4.4 Results

4.4.1 Overview

In S1, hunger ratings were higher in Wave 1 (M = 4.3, SD = 2.5, range = 0-8) than in Wave 2 (M = 3.2, SD = 2.7, range = 0-9; median difference = 1.0, 89% CI [0.1, 1.8], pd = 96.7, ROPE = 0.8). In S2, participants in the hungry condition had higher average self-reported hunger ratings (M = 5.4, SD = 1.3, range = 1-7) than participants in the sated condition (M = 0.6, SD = 0.9, range = 0-5; median difference = 4.7, 89% CI [4.4, 5.1], pd = 100.0, ROPE = 0.0).

Summary statistics for the image recognition and price recall task outcome variables in S1 and S2 are presented in Table 4.4 and Table 4.5, respectively.

4.4.2 Image Recognition

Overall, participants performed well on the image recognition task in S1 and S2 (as shown by a high hit and correct rejection totals in Figure 4.1A), but we did not observe ceiling effects.

Discrimination (d') In S1, we found no evidence to support a main effect of hunger rating (median difference = -0.01, 89% CI [-0.07, 0.05], pd = 66.5, ROPE = 97.0), or item type (median difference = -0.04, 89% CI [-0.21, 0.15], pd = 62.3, ROPE = 53.0), or a hunger rating X item type interaction effect (median difference = 0.03, 89% CI [-0.01, 0.07], pd = 91.2, ROPE = 98.2), on d' (Figure 4.1A; Appendix C.5). Thus, hunger rating and item type are unlikely to have affected participant's ability to discriminate between "old" and "new" items in S1.

While participants could better discriminate between "old" and "new" food images than "old" and "new" non-food images in S2 (main effect of item type; median difference = -0.23, 89% CI [-0.35, -0.12], pd = 100.00, ROPE = 1.1), we found no evidence of a main effect of state (median difference = 0.20, 89% CI [-0.02, 0.40], pd = 92.7, ROPE = 14.3), or a state X item type interaction effect (median difference = 0.05, 89% CI [-0.11, 0.21], pd = 70.9, ROPE = 43.0), on d' (Figure 4.1B; Appendix C.5). Thus, in line with S1, state is unlikely to have affected participants' ability to discriminate between "old" and "new" items, regardless of whether they were food or non-food in S2.

Effect of kcal content on discrimination (d') In S1, kcal content affected participants' ability to correctly respond to "old" food images (i.e., food hit rate); participants were more likely to get a hit in food trials with "old" images when the kcal content of the image was higher (median (OR) = 2.30, 89% CI [1.33, 4.33], pd = 98.9, ROPE = 2.8; Appendix C.6). Kcal content did not interact with hunger rating (median difference = 0.91, 89% CI [0.83, 1.00], pd = 94.1, ROPE = 92.1).

In contrast, in S2, kcal content did not affect participants' ability to correctly respond to "old" food images (median (OR) = 1.02, 89% CI [0.87, 1.20], pd = 57.0, ROPE = 92.6) and kcal content did not interact with state (median difference = 0.83, 89% CI [0.70, 0.98], pd = 96.9, ROPE = 46.6; Appendix C.6).

Bias and c There was no effect of hunger rating on how many food items participants said "yes" to in S1 (median difference = -0.07, 89% CI [-0.16, 0.01], pd = 91.7, ROPE = 84.9; Figure 4.1C; Appendix C.5). In S2, participants said "yes" more often to food items than non-food items (main effect of item type on c; median difference = 0.16, 89% CI [0.09, 0.22], pd = 100.0, ROPE = 0.4). However, there was no effect of state (median difference = -0.05, 89% CI [-0.16, 0.07], pd = 74.1, ROPE = 33.3), nor was there an interaction effect between state and item type on c (median difference = -0.09, 89% CI [-0.19, 0.00], pd = 94.1, ROPE = 14.3; Figure 4.1D; Appendix C.5).

Table 4.4. Summary statistics of outcome variables in S1 tasks.

	Outcome	Item	Mean (SD)	Minimum	Maximum
Image recognition	*d*' (discrimination)	Food	2.63 (0.88)	0.48	3.73
		Non-food	2.73 (0.85)	0.64	3.73
	Bias	Food	15.4 (1.3)	11	20
Price recall	Score	Food	4.3 (2.9)	0	15
		Non-food	4.9 (2.9)	0	12
	Accuracy	Food	0.31 (0.33)	0.00	2.22
		Non-food	0.83 (0.80)	0.07	6.73

Table 4.5. Summary statistics of outcome variables in S2 tasks.

	Outcome	Item	State	Mean (SD)	Minimum	Maximum
Image recognition	d' (discrimination)	Food	Hungry	1.99 (0.55)	0.89	3.80
			Sated	2.17 (0.74)	0.65	3.80
		Non-food	Hungry	1.76 (0.56)	0.76	3.35
			Sated	1.99 (1.70)	0.43	3.35
	c (bias)	Food	Hungry	-0.01 (0.35)	-0.83	0.66
			Sated	-0.07 (0.37)	-1.71	0.57
		Non-food	Hungry	0.15 (0.33)	-0.49	0.82
			Sated	-0.01 (0.38)	-1.81	0.74
Price recall	Score	Food	Hungry	4.88 (2.69)	1	11
			Sated	5.17 (2.55)	0	10
		Non-food	Hungry	4.51 (2.73)	0	12
			Sated	5.17 (2.87)	0	12
	Accuracy	Food	Hungry	0.35 (0.26)	0.01	1.12
			Sated	0.31 (0.24)	0.06	1.27
		Non-food	Hungry	0.60 (0.40)	0.02	1.70
			Sated	0.50 (0.33)	0.10	1.73

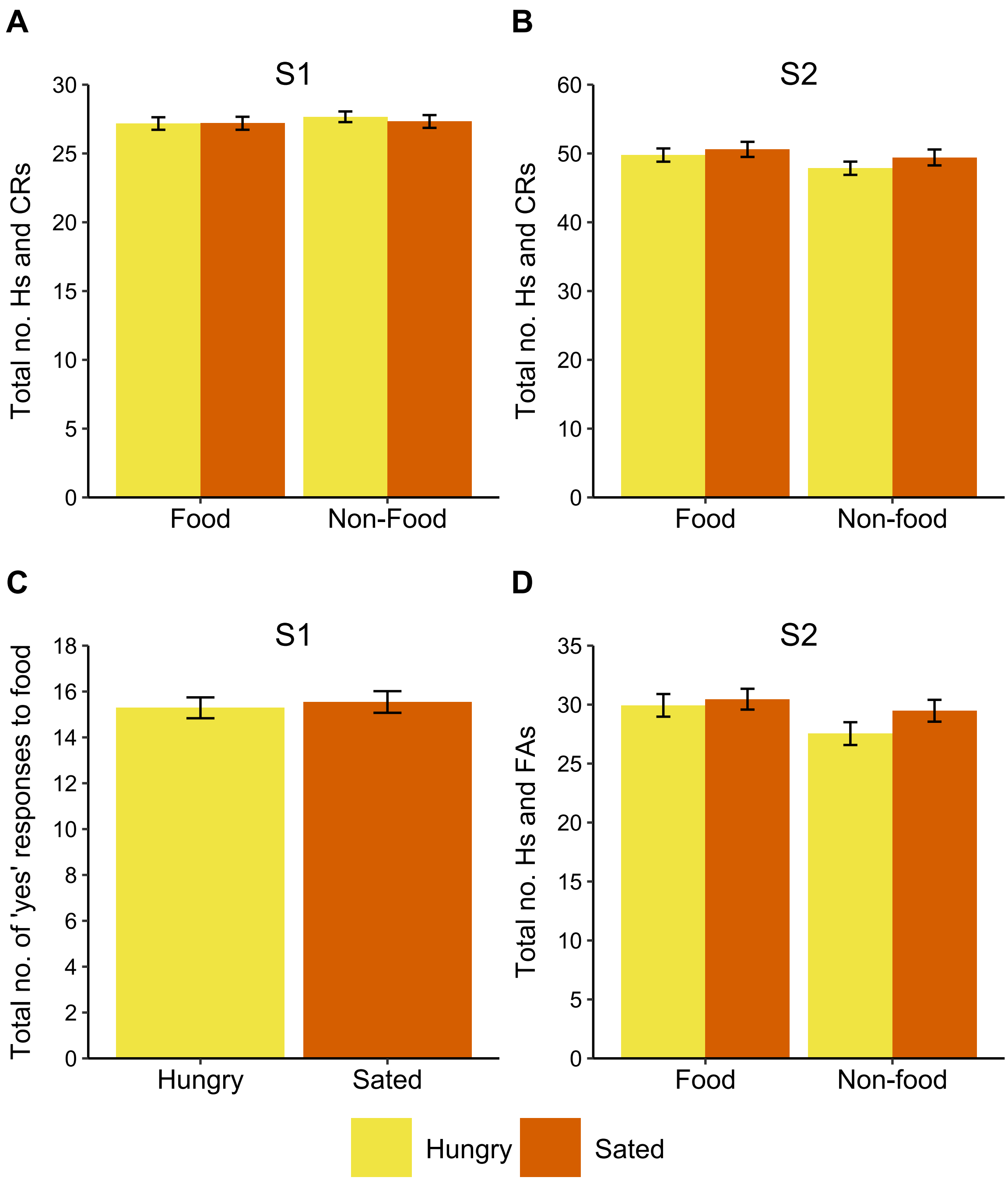


Figure 4.1. Results from S1 and S2 image recognition tasks: (**A**) S1, number of correct trials (hits + correct rejections) in food and non-food trials; (**B**) S2, number of correct trials (hits + correct rejections) in food and non-food trials; (**C**) S1, number of "yes" trials (hits + false alarms) in food trials; and (**D**) S2, number of "yes" trials (hits + false alarms) in food and non-food trials.

Note. In **A** and **C**, a median split on hunger rating was used to categorise S1 participants as hungry or sated as hunger rating was recorded as a continuous measure. Error bars indicate one standard error of the mean. Hits (H) and correct rejections (CR) are trials in which the participants correctly identified whether they had or had not previously seen an old or new image, respectively (the maximum possible value was 30 for S1 and 60 for S2). The maximum number of food items a participant could have said 'yes' to was 30 in S1. Hits (H) and false alarms (FAs) are trials in which the participant responded 'yes' (the maximum possible value was 30 in S1 and 60 in S2).

4.4.3 Price Recall

Number of prices correctly recalled In S1, participants correctly recalled the prices of more non-food items than food items (main effect of item type; median difference = 0.98, 89% CI [0.33, 1.63], pd = 99.1, ROPE = 4.3; Figure 4.2A; Appendix C.7). However, there was no evidence of a main effect of hunger rating (median difference = 0.09, 89% CI [-0.10, 0.27], pd = 76.5, ROPE = 95.1), or a hunger rating X item type interaction effect (median difference = -0.10, 89% CI [-0.24, 0.04], pd = 87.1, ROPE = 98.5). In S2, there was no main effect of item type (median difference = -0.38, 89% CI [-0.90, 0.16], pd = 87.2, ROPE = 35.2), state (median difference = 0.30, 89% CI [-0.56, 1.18], pd = 71.0, ROPE = 33.5), or a state X item type interaction (median difference = 0.36, 89% CI [-0.36, 1.09], pd = 79.3, ROPE = 33.6) on the number of prices correctly recalled (Figure 4.2B; Appendix C.7).

Effect of kcal content on the number of prices correctly recalled There was no evidence that the total kcal content of a food image predicted whether a participant correctly recalled its price in S1 (median (OR) = 0.74, 89% CI [0.44, 1.27], pd = 83.4, ROPE = 27.0; Appendix C.6) or S2 (median (OR) = 1.00, 89% CI [0.77, 1.29], pd = 50.1, ROPE = 73.2; Appendix C.6). Furthermore, there was no interaction effect in S1 between kcal content and hunger rating (median (OR) = 1.01, 89% CI [0.97, 1.06], pd = 68.3, ROPE = 100.0) or between kcal content and state in S2 (median (OR) = 0.93, 89% CI [0.77, 1.12], pd = 73.0, ROPE = 80.4).

Effect of whole number pricing on the number of prices correctly recalled S1 participants correctly recalled higher proportions of items that had whole number prices than non-whole number prices (median difference = 1.06, 89% CI [1.02, 1.11], pd = 98.7, ROPE = 11.8) and whole number price interacted with item type (median difference = 0.85, 89% CI [0.80, 0.90], pd = 100.0, ROPE = 0.0). The difference between the proportion of whole number and non-whole number prices correctly recalled was greater for food items than for non-food items (Appendix C.8).

Price recall accuracy Participants were more accurate when recalling prices of food items than non-food items in S1 (median difference = 0.59, 89% CI [0.40, 0.77], pd = 100.0, ROPE = 0.0) and S2 (median difference = 0.25, 89% CI [0.17, 0.34], pd = 100.0, ROPE = 0.0). However, in S1, there was no evidence of a main effect of hunger rating (median difference = 0.00, 89% CI [-0.04, 0.04], pd = 55.0, ROPE = 99.2) or a hunger rating X item type interaction effect on price recall accuracy (median difference = -0.02, 89% CI [-0.06, 0.02], pd = 81.0, ROPE = 96.9; Figure 4.2C; Appendix C.7). Similarly, in S2, there was no evidence of a main effect of state (median difference = -0.03, 89% CI [-0.13, 0.07], pd = 68.2, ROPE = 36.5) or a state X item type interaction effect (median difference = -0.06, 89% CI [-0.18, 0.06], pd = 80.0, ROPE = 25.7) on price recall accuracy (Figure 4.2D; Appendix C.7).

Effect of kcal content on price recall accuracy The total kcal content of a food image did not predict participants' accuracy in recalling its price in S1 (median difference = 0.02, 89% CI [-0.09, 0.13], pd = 62.5, ROPE = 72.3) or S2 (median difference = -0.05, 89% CI [-0.20, 0.10], pd = 70.0, ROPE = 56.3). Furthermore, there was no evidence of an interaction effect between kcal content and hunger rating in S1 (median difference = 0.00, 89% CI [-0.01, 0.01], pd = 52.6, ROPE = 100.0) or between kcal content and state in S2 (median difference = 0.01, 89% CI [-0.06, 0.07], pd = 56.0, ROPE = 95.9; Appendix C.9)

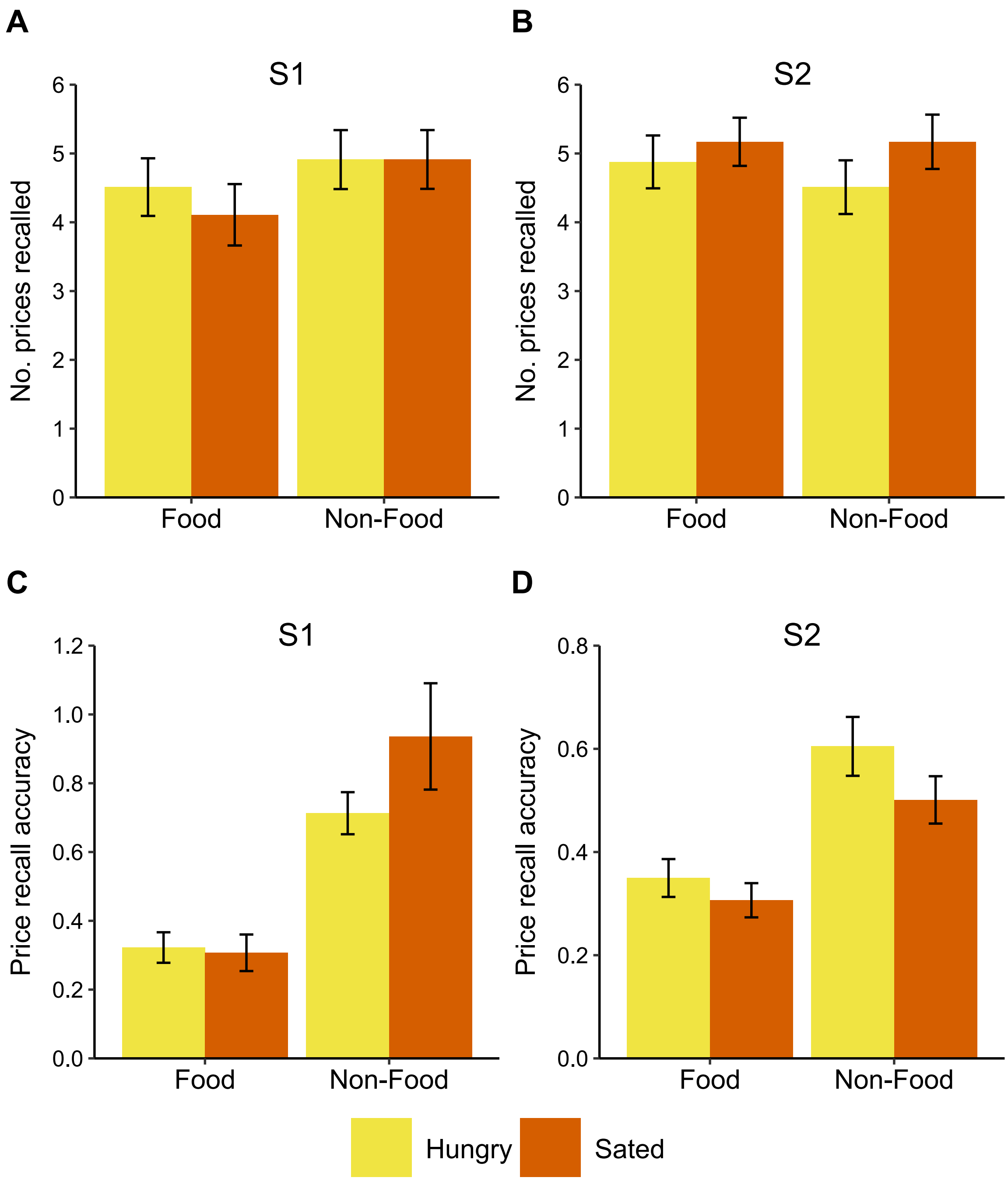


Figure 4.2. Results from S1 and S2 price recall tasks: (**A**) and (**B**), the total number of prices recalled correctly in S1 and S2, respectively, and (**C**) and (**D**), the accuracy of price recalled in food and non-food trials in S1 and S2, respectively.

Note. In **A** and **C**, a median split on hunger rating was used to categorise S1 participants as hungry or sated as hunger rating was recorded as a continuous measure. Error bars indicate one standard error of the mean. The accuracy of price recall was calculated using the sum of squares of the difference between the actual and reported prices; a higher value indicates poorer accuracy.

4.5 Discussion

We have reported two studies (S1 and S2) investigating whether current hunger levels enhance memory for food information in image recognition and price recall tasks. S1 was conducted online, and subjective hunger rating was measured as a continuous variable. In contrast, S2 was conducted in person, and hunger was experimentally manipulated between subjects. There was no evidence of our primary outcome of interest in this study: an interaction effect between hunger and item type (food or non-food stimuli) in the image recognition and price recall tasks of S1 and S2. Below, we individually consider results from the image recognition and price recall tasks and discuss their strengths and limitations in the context of previous research.

4.5.1 Image Recognition

Despite the lack of interaction between hunger and item type in both studies, there were notable differences in ancillary results from the image recognition task between S1 and S2. In S1, there were no main effects of hunger or item type on discrimination (d') between old and new items nor food item response bias. However, we observed that participants were more likely to respond positively to old food items (i.e., hits) with higher calorie content. Conversely, in S2, participants were better at discriminating between old and new food items than between old and new non-food items and responded “yes” more to food items than non-food items. However, item calorie content did not impact the likelihood that a participant achieved a hit in trials with old food items.

S1 suffered from limitations, many of which we aimed to rectify in S2. In the image recognition task, response modality was the biggest methodological difference between S1 and S2. In S1, the entire study was presented to participants on Qualtrics rather than an online platform better suited to cognitive testing due to unavoidable constraints on our time and software access for online studies. While we aimed to reproduce the yes/no image recognition task design of Morris and Dolan (2001), we were unable to because of these constraints. As a result, participants were presented with a sequence of images during encoding but responded to the task using a grid of words. Some studies that report an interaction between hunger and food memory using image recognition tasks were not limited by this difference in stimuli modality between encoding and responding (e.g., Montagrin et al., 2021; Morris & Dolan, 2001). Therefore, this may be why we did not find main effects of item type.

While we ensured that this issue with response modality was rectified in S2, we still did not find evidence for an interaction effect between hunger and item type on discrimination or response bias. Despite this, in S2 (unlike S1), there was a main effect of item type on discrimination and response bias: food items were more memorable than non-food items. This evidence aligns with existing research, indicating that food images are more memorable than other inanimate objects (see Seitz et al. (2019)) for discussion). In evolutionary terms, the higher memorability of food stimuli is advantageous for environments where food resources may become sparse (the survival processing effect; Nairne et al., 2007). However, in modern society, people are more exposed to food availability and marketing than ever (Swinburn et al., 2011). Understanding the factors that contribute to a food image’s memorability is paramount to addressing current public health concerns linked to eating behaviour.

One such factor may be the calorie content of the food image. In S1, we found that the higher the calorie content of food images, the more likely a participant would remember them. This was not true in S2. Past research has shown that participants are more likely to remember information about food with higher calorie content (e.g., New et al., 2007), which our finding in S1 supports. However, we suggest this finding should be interpreted cautiously because of the lack of the equivalent effect in S2 and the differences mentioned above in response modality in S1.

4.5.2 Price Recall

Price recall task results were more consistent between S1 and S2 than image recognition results. This likely reflects that the recall tasks in S1 and S2 were more similar than the image recognition tasks in S1 and S2. There was a main effect of item type on price recall accuracy in both studies; participants were more accurate when recalling the prices of food items compared to non-food items. However, accuracy for food items was not affected by the item's calorie content, which suggests that the properties of the individual foods may not matter as much as their status as a food; typically, stimuli essential for survival are better remembered than stimuli that are not (Nairne et al., 2007; Seitz et al., 2019). This may explain why participants were more accurate in recalling food items than non-food items. However, another explanation could be rooted in exposure and recency bias (Castel, 2005). Individuals probably shop more frequently for food than the non-food items used in the tasks. Consequently, they are more likely to have more frequent exposure to prices of the food items in the task than to the non-food prices and have their memories of how much these foods typically cost. Given that we used ecologically valid prices, this alone could improve their accuracy of price recall for food items, but not how many prices they correctly recall, which was the case in S2. Participants were more accurate at recalling food prices than non-food prices, but there was no difference between food and non-food prices in how many participants recalled exactly.

In contrast, in S1, participants exactly recalled more non-food than food prices. However, this was likely due to whole number pricing. We found evidence that whole number pricing increased the likelihood of a participant correctly recalling a price. In S1, there were three food and four non-food items with whole number prices. Therefore, participants were likely better at recalling non-food prices in S1 because more of these were whole numbers, which were more memorable. Because of this, we did not use whole number pricing in S2, which may explain the absence of a main effect of item type in the S2 correct price recall analysis. Furthermore, the calorie content of a food item is unlikely to be responsible for this discrepancy in results between S1 and S2, as items with higher calorie content were not more likely to be correctly recalled in S1 or S2.

4.5.3 Hunger Variables and Demographics

The two previous sections considered limitations which affect the image recognition and price recall tasks independently. However, other factors affecting both tasks and studies need to be considered when contextualising the results of this study.

Firstly, S1 used a subjective hunger rating as a continuous predictor rather than an experimental, between-subjects hunger manipulation, as in S2. In S1, we tried to overcome this potential limitation by recruiting participants in two waves (one prior to and one following a typical lunchtime). However, this is unlikely to have provided us with a hunger variable as robust as an experimental manipulation could have, and it introduced additional confounds, such as time of day. In previous studies, researchers have typically opted to manipulate hunger by requiring participants to fast for a given amount of time prior to the study (Montagrin et al., 2021; Morris & Dolan, 2001; Talmi et al., 2013; Thieleking et al., 2023). We used a strategy similar to Montagrin et al. (2021) in S2 and a comparison of hunger ratings in the sated and hungry conditions showed that we successfully created a difference in hunger between our conditions. It is worth noting that while there was a difference in the hunger rating scales we used in S1 and S2, this would not have influenced our outcome as, in S2, hunger ratings were only used to assess the success of our hunger manipulation, not as a predictor of task performance (as in S1).

In all analyses for both studies, we controlled for age and gender and found no evidence to suggest that either affected our outcome variables. Thus, it is unlikely that these demographic variables were responsible for the differences in S1 and S2 results. However, we did not record participants' BMI or level of dietary restraint (Herman & Mack, 1975). Consequently, we cannot comment on how these factors impacted our findings. However, a growing body of evidence suggests associations between general memory impairments, appetite control problems and, potentially, increased weight (Higgs & Spetter, 2018). Furthermore, higher dietary restraint in women has been linked to poorer memory of control stimuli than food stimuli, which was not observed in those with lower dietary restraint (Israeli & Stewart, 2001). Therefore, future studies should consider weight status and dietary restraint when designing recruitment strategies and analytical approaches so that the potential effects of these variables can be investigated and accounted for.

4.5.4 Study Design Across S1 and S2

There were several limitations in S1 that we aimed to improve upon in S2. As S1 was run online, display conditions were not consistent between participants. We rectified this in S2 by running the study in a controlled laboratory setting, and we did not see a difference in our primary effect of interest (hunger X item type interaction). Thus, the display conditions in S1 are unlikely to have impacted our findings.

Furthermore, although we did not observe ceiling effects in the S1 image recognition task, variation in task performance was relatively low; participants performed very well in the image recognition task. This may have made detecting an interaction effect between hunger and image type on recognition performance more difficult. Therefore, in S2, we made the image recognition task harder by doubling the number of images from 30 (15 food, 15 non-food) to 60 (30 food, 30 non-food). We also split up the image recognition and price recall tasks, as keeping the encoding sequence for both images and prices combined in S2 would have made the price recall task too difficult and risked creating floor effects. Although this reduced the comparability of tasks between S1 and S2, it improved the comparability with existing literature (Morris & Dolan, 2001; Thieleking et al., 2023). Interestingly, the results of these two studies show conflicting support for a hunger-related enhancement of food memory in a yes/no image recognition task. However, our results align with those of Thieleking et al. (2023) rather than the effect reported originally by Morris and Dolan (2001).

Additionally, in S2, we matched the prices of food and non-food items (see Appendix C.2) to the best of our ability whilst maintaining ecological validity. We did this to overcome any potential impact that value may have on improved memorability of an item (Villaseñor et al., 2021).

4.5.5 State of Current Evidence

Given the current state of evidence exploring memory enhancements for food stimuli when hungry is mixed, implications for our hypotheses are unclear. The present studies (and Thieleking et al. (2023), whose results correspond to our own) have larger sample sizes than those that report evidence of an effect (Montagrin et al., 2021; Morris & Dolan, 2001; Talmi et al., 2013).

We and Thieleking et al. (2023) also use images from the food-pics database (Blechert et al., 2014), a resource of images intended for use in experimental research on appetite. In contrast, Talmi et al. (2013) and Montagrin et al. (2021) created their own image sets. Experiments with image stimuli can be highly sensitive to image composition (e.g., Knebel et al., 2008), which may explain some of the discrepancies in our findings. Furthermore, the recognition task response modality of our studies and Thieleking et al. (2023) (i.e., yes/no image recognition) is most similar to that of Morris and Dolan (2001). However, Morris and Dolan (2001) had a particularly small sample size (N = 10), increasing the likelihood that their results were spurious (although we cannot directly compare our sample sizes in S1 (N = 91) and S2 (N = 102) to Morris and Dolan (2001) because their study used a crossover design). As previously mentioned, in the image recognition tasks reported by Talmi et al. (2013) and Montagrin et al. (2021), an interaction effect was only found between hunger condition and item interaction when participants responded using written free recall (note, in Montagrin et al. (2021), the authors report a "marginally significant interaction between hunger condition and item type" in a forced-choice image recognition task, but with $p = .06$). It is likely that the two different response modalities we have described (yes/no image recognition and free recall) require different memory processes (familiarity and recollection, respectively; Yonelinas, 2001). We may see differences in our outcomes because of this if hunger only impacts specific components of memory for food stimuli.

To summarise, the larger sample sizes, image sets, and image recognition response modality used in the present studies and Thieleking et al. (2023) highlight the need to consider and investigate whether previous results could be (i) false positives, (ii) confounded by stimuli selection, or (iii) a product of their task response modality. We suspect that a critical driver in the differences in results we have observed here is response modality. However, if hunger does, in fact, improve the memory of food stimuli, the mixed evidence suggests the effect is highly sensitive to deviations in study design and may not be easily replicated. Thus, the effect may only have minimal significance to and impact on real-world food decision-making.

4.6 Conclusion

We did not find a relationship between hunger and food-related memory in the present studies. While our findings align with those of Thieleking et al. (2023), they contrast with findings from other studies in the area (Montagrin et al., 2021; Morris & Dolan, 2001; Talmi et al., 2013). This may be due to methodological limitations in the studies presented and discussed here. However, it may also suggest that if humans experience hunger-driven enhancements for food recognition, this effect may not be as robust or generalisable as previous evidence might imply. We propose that further attempts be made to characterise the effect we have attempted to reproduce in this paper. Using better controls in the hunger manipulation (e.g., controlled feeding), employing within-subjects designs, counterbalancing the order of hungry and sated sessions, testing different recognition task response modalities, and matching participants based on BMI would strengthen future studies.

Chapter 5 Summary, Implications and Future Research

5.1 Thesis Summary

The previous chapters have explored the experience of hunger under conditions of FI and how hunger may impact dietary cognition. Chapter 2 showed that females experiencing FI have higher variation in daily mean hunger and hunger variation across days than those experiencing FS. However, there was no difference between groups in average hunger or hunger variation within a day. These novel results suggest that the experience of hunger differs more day-to-day under conditions of FI than FS, yet the total amount of hunger within a single day does not significantly differ. Furthermore, exploratory analyses suggest that the timing of hunger throughout the day may be different between the groups. Greater variation in hunger experiences between days likely reflects greater environmental uncertainty in food availability in the lives of those experiencing FI compared to those with FS.

Chapters 3 and 4 tested the hypothesis that hunger is associated with improved cognitive processing of food stimuli and information relevant to food acquisition. Chapter 3 comprised two conceptual replications of a study which used an Emotional Blink of Attention paradigm and an experimental hunger manipulation to show that hunger increases the perception of food cues via an attentional mechanism (Piech et al., 2010). However, I did not find evidence to support the effect reported by PPZ in either of the replications. While the failure to replicate the effect may reflect discrepancies in demographics or methods, it raises questions about the replicability of the effect and the suitability of the EBA paradigm for experimental designs such as those described in Chapter 3.

Chapter 4 was inspired by several papers which found that hunger improved memory of food images relative to non-food images (Montagrin et al., 2021; Morris & Dolan, 2001; Talmi et al., 2013). In Chapter 4, I aimed to replicate and extend their findings by investigating whether the effect extended to information ecologically relevant to food acquisition (i.e., price). I presented two studies, each comprising an image recognition and price recall task with food and non-food stimuli. Study 1 used a continuous measure of hunger, whereas Study 2 used a between-subjects experimental manipulation of hunger. There was no evidence of an effect of hunger on task performance with food stimuli in either of the studies, on either the image recognition or the price recall task. These results suggest that the positive effect of hunger on food-related memory may not be as robust or generalisable as previously implied, as it appears sensitive to deviations in the response modality of the task. As such, the effect of hunger on the memory of food may have limited impacts on dietary decision-making in real-world settings at most, but it may not have an impact at all.

5.2 Implications of Findings

As outlined in Chapter 1, the aims of this thesis were to (i) quantitatively describe and compare the experience of hunger in women with and without FI and (ii) investigate the effect of hunger on food-related cognition. The implications of the findings from each specific study are discussed in their relevant chapters. Here, I first consider additional implications that go beyond the scope of the paper presented in Chapter 2. Then, I tie the results of Chapters 3 and 4 together to examine what their null results mean for the hypothesis that hunger leads to the improved processing of food-related stimuli. Finally, I consider the implications of Chapters 2-4 together.

5.2.1 Implications of Chapter 2

The results in Chapter 2 provide new insights into the experience of hunger under conditions of FI in females. Females with FI and FS seem to differ in their experiences of hunger across days but not in the total amount of hunger experienced in the day. These findings suggest that the uncertainty and instability of FI are detectable in the hunger profile of individuals experiencing FI and highlight the importance of the temporal scale (e.g., day or week) at which hunger is analysed in FI. Considering the temporal scale of analysis may also be relevant to studying other variables that fluctuate throughout the day. For example, FI is associated with higher levels of perceived stress (Chiu et al., 2024). The results in Chapter 2 suggest that if one is interested in how the uncertainty and instability of FI impact perceived stress, it would be beneficial to assess both within- and between-day variables of perceived stress. Only assessing within-day variables of perceived stress may omit important insights relevant to how the environmental uncertainty of FI affects perceived stress through time. Thus, future research considering the psychological and physiological impacts of FI should analyse variables at different temporal scales to comprehensively capture the effects of FI over time. I collected data on perceived stress in Chapter 2 but did not present them as they were beyond the scope of the thesis, and they were not included in the manuscript submitted for peer review. However, I intend to conduct the analyses I have described here (see ***5.4 Future Directions***).

While the results in Chapter 2 do not provide any direct implications for policy, they contribute to a growing body of literature across several areas of research that suggest policymakers should minimise the uncertainty people face to improve public health. For example, stress responses are tuned to environmental uncertainty, and experiencing uncertainty in everyday life increases allostatic load and leads to negative health outcomes (de Berker et al., 2016; Peters et al., 2017). Furthermore, uncertainty or volatility in income is also detrimental to health, particularly for those with low incomes (Akanni et al., 2022). While the ultimate goal of policy should be to eradicate FI altogether, FI is likely to persist in the near future as additional drivers of FI begin to exert their influence. Current predictions about the impact of climate change on food systems security support that FI is likely to be more widespread in the coming years due to increased food prices (Abdallah et al., 2021; Dasgupta & Robinson, 2022). Even if we were to resolve FI through policy change now, the same policies are unlikely to be effective against FI that is driven by disruptions to food systems caused by climate change. In the UK, the increases in FI that have occurred since 2020 have been predominantly driven by the COVID-19 pandemic, energy price increase and the cost-of-living crisis (FAO, 2023; The Food Foundation, 2024). As the root of the problem is different, policies that would reduce current FI prevalence would likely be ineffective in tackling increased FI prevalence caused by reduced food systems security.

Therefore, it is necessary to also develop evidence-based policies which mitigate the negative health impacts of FI and prevent the further exacerbation of health inequalities with increasing FI prevalence. To do this, future research must focus on establishing causal pathways between FI and negative health outcomes. Doing so will evidence the need for policy change and provide insight into which policies are more likely to be beneficial. As discussed in Chapter 2, to achieve these goals, future research should investigate the relationships between psychological (e.g., hunger, perceived stress), behavioural (e.g., eating behaviour, physical activity), and physiological (e.g., blood glucose, energy expenditure) components of FI. The results of Chapter 2 highlight the need to consider variability and uncertainty in experience as causal drivers of health outcomes both within and between individuals and not just the average level of experience.

5.2.2 Implications of Chapters 3 and 4

In Chapters 3 and 4, hunger did not affect food-related attention or memory, respectively. Therefore, the results presented in this thesis do not support the hypothesis that hunger leads to the improved processing of food-related stimuli. Such findings are contrary to much of the published research in the area (e.g., Davidson et al., 2018; Montagrin et al., 2021; Morris & Dolan, 2001; Piech et al., 2010; Talmi et al., 2013; with few exceptions, e.g., Arumäe et al., 2019; Thieleking et al., 2023). I have considered potential reasons for this in detail in Chapters 3 and 4. Here, I will briefly discuss these null results within the broader research landscape of hunger and food-related cognition.

It is possible that other researchers have attempted similar studies to those in Chapters 3 and 4 but have not published their results if they were null. Publication bias against null results has been well-documented in the psychological sciences (e.g., Ferguson & Brannick, 2012), and this may inhibit progress in understanding the relationship between hunger and food-related cognition if results are often null, as in this thesis. If null results are produced and not published, predictions, study designs and research questions cannot be accordingly updated to reflect new understanding. It is for this reason that I have published (or submitted for publication) all of my null results.

Additionally, researcher biases for citing positive results may hinder advances in the research area. For example, since publication, Chapter 3 (Neal et al., 2023) has not been cited, while the original PPZ paper has been cited eight times in the same time frame (Scopus, 2024). This is despite the fact that PPZ is only accessible to journal subscription holders while Neal et al. (2023) is open access. These differences in citations are likely due to timing (i.e., citing authors may not have read Neal et al. (2023) before writing their manuscripts). Yet, it is worth noting that if this pattern does not change in the coming years, it could suggest that there is a selection bias for which evidence researchers choose to cite in their manuscripts. While these behaviours may not be intentional, they limit the progression of research by overinflating the coverage of results which have been shown to have questionable reproducibility; this may also contribute to increased perceived credibility of the research. As a result, researchers designing new studies may be more likely to attempt to use the methods or build upon the findings of research that may not be robust, often producing null results which have a lower probability of publication. Thus, we should be conscious and cautious of these potential biases in research investigating the effects of hunger on food-related cognition and continue to pursue replications before extending previous research.

5.2.3 Overall Implications of Thesis

In Chapter 1, I outlined the rationale for the combination of studies presented in this thesis. I argued that if hunger alters the processing of food and food-related information, it may impact dietary decision-making, particularly in a modern food environment where exposure to and availability of food is very high. If hunger impacts dietary decision-making in a way which leads to reductions in diet quality, then these changes are likely to be maladaptive for weight status and health. Thus, if hunger is experienced differently under conditions of FI, then those experiencing FI may be disproportionately vulnerable to the health consequences of hunger-driven alterations in dietary decision-making.

When considering my rationale for the body of work in this thesis, it is difficult to cohesively interpret the implications of the null results of Chapters 3 and 4 with the results of Chapter 2. It is particularly challenging because, as I have pointed out in Chapters 3 and 4, the null results may be down to methodological limitations. However, as I argued in Chapter 4, if hunger effects on dietary cognition are sensitive to minor changes in study design, then their real-world impact on dietary decision-making via alterations in the cognitive processing of food may be negligible. If their real-world impact is negligible, differences in eating behaviour and dietary decision-making in conditions of FI are unlikely to be due to different cognitive processing of food stimuli driven by the shifts in the experience of hunger. In ***5.4 Future Directions*** below, I discuss how future research might investigate the mechanisms driving differential outcomes in weight status in FS and FI.

5.3 Reflections

As I have extensively discussed the strengths and limitations of each study at the end of each chapter, this information will not be repeated here. Instead, I will discuss the challenges I have encountered while completing my doctoral studies, as I believe these reflect the strengths and limitations of this thesis as a whole as opposed to its component parts.

5.3.1 Replication, Null Results and Open Research Practices

As previously mentioned, the chapter presentation order in this thesis does not reflect the chronological order in which the studies were conducted. I began my doctoral studies by aiming to replicate the findings of PPZ in as close a replication as possible, circumstances permitting. I chose this paper specifically as I wanted to begin by investigating the effect of hunger on attention for food rather than memory for reasons described earlier (***1.3 Thesis Outline***). Furthermore, the decision to replicate PPZ's methods – rather than extend them – was guided by concerns shared between me, my supervisory team and the broader research community regarding psychology's reproducibility crisis (Open Science Collaboration, 2015). We believed it was appropriate and just to begin my doctoral research programme by attempting to replicate an effect that had been previously reported but not reproduced.

I contacted PPZ to gather methodological information about the original study, as key experimental details were missing from their paper, and to request the stimuli sets they had used. I also uploaded a comprehensive pre-registration of my protocol and predictions to ensure there was an audit trail of my approach. When E1 did not support the findings of PPZ, I recognised the value of working with the original authors and pre-registering ahead of data collection despite the delays that were associated with doing so. Taking these actions increased the credibility of the null result of E1, which became even more pertinent in light of the second null result of E2. Engaging in open research practices in Chapter 3 made me more inclined to do so when I conducted S2 in Chapter 4, as I had firsthand experience of the benefits of pre-registration. When Chapter 4 also produced null results despite meticulous planning and pre-registration, I was grateful I, again, had a public record of research intentions for the purposes of credibility and publishability. I am determined to publish the null results in Chapter 4 in a peer-reviewed journal (as I have with Chapter 3) so that others can build upon my work and minimise my contribution to publication bias and the 'file-drawer problem'.

Producing a series of null results during my doctoral studies has, at times, been extremely challenging, but it has taught me much about the scientific process and the need for replication and transparency in research. Having pre-registrations that accompany null results has given me more confidence in the results despite the outcome. Furthermore, when I have deviated from pre-registered analyses, I have been required to justify my reasons for doing so because I previously committed to something different; I now believe this to be a gold standard in scientific practice as it promotes a culture of transparency without penalty. I have found that the very act of pre-registering has encouraged me to be more thoughtful and deliberate at every step of the research process because I must justify why I have made each decision at some stage, whether it be in the pre-registration itself or explaining why I deviated from a pre-registered plan in a manuscript. Additionally, planning analyses before having a dataset to work with has made me more proficient in handling and analysing data than I would have otherwise been. It has also trained me to narrow my focus when designing a study so that I only collect data critical to the key variables of interest without getting distracted by the opportunity to collect peripheral, secondary measures. Focussing on fewer variables reduces the potential opportunity for p hacking and prevents rushing data collection to record a series of measurements, rather than giving each measure the time required to ensure high data quality. Overall, my engagement with open research practices has convinced me that they are the route to increasing scientific rigour and reproducibility in the psychological sciences and an excellent framework for training early career researchers to conduct high-quality research.

5.3.2 Statistical Approach

I used a combination of frequentist and Bayesian statistical approaches to analyse data during my doctoral studies. In Chapter 2, I was concerned with describing patterns in the data rather than testing a hypothesis. I used frequentist methods to do this and established a sample size based on a similar study. In Chapters 3 and 4, I tested hypotheses about particular effects without any idea of what the effect sizes might be, meaning a sample size based on an *a priori* power calculation would be unsuitable. Using a Bayesian approach enabled me to monitor the strength of supporting evidence for the null or alternative hypotheses during data collection without inflating the rate of false positives. Doing so meant that I did not need to pre-specify a sample size and run the risk of the studies being underpowered. It also meant that I did not collect surplus data to address my research question, which could be argued to be unethical. Instead, I pre-registered the Bayesian stopping rules that I would use to determine when I had collected enough data to confidently support the null or alternative hypothesis.

Furthermore, I used different Bayesian inferential statistics to support or reject hypotheses in Chapters 3 (Bayes factors) and 4 (pd and ROPE). I changed my approach after learning that, for small effect sizes or modest sample sizes, BFs may swing from inconclusive or supporting the null to decidedly supporting the alternative hypothesis as the amount of data supporting the alternative hypothesis increases (Huisman, 2023). In practice, if one uses a Bayesian stopping rule based on the BF, such instability of the BF could mean that data collection is prematurely stopped if the BF (incorrectly) indicates support for the null in line with one’s pre-determined stopping rule. In my experience analysing the data presented in Chapter 4, I found that conclusions drawn from the BFs often did not agree with conclusions drawn from other Bayesian inferential statistics (pd and ROPE), which describe the likelihood of effect existence and significance. I found that when the pd and ROPE supported the presence of an effect, the BF was inconclusive or null. Given these experiences and recommendations from other researchers (Makowski et al., 2019a; Makowski et al., 2019b), I used the pd and ROPE to infer support for my hypotheses in Chapter 4. As Chapter 3 was already published at this stage of my learning and reasoning, Chapter 3 retains its analyses using BFs. It is also worth noting that I did not experience these same issues with BFs in Chapter 3.

I recognise that switching between statistical approaches within a thesis is unusual, particularly given the fundamental differences in the ideas underpinning frequentist and Bayesian statistics (for example, frequentist approaches assign probabilities to observing data under a null hypothesis, whereas Bayesian approaches assign probabilities to hypotheses given the data). However, I chose the most appropriate analytical method based on each study's context and predictions, which I consider better scientific practice than strictly following one approach in the name of consistency across studies.

5.3.3 Methodological Skills Development

The EMA study presented in Chapter 2 was the last study I conducted during my doctoral studies. It allowed me to develop additional skills that have significantly contributed to my growth as an independent researcher. Firstly, I recruited participants through external partners (e.g., schools and community organisations) for the first time. Secondly, I was recruiting participants who were likely to be financially vulnerable as I was recruiting people experiencing FI. The combination of these factors meant I had to develop trusting, productive working relationships with the partner organisations so that they felt comfortable enough to advertise the study opportunity to their parents or patrons who may be vulnerable. It also meant I had to be conscious of sensitivity when designing the study and materials so that participants did not feel stigmatised by the study or while participating. The EMA study also allowed me to develop expertise in running large, remote, longitudinal studies using highly intensive data collection techniques; these skills have already proven beneficial to my future career development, as others conducting similar research have sought my advice. Having such a rich dataset has encouraged me to learn more about longitudinal data analysis, and I will be developing these skills further as the next steps I plan to take with the Chapter 2 dataset require me to use more complex analytical methods.

5.4 Future Directions

The null results presented in Chapters 3 and 4 were unexpected. For Chapter 3, there is likely little to be gained from further replication of Piech et al. (2010) in light of the issues with the EBA paradigm reported by Santacroce et al. (2023). However, it may be beneficial to conduct and replicate studies with similar experimental designs but instead use better-established attentional bias tasks, such as simpler attentional blink tasks (e.g. Ballestero-Arnau et al., 2021), to understand if hunger impacts attention for food and the extent the effect has on dietary decision-making. For Chapter 4, there are improvements to be made to the study designs of Study 1 and Study 2 before one can confidently support the null hypothesis (that hunger does not improve food-related memory). To address this, I have designed a within-subjects study with controlled feeding which otherwise uses a similar protocol to Study 2 in Chapter 4. The time constraints of my doctoral studies did not allow for this study to be conducted and included in this thesis. Studying the effect of hunger on food memory within subjects and standardising feeding before the experiment should remove noise in the data related to these factors, making the effect easier to observe if it is, in fact, present. I also hope to explore how response modality in an image recognition task impacts (as discussed in ***4.5.1 Image Recognition***) performance with food items when participants are hungry to understand if this dictates whether an effect of hunger on food image recognition is observed.

Lastly, as I have previously discussed, the causal mechanisms underlying the association between obesity and food insecurity in females are currently poorly understood (Bateson & Pepper, 2023). Furthermore, FI is associated with increased levels of perceived stress, which in turn is linked to negative health outcomes (Chiu et al., 2024; Peters et al., 2017). Thus, my next step is to use the data I have on perceived stress from the study in Chapter 2 to explore temporal differences in stress and stress variation in FI. I plan to develop methods to analyse the longitudinal relationship between hunger and stress over the course of a day and a week, as evidence suggests that increased hunger leads to increased stress, particularly in individuals with FI (Dzubur et al., 2022). Ultimately, I hope to establish whether the hunger-stress relationship differs between the conditions of FI and FS.

Moving forward in my research career, my goal is to continue investigating the links between food insecurity and health. I am particularly interested in exploring how the uncertainty of food access and temporal variability of eating behaviour impacts adiposity and trying to unpick the mechanisms which lead to sex differences in weight status under conditions of FI. To establish the causal pathways that link FI and obesity, it is critical to study biological, psychological and behavioural components of FI longitudinally rather than cross-sectionally. The results of this thesis support that we need to consider and investigate the influence that the uncertainty of experience and the environment have on health in FI. Doing so will provide the evidence required for effective policy change to overcome the growing public health challenge that FI presents.

References

Abdi, H. (2009). Signal detection theory. In McGaw, B., Peterson, P. L., Baker, E. (Eds.): *Encyclopedia of Education* (*3rd Ed*). New York: Elsevier.

Abrams, P. A. (1993). Optimal traits when there are several costs: The interaction of mortality and energy costs in determining foraging behavior. *Behavioral Ecology*, *4*(3), 246–259. https://doi.org/10.1093/beheco/4.3.246

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, *19*(6), 586–592. https://doi.org/10.1037//0278-6133.19.6.586

Akanni, L., Lenhart, O., & Morton, A. (2022). Income trajectories and self-rated health status in the UK. *SSM - Population Health*, *17*, 101035. https://doi.org/10.1016/j.ssmph.2022.101035

Allen, C., & Nettle, D. (2021). Hunger and socioeconomic background additively predict impulsivity in humans. *Current Psychology*, *40*(5), 2275–2289. https://doi.org/10.1007/s12144-019-0141-7

Al-Shawaf, L. (2016). The evolutionary psychology of hunger. *Appetite*, *105*, 591–595. https://doi.org/10.1016/j.appet.2016.06.021

Apovian, C. M. (2016). Obesity: definition, comorbidities, causes, and burden. *The American Journal Of Managed Care*, *22*(7).

Arumäe, K., Kreegipuu, K., & Vainik, U. (2019). Assessing the overlap between three measures of food reward. *Frontiers in Psychology*, *10*, 883. https://doi.org/10.3389/fpsyg.2019.00883

Atanasova, P., Kusuma, D., Pineda, E., Frost, G., Sassi, F., & Miraldo, M. (2022). The impact of the consumer and neighbourhood food environment on dietary intake and obesity-related outcomes: A systematic review of causal impact studies. *Social Science & Medicine*, *299*, 114879. https://doi.org/10.1016/j.socscimed.2022.114879

Ballestero-Arnau, M., Moreno-Sánchez, M., & Cunillera, T. (2021). Food is special by itself: Neither valence, arousal, food appeal, nor caloric content modulate the attentional bias induced by food images. *Appetite*, *156*, 104984. https://doi.org/10.1016/j.appet.2020.104984

Banerjee, S., Radak, T., Khubchandani, J., & Dunn, P. (2021). Food insecurity and mortality in American adults: results from the NHANES-linked mortality study. *Health Promotion Practice*, *22*(2), 204–214. https://doi.org/10.1177/1524839920945927

Bateson, M., & Pepper, G. V. (2023). Food insecurity as a cause of adiposity: Evolutionary and mechanistic hypotheses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1888), 20220228. https://doi.org/10.1098/rstb.2022.0228

Beaulieu, K., & Blundell, J. (2021). The Psychobiology of Hunger – A Scientific Perspective. *Topoi*, *40*(3), 565–574. https://doi.org/10.1007/s11245-020-09724-z

Ben Abdallah, M., Fekete-Farkas, M., & Lakner, Z. (2021). Exploring the link between food security and food price dynamics: a bibliometric analysis. *Agriculture*, *11*(3), 3. https://doi.org/10.3390/agriculture11030263

Benau, E. M., Orloff, N. C., Janke, E. A., Serpell, L., & Timko, C. A. (2014). A systematic review of the effects of experimental fasting on cognition☆. *Appetite*, *77*, 52–61. https://doi.org/10.1016/j.appet.2014.02.014

Bender, A. R., Naveh-Benjamin, M., & Raz, N. (2010). Associative deficit in recognition memory in a lifespan sample of healthy adults. *Psychology and Aging*, *25*(4), 940–948. https://doi.org/10.1037/a0020595

Bennett, R., Keeble, M., Zorbas, C., Sacks, G., Driessen, C., Grigsby-Duffy, L., Adams, J., Burgoine, T., & Backholer, K. (2024). The potential influence of the digital food retail environment on health: A systematic scoping review of the literature. *Obesity Reviews*, *25*(3), e13671. https://doi.org/10.1111/obr.13671

Blechert, J., Meule, A., Busch, N. A., & Ohla, K. (2014). Food-pics: An image database for experimental research on eating and appetite. *Frontiers in Psychology*, *5*, 617. https://doi.org/10.3389/fpsyg.2014.00617

Blüher, M. (2019). Obesity: global epidemiology and pathogenesis. *Nature Reviews Endocrinology*, *15*(5), 288–298. https://doi.org/10.1038/s41574-019-0176-8

Blundell, J., de Graaf, C., Hulshof, T., Jebb, S., Livingstone, B., Lluch, A., Mela, D., Salah, S., Schuring, E., van der Knaap, H., & Westerterp, M. (2010). Appetite control: methodological aspects of the evaluation of foods. *Obesity Reviews : An Official Journal of the International Association for the Study of Obesity*, *11*(3), 251–270. https://doi.org/10.1111/j.1467-789X.2010.00714.x

Blundell, J. E., Caudwell, P., Gibbons, C., Hopkins, M., Naslund, E., King, N., & Finlayson, G. (2012). Role of resting metabolic rate and energy expenditure in hunger and appetite control: a new formulation. *Disease Models & Mechanisms*, *5*(5), 608–613. https://doi.org/10.1242/dmm.009837

Booker, J. M., Chang, D. C., Stinson, E. J., Mitchell, C. M., Votruba, S. B., Krakoff, J., Gluck, M. E., & Cabeza de Baca, T. (2022). Food insecurity is associated with higher respiratory quotient and lower glucagon-like peptide 1. *Obesity*, *30*(6), 1248–1256. https://doi.org/10.1002/oby.23437

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., & van 't Veer, A. (2014). The replication recipe: what makes for a convincing replication? *Journal of Experimental Social Psychology*, *50*, 217–224. https://doi.org/10.1016/j.jesp.2013.10.005

Case, A., & Kraftman, L. (2022). *Health inequalities*. Institute for Fiscal Studies Deaton Review of Inequalities. https://ifs.org.uk/sites/default/files/2022-11/Health-inequalities-The-IFS-Deaton-Review-of-Inequalities.pdf

Castel, A. D. (2005). Memory for grocery prices in younger and older adults: The role of schematic support. *Psychology and Aging*, *20*(4), 718–721. https://doi.org/10.1037/0882-7974.20.4.718

Castellanos, E. H., Charboneau, E., Dietrich, M. S., Park, S., Bradley, B. P., Mogg, K., & Cowan, R. L. (2009). Obese adults have visual attention bias for food cue images: evidence for altered reward system function. *International Journal of Obesity*, *33*(9), Article 9. https://doi.org/10.1038/ijo.2009.138

Cepeda-Benito, A., Gleaves, D., Williams, T., & Erath, S. (2000). The development and validation of the state and trait food-cravings questionnaire. *Behavior Therapy*, *31*, 151–173. https://doi.org/10.1016/S0005-7894(00)80009-X

Channon, S., & Hayward, A. (1990). The effect of short-term fasting on processing of food cues in normal subjects. *International Journal of Eating Disorders*, *9*(4), 447–452. https://doi.org/10.1002/1098-108X(199007)9:4<447::AID-EAT2260090411>3.0.CO;2-0

Chao, A. M., Loughead, J., Bakizada, Z. M., Hopkins, C. M., Geliebter, A., Gur, R. C., & Wadden, T. A. (2017). Sex/gender differences in neural correlates of food stimuli: a systematic review of functional neuroimaging studies. *Obesity Reviews*, *18*(6), 687–699. https://doi.org/10.1111/obr.12527

Cheon, B. K., & Hong, Y.-Y. (2017). Mere experience of low subjective socioeconomic status stimulates appetite and food intake. *Proceedings of the National Academy of Sciences*, *114*(1), 72–77. https://doi.org/10.1073/pnas.1607330114

Chiu, D. T., Parker, J. E., Wiley, C. R., Epel, E. S., Laraia, B. A., Leung, C. W., & Tomiyama, A. J. (2024). Food insecurity, poor diet, and metabolic measures: the roles of stress and cortisol. *Appetite*, 107294. https://doi.org/10.1016/j.appet.2024.107294

Ciampolini, M., & Bianchi, R. (2006). Training to estimate blood glucose and to form associations with initial hunger. *Nutrition & Metabolism*, *3*(1), 42. https://doi.org/10.1186/1743-7075-3-42

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences* (Revised Edition). New York: Academic Press.

Communications Affordability Tracker. (2024, February 27). Ofcom. https://www.ofcom.org.uk/research-and-data/multi-sector-research/affordability-tracker

Conger, R. D., Ge, X., Elder Jr., G. H., Lorenz, F. O., & Simons, R. L. (1994). Economic stress, coercive family process, and developmental problems of adolescents. *Child Development*, *65*(2), 541–561. https://doi.org/10.1111/j.1467-8624.1994.tb00768.x

Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In *Handbook of Emotions* (2nd ed., pp. 91–115). New York: The Guilford Press.

Dasgupta, S., & Robinson, E. J. Z. (2022). Attributing changes in food insecurity to a changing climate. *Scientific Reports*, *12*(1), 4709. https://doi.org/10.1038/s41598-022-08696-x

Davidson, G. R., Giesbrecht, T., Thomas, A. M., & Kirkham, T. C. (2018). Pre- and postprandial variation in implicit attention to food images reflects appetite and sensory-specific satiety. *Appetite*, *125*, 24–31. https://doi.org/10.1016/j.appet.2018.01.028

de Berker, A. O., Rutledge, R. B., Mathys, C., Marshall, L., Cross, G. F., Dolan, R. J., & Bestmann, S. (2016). Computations of uncertainty mediate acute stress responses in humans. *Nature Communications*, *7*(1), 10996. https://doi.org/10.1038/ncomms10996

de Vries, R., de Vet, E., de Graaf, K., & Boesveldt, S. (2020). Foraging minds in modern environments: high-calorie and savory-taste biases in human food spatial memory. *Appetite*, *152*, 104718. https://doi.org/10.1016/j.appet.2020.104718

Department for Work & Pensions. (2023). *Family Resources Survey: Financial year 2022 to 2023.* GOV.UK. https://www.gov.uk/government/statistics/family-resources-survey-financial-year-2021-to-2022/family-resources-survey-financial-year-2021-to-2022

Dietz, W. H. (1995). Does Hunger Cause Obesity? *Pediatrics*, *95*(5), 766–767. https://doi.org/10.1542/peds.95.5.766

Dinour, L. M., Bergen, D., & Yeh, M.-C. (2007). The food insecurity–obesity paradox: a review of the literature and the role food stamps may play. *Journal of the American Dietetic Association*, *107*(11), 1952–1961. https://doi.org/10.1016/j.jada.2007.08.006

Dzubur, E., Semborski, S., Redline, B., Hedeker, D., Dunton, G. F., & Henwood, B. F. (2022). Food insecurity, hunger, stress, and homelessness among young adults. *Health Psychology*, *41*(8), 559–565. https://doi.org/10.1037/hea0001214

Eberly, L. A., Yang, L., Essien, U. R., Eneanya, N. D., Julien, H. M., Luo, J., Nathan, A. S., Khatana, S. A. M., Dayoub, E. J., Fanaroff, A. C., Giri, J., Groeneveld, P. W., & Adusumalli, S. (2021). Racial, ethnic, and socioeconomic inequities in glucagon-like peptide-1 receptor agonist use among patients with diabetes in the US. *JAMA Health Forum*, *2*(12), e214182. https://doi.org/10.1001/jamahealthforum.2021.4182

Eicher-Miller, H. A., Graves, L., McGowan, B., Mayfield, B. J., Connolly, B. A., Stevens, W., & Abbott, A. (2023). A scoping review of household factors contributing to dietary quality and food security in low-income households with school-age children in the United States. *Advances in Nutrition*, *14*(4), 914–945. https://doi.org/10.1016/j.advnut.2023.05.006

Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2022). The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment*, *29*(2), 136–151. https://doi.org/10.1177/1073191120957102

Ekman, J. B., & Hake, M. K. (1990). Monitoring starvation risk: adjustments of body reserves in greenfinches (Carduelis Chloris L.) during periods of unpredictable foraging success. *Behavioral Ecology*, *1*(1), 62–67. https://doi.org/10.1093/beheco/1.1.62

Emilien, C., & Hollis, J. H. (2017). A brief review of salient factors influencing adult eating behaviour. *Nutrition Research Reviews*, *30*(2), 233–246. https://doi.org/10.1017/S0954422417000099

Eskandari, F., Lake, A. A., Rose, K., Butler, M., & O'Malley, C. (2022). A mixed-method systematic review and meta-analysis of the influences of food environments and food insecurity on obesity in high-income countries. *Food Science & Nutrition*, *10*(11), 3689–3723. https://doi.org/10.1002/fsn3.2969

Ethica (Version 630). (2023). [Mobile App]. Ethica Data. https://ethicadata.com/

Fang, D., Thomsen, M. R., & Nayga, R. M. (2021). The association between food insecurity and mental health during the COVID-19 pandemic. *BMC Public Health*, *21*(1), 607. https://doi.org/10.1186/s12889-021-10631-0

FAO, I. (2023). *The state of food security and nutrition in the world 2023: urbanization, agrifood systems transformation and healthy diets across the rural–urban continuum*. FAO, IFAD, UNICEF, WFP, WHO. https://doi.org/10.4060/cc3017en

Farshchi, H. R., Taylor, M. A., & Macdonald, I. A. (2004a). Decreased thermic effect of food after an irregular compared with a regular meal pattern in healthy lean women. *International Journal of Obesity*, *28*(5), 653–660. https://doi.org/10.1038/sj.ijo.0802616

Farshchi, H. R., Taylor, M. A., & Macdonald, I. A. (2004b). Regular meal frequency creates more appropriate insulin sensitivity and lipid profiles compared with irregular meal frequency in healthy lean women. *European Journal of Clinical Nutrition*, *58*(7), 1071–1077. https://doi.org/10.1038/sj.ejcn.1601935

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods*, *17*(1), 120–128. https://doi.org/10.1037/a0024445

Field, M., Werthmann, J., Franken, I., Hofmann, W., Hogarth, L., & Roefs, A. (2016). The role of attentional bias in obesity and addiction. *Health Psychology*, *35*(8), 767–780. https://doi.org/10.1037/hea0000405

Finucane, M. M., Stevens, G. A., Cowan, M. J., Danaei, G., Lin, J. K., Paciorek, C. J., Singh, G. M., Gutierrez, H. R., Lu, Y., Bahalim, A. N., Farzadfar, F., Riley, L. M., & Ezzati, M. (2011). National, regional, and global trends in body-mass index since 1980: Systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9·1 million participants. *The Lancet*, *377*(9765), 557–567. https://doi.org/10.1016/S0140-6736(10)62037-5

Frankenhuis, W. E., & Nettle, D. (2020). The strengths of people in poverty. *Current Directions in Psychological Science*, *29*(1), 16–21. https://doi.org/10.1177/0963721419881154

Franklin, B., Jones, A., Love, D., Puckett, S., Macklin, J., & White-Means, S. (2012). Exploring mediators of food insecurity and obesity: a review of recent literature. *Journal of Community Health*, *37*(1), 253–264. https://doi.org/10.1007/s10900-011-9420-4

Graph Data Extractor. (2010). [Computer software]. https://sourceforge.net/projects/graphdataextrac/

Gundersen, C., & Ziliak, J. P. (2015). Food insecurity and health outcomes. *Health Affairs*, *34*(11), 1830–1839. https://doi.org/10.1377/hlthaff.2015.0645

Hagan, K. E., Alasmar, A., Exum, A., Chinn, B., & Forbush, K. T. (2020). A systematic review and meta-analysis of attentional bias toward food in individuals with overweight and obesity. *Appetite*, *151*, 104710. https://doi.org/10.1016/j.appet.2020.104710

Hardman, C. A., Jones, A., Burton, S., Duckworth, J. J., McGale, L. S., Mead, B. R., Roberts, C. A., Field, M., & Werthmann, J. (2021). Food-related attentional bias and its associations with appetitive motivation and body weight: a systematic review and meta-analysis. *Appetite*, *157*, 104986. https://doi.org/10.1016/j.appet.2020.104986

Herman, C. P., Janet, P., Patricia, P., Joyce, T., & Donna, M. (1978). Distractibility in dieters and nondieters: an alternative view of ‘externality.’ *Journal of Personality and Social Psychology*, *36*(5), 536–548. https://doi.org/10.1037/0022-3514.36.5.536

Herman, C. P., & Mack, D. (1975). Restrained and unrestrained eating. *Journal of Personality*, *43*(4), 647–660. https://doi.org/10.1111/j.1467-6494.1975.tb00727.x

Herman, C. P., & Polivy, J. (1975). Anxiety, restraint, and eating behavior. *Journal of Abnormal Psychology*, *84*(6), 666–672. https://doi.org/10.1037/0021-843X.84.6.666

Higgs, S., & Spetter, M. S. (2018). Cognitive control of eating: the role of memory in appetite and weight gain. *Current Obesity Reports*, *7*(1), 50–59. https://doi.org/10.1007/s13679-018-0296-9

Hill, B., Bergmeier, H., Incollingo Rodriguez, A. C., Barlow, F. K., Chung, A., Ramachandran, D., Savaglio, M., & Skouteris, H. (2021). Weight stigma and obesity-related policies: A systematic review of the state of the literature. *Obesity Reviews*, *22*(11), e13333. https://doi.org/10.1111/obr.13333

Holliday, A., Johnson, K. O., Kaiseler, M., & Crabtree, D. R. (2021). APPetite: Validation of a smartphone app-based tool for the remote measure of free-living subjective appetite. *British Journal of Nutrition*, *129*(9), 1–30. https://doi.org/10.1017/S0007114521003512

Hopkins, M., Beaulieu, K., Myers, A., Gibbons, C., & Blundell, J. E. (2017). Mechanisms responsible for homeostatic appetite control: theoretical advances and practical implications. *Expert Review of Endocrinology & Metabolism*, *12*(6), 401–415. https://doi.org/10.1080/17446651.2017.1395693

Hoyland, A., Dye, L., & Lawton, C. L. (2009). A systematic review of the effect of breakfast on the cognitive performance of children and adolescents. *Nutrition Research Reviews*, *22*(2), 220–243. https://doi.org/10.1017/S0954422409990175

Hruby, A., Manson, J. E., Qi, L., Malik, V. S., Rimm, E. B., Sun, Q., Willett, W. C., & Hu, F. B. (2016). Determinants and consequences of obesity. *American Journal of Public Health*, *106*(9), 1656–1662. https://doi.org/10.2105/AJPH.2016.303326

Huisman, L. (2023). Are p-values and Bayes factors valid measures of evidential strength? *Psychonomic Bulletin & Review*, *30*(3), 932–941. https://doi.org/10.3758/s13423-022-02205-x

Israeli, A. L., & Stewart, S. H. (2001). Memory bias for forbidden food cues in restrained eaters. *Cognitive Therapy and Research*, *25*(1), 37–48. https://doi.org/10.1023/A:1026422731313

James, J. E., & Rogers, P. J. (2005). Effects of caffeine on performance and mood: withdrawal reversal is the most plausible explanation. *Psychopharmacology*, *182*(1), 1–8. https://doi.org/10.1007/s00213-005-0084-6

Kantar. (2024). *Grocery Market Share*. https://www.kantarworldpanel.com/en/grocery-market-share/great-britain

Kaufman, M. R., Eschliman, E. L., & Karver, T. S. (2023). Differentiating sex and gender in health research to achieve gender equity. *Bulletin of the World Health Organization*, *101*(10), 666–671. https://doi.org/10.2471/BLT.22.289310

Knebel, J.-F., Toepel, U., Hudry, J., Le Coutre, J., & Murray, M. M. (2008). Generating controlled image sets in cognitive neuroscience research. *Brain Topography*, *20*(4), 284–289. https://doi.org/10.1007/s10548-008-0046-5

Kowaleski-Jones, L., Wen, M., & Fan, J. X. (2019). Unpacking the paradox: testing for mechanisms in the food insecurity and BMI association. *Journal of Hunger & Environmental Nutrition*, *14*(5), 683–697. https://doi.org/10.1080/19320248.2018.1464997

Kruschke, J. (2014). *Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan*. New York: Academic Press.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual.* (Technical Report A-8). University of Florida. https://www2.unifesp.br/dpsicobio/adap/instructions.pdf

Laraia, B. A. (2013). Food insecurity and chronic disease. *Advances in Nutrition*, *4*(2), 203–212. https://doi.org/10.3945/an.112.003277

Larson, N. I., & Story, M. T. (2011). Food insecurity and weight status among U.S. children and families: a review of the literature. *American Journal of Preventive Medicine*, *40*(2), 166–173. https://doi.org/10.1016/j.amepre.2010.10.028

Leung, C. W., Epel, E. S., Ritchie, L. D., Crawford, P. B., & Laraia, B. A. (2014). Food insecurity is inversely associated with diet quality of lower-income adults. *Journal of the Academy of Nutrition and Dietetics*, *114*(12), 1943-1953.e2. https://doi.org/10.1016/j.jand.2014.06.353

Loewenstein, G. (1996). Out of control: visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, *65*(3), 272–292. https://doi.org/10.1006/obhd.1996.0028

Lucas, R. E., Wallsworth, C., Anusic, I., & Donnellan, M. B. (2021). A direct comparison of the day reconstruction method (DRM) and the experience sampling method (ESM). *Journal of Personality and Social Psychology*, *120*(3), 816–835. https://doi.org/10.1037/pspp0000289

Lucero-Prisno III, D. E., Shomuyiwa, D. O., Kouwenhoven, M. B. N., Dorji, T., Odey, G. O., Miranda, A. V., Ogunkola, I. O., Adebisi, Y. A., Huang, J., Xu, L., Obnial, J. C., Huda, A., Thepanondh, S., Dayrit, M. M., Evardone, S. B., Lamawansa, M. D., Solomon Ethiopia, S., Aziato, L., Adongo, P. B., … Wong, M. C. S. (2023). Top 10 public health challenges to track in 2023: shifting focus beyond a global pandemic. *Public Health Challenges*, *2*(2), e86. https://doi.org/10.1002/puh2.86

Lutter, M., & Nestler, E. J. (2009). Homeostatic and hedonic signals interact in the regulation of food intake. *The Journal of Nutrition*, *139*(3), 629–632. https://doi.org/10.3945/jn.108.097618

Makowski, D., Ben-Shachar, M. S., Chen, S. H. A., & Lüdecke, D. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology*, *10*, 2767. https://doi.org/10.3389/fpsyg.2019.02767

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, *40*, 1541. https://doi.org/10.21105/joss.01541

Malik, V. S., Willet, W. C., & Hu, F. B. (2020). Nearly a decade on—trends, risk factors and policy implications in global obesity. *Nature Reviews Endocrinology*, *16*(11), 615–616. https://doi.org/10.1038/s41574-020-00411-y

Martin, M. S., Maddocks, E., Chen, Y., Gilman, S. E., & Colman, I. (2016). Food insecurity and mental illness: disproportionate impacts in the context of perceived stress and social isolation. *Public Health*, *132*, 86–91. https://doi.org/10.1016/j.puhe.2015.11.014

Mattes, R. (1990). Hunger ratings are not a valid proxy measure of reported food intake in humans. *Appetite*, *15*(2), 103–113. https://doi.org/10.1016/0195-6663(90)90043-8

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, *70*(6), 487–498. https://doi.org/10.1037/a0039400

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and STAN* (2nd ed.). New York: Chapman and Hall/CRC. https://doi.org/10.1201/9780429029608

McGill, R., Anwar, E., Orton, L., Bromley, H., Lloyd-Williams, F., O'Flaherty, M., Taylor-Robinson, D., Guzman-Castillo, M., Gillespie, D., Moreira, P., Allen, K., Hyseni, L., Calder, N., Petticrew, M., White, M., Whitehead, M., & Capewell, S. (2015). Are interventions to promote healthy eating equally effective for all? Systematic review of socioeconomic inequalities in impact. *BMC Public Health*, *15*(1), 457. https://doi.org/10.1186/s12889-015-1781-7

McKiernan, F., Houchins, J. A., & Mattes, R. D. (2008). Relationships between human thirst, hunger, drinking, and feeding. *Physiology & Behavior*, *94*(5), 700–708. https://doi.org/10.1016/j.physbeh.2008.04.007

Mogg, K., Bradley, B. P., Hyare, H., & Lee, S. (1998). Selective attention to food-related stimuli in hunger: Are attentional biases specific to emotional and psychopathological states, or are they also found in normal drive states? *Behaviour Research and Therapy*, *36*(2), 227–237. https://doi.org/10.1016/S0005-7967(97)00062-4

Monk, R. L., Heim, D., Qureshi, A., & Price, A. (2015). “I have no clue what I drunk last night” Using smartphone technology to compare in-vivo and retrospective self-reports of alcohol consumption. *PLOS ONE*, *10*(5), e0126209. https://doi.org/10.1371/journal.pone.0126209

Montagrin, A., Brosch, T., & Sander, D. (2013). Goal conduciveness as a key determinant of memory facilitation. *Emotion*, *13*(4), 622–628. https://doi.org/10.1037/a0033066

Montagrin, A., Martins-Klein, B., Sander, D., & Mather, M. (2021). Effects of hunger on emotional arousal responses and attention/memory biases. *Emotion*, *21*(1), 148–158. https://doi.org/10.1037/emo0000680

Morris, J. S., & Dolan, R. J. (2001). Involvement of human amygdala and orbitofrontal cortex in hunger-enhanced memory for food stimuli. *The Journal of Neuroscience*, *21*(14), 5304–5310. https://doi.org/10.1523/JNEUROSCI.21-14-05304.2001

Morselli, L. L., Amjad, R., James, R., Kindel, T. L., Kwitek, A. E., Williams, J. S., Grobe, J. L., & Kidambi, S. (2024). Diet in food insecurity: a mediator of metabolic health? *Journal of the Endocrine Society*, *8*(6), bvae062. https://doi.org/10.1210/jendso/bvae062

Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review*, *12*(4), 654–661. https://doi.org/10.3758/BF03196754

Murray, C. J. L., Aravkin, A. Y., Zheng, P., Abbafati, C., Abbas, K. M., Abbasi-Kangevari, M., Abd-Allah, F., Abdelalim, A., Abdollahi, M., Abdollahpour, I., Abegaz, K. H., Abolhassani, H., Aboyans, V., Abreu, L. G., Abrigo, M. R. M., Abualhasan, A., Abu-Raddad, L. J., Abushouk, A. I., Adabi, M., … Lim, S. S. (2020). Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, *396*(10258), 1223–1249. https://doi.org/10.1016/S0140-6736(20)30752-2

Nairne, J., Thompson, S., & Pandeirada, J. (2007). Adaptive memory: survival processing enhances retention. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*, 263–273. https://doi.org/10.1037/0278-7393.33.2.263

Neal, C., Pepper, G. V., Allen, C., & Nettle, D. (2023). No effect of hunger on attentional capture by food cues: Two replication studies. *Appetite*, *191*, 107065. https://doi.org/10.1016/j.appet.2023.107065

Nettle, D. (2017). Does hunger contribute to socioeconomic gradients in behavior? *Frontiers in Psychology*, *8*, 229550.

Nettle, D., Andrews, C., & Bateson, M. (2017). Food insecurity as a driver of obesity in humans: the insurance hypothesis. *The Behavioral and Brain Sciences*, *40*, e105. https://doi.org/10.1017/S0140525X16000947

Nettle, D., & Bateson, M. (2019). Food-insecure women eat a less diverse diet in a more temporally variable way: evidence from the US National Health and Nutrition Examination

Survey, 2013-4. *Journal of Obesity*, *2019*, e7174058. https://doi.org/10.1155/2019/7174058

New, J., Krasnow, M. M., Truxaw, D., & Gaulin, S. J. C. (2007). Spatial adaptations for plant foraging: women excel and calories count. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1626), 2679–2684. https://doi.org/10.1098/rspb.2007.0826

NHS. (2019). *NHS Long Term Plan*. https://www.longtermplan.nhs.uk/online-version/

Norgan, N. G. (1997). The beneficial effects of body fat and adipose tissue in humans. *International Journal of Obesity and Related Metabolic Disorders*, *21*(9), 738–746. https://doi.org/10.1038/sj.ijo.0800473

Office for Health Improvement & Disparities. (2023). *Inequality Tools—Health Inequalities Dashboard*. https://fingertips.phe.org.uk/profile/inequality-tools

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Orquin, J. L., & Kurzban, R. (2016). A meta-analysis of blood glucose effects on human decision making. *Psychological Bulletin*, *142*(5), 546–567. https://doi.org/10.1037/bul0000035

Paquet, C. (2019). Environmental influences on food behaviour. *International Journal of Environmental Research and Public Health*, *16*(15), 2763. https://doi.org/10.3390/ijerph16152763

Peters, A., McEwen, B. S., & Friston, K. (2017). Uncertainty and stress: why it causes diseases and how it is mastered by the brain. *Progress in Neurobiology*, *156*, 164–188. https://doi.org/10.1016/j.pneurobio.2017.05.004

Piech, R. M., Pastorino, M. T., & Zald, D. H. (2010). All I saw was the cake: hunger effects on attentional capture by visual food cues. *Appetite*, *54*(3), 579–582. https://doi.org/10.1016/j.appet.2009.11.003

Piperata, B. A., Scaggs, S. A., Dufour, D. L., & Adams, I. K. (2023). Measuring food insecurity: an introduction to tools for human biologists and ecologists. *American Journal of Human Biology*, *35*(2), e23821. https://doi.org/10.1002/ajhb.23821

Placanica, J. L., Faunce, G. J., & Soames Job, R. F. (2002). The effect of fasting on attentional biases for food and body shape/weight words in high and low eating disorder inventory scorers. *International Journal of Eating Disorders*, *32*(1), 79–90. https://doi.org/10.1002/eat.10066

R Core Team. (2023). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Redlich, D., Memmert, D., & Kreitz, C. (2021). Does hunger promote the detection of foods? The effect of value on inattentional blindness. *Psychological Research*. https://doi.org/10.1007/s00426-021-01480-y

Reeder, N., Tolar-Peterson, T., Bailey, R. H., Cheng, W.-H., & Evans, M. W. (2022). Food insecurity and depression among US Adults: NHANES 2005–2016. *Nutrients*, *14*(15), 15. https://doi.org/10.3390/nu14153081

Reichenberger, J., Pannicke, B., Arend, A.-K., Petrowski, K., & Blechert, J. (2020). Does stress eat away at you or make you eat? EMA measures of stress predict day to day food craving and perceived food intake as a function of trait stress-eating. *Psychology & Health*, *36*(2), 129–147. https://doi.org/10.1080/08870446.2020.1781122

Rincón-Gallardo Patiño, S., Zhou, M., Da Silva Gomes, F., Lemaire, R., Hedrick, V., Serrano, E., & Kraak, V. I. (2020). Effects of menu labeling policies on transnational restaurant chains to promote a healthy diet: a scoping review to inform policy and research. *Nutrients*, *12*(6), 6. https://doi.org/10.3390/nu12061544

Santacroce, L. A., Swami, A. L., & Tamber-Rosenau, B. J. (2023). More than a feeling: the emotional attentional blink relies on non-emotional “pop out,” but is weak compared to the attentional blink. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-023-02677-6

Scheer, F. A. J. L., Morris, C. J., & Shea, S. A. (2013). The internal circadian clock increases hunger and appetite in the evening independent of food intake and other behaviors. *Obesity*, *21*(3), 421–423. https://doi.org/10.1002/oby.20351

Schmidt, C., Collette, F., Cajochen, C., & Peigneux, P. (2007). A time to think: circadian rhythms in human cognition. *Cognitive Neuropsychology*, *24*(7), 755–789. https://doi.org/10.1080/02643290701754158

Scopus. (2024). *Scopus—125 documents that cite: All I saw was the cake. Hunger effects on attentional capture by visual food cues*. Scopus- Elsevier. https://www.scopus.com/results/citedbyresults.uri?sort=plf-f&cite=2-s2.0-77953612230&src=s&imp=t&sid=36f4926add1b300e5e26dec5dc032afa&sot=cite&sdt=a&sl=0&origin=inward&editSaveSearch=&txGid=dd0ce198b055fa22402254be80f0d0c4

Seitz, B. M., Blaisdell, A. P., Polack, C. W., & Miller, R. R. (2019). The role of biological significance in human learning and memory. *International Journal of Comparative Psychology*, *32*. https://doi.org/10.46867/ijcp.2019.32.03.02

Seitz, B. M., Blaisdell, A. P., & Tomiyama, A. J. (2021). Calories count: memory of eating is evolutionarily special. *Journal of Memory and Language*, *117*, 104192. https://doi.org/10.1016/j.jml.2020.104192

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Shinwell, J., Bateson, M., Nettle, D., & Pepper, G. V. (2022). Food insecurity and patterns of dietary intake in a sample of UK adults. *British Journal of Nutrition*, *128*(4), 770–777. https://doi.org/10.1017/S0007114521003810

Stevenson, R. J. (2023). The psychological basis of hunger and its dysfunctions. *Nutrition Reviews*, nuad092. https://doi.org/10.1093/nutrit/nuad092

Stevenson, R. J., Yeomans, M. R., & Francis, H. M. (2023). Human hunger as a memory process. *Psychological Review*. https://doi.org/10.1037/rev0000413

Stinson, E. J., Votruba, S. B., Venti, C., Perez, M., Krakoff, J., & Gluck, M. E. (2018). Food insecurity is associated with maladaptive eating behaviors and objectively measured overeating. *Obesity*, *26*(12), 1841–1848. https://doi.org/10.1002/oby.22305

Stockburger, J., Schmälzle, R., Flaisch, T., Bublatzky, F., & Schupp, H. T. (2009). The impact of hunger on food cue processing: An event-related brain potential study. *NeuroImage*, *47*(4), 1819–1829. https://doi.org/10.1016/j.neuroimage.2009.04.071

Stoet, G. (2010). PsyToolkit: a software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: a novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Stone, A. A., Schneider, S., & Smyth, J. M. (2023). Evaluation of pressing issues in ecological momentary assessment. *Annual Review of Clinical Psychology*, *19*(1), 107–131. https://doi.org/10.1146/annurev-clinpsy-080921-083128

Stubbs, R. J., Hughes, D. A., Johnstone, A. M., Rowley, E., Reid, C., Elia, M., Stratton, R., Delargy, H., King, N., & Blundell, J. E. (2000). The use of visual analogue scales to assess motivation to eat in human subjects: a review of their reliability and validity with an evaluation of new hand-held computerized systems for temporal tracking of appetite ratings. *British Journal of Nutrition*, *84*(4), 405–415. https://doi.org/10.1017/S0007114500001719

Swinburn, B. A., Sacks, G., Hall, K. D., McPherson, K., Finegood, D. T., Moodie, M. L., & Gortmaker, S. L. (2011). The global obesity pandemic: shaped by global drivers and local environments. *The Lancet*, *378*(9793), 804–814. https://doi.org/10.1016/S0140-6736(11)60813-1

Talmi, D., Ziegler, M., Hawksworth, J., Lalani, S., Herman, C. P., & Moscovitch, M. (2013). Emotional stimuli exert parallel effects on attention and memory. *Cognition and Emotion*, *27*(3), 530–538. https://doi.org/10.1080/02699931.2012.722527

Tesco. (2020). *Online food shopping and grocery delivery—Tesco Groceries*. https://www.tesco.com/groceries/

The Food Foundation. (2024). *Food Insecurity Tracking*. https://foodfoundation.org.uk/initiatives/food-insecurity-tracking

Theis, D. R., & White, M. (2021). Is obesity policy in England fit for purpose? Analysis of government strategies and policies, 1992–2020. *The Milbank Quarterly*, *99*(1), 126–170. https://doi.org/10.1111/1468-0009.12498

Thieleking, R., Medawar, E., Villringer, A., Beyer, F., & Witte, A. V. (2023). Neurocognitive predictors of food memory in healthy adults – a preregistered analysis. *Neurobiology of Learning and Memory*, *205*, 107813. https://doi.org/10.1016/j.nlm.2023.107813

UK Government. (2022). *Regional ethnic diversity*. GOV.UK. https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/regional-ethnic-diversity/latest/

UK Government. (2024). *Healthy eating among adults*. GOV.UK. https://www.ethnicity-facts-figures.service.gov.uk/health/diet-and-exercise/healthy-eating-of-5-a-day-among-adults/latest/

USDA Economic Research Services. (2023a). *USDA ERS - Measurement*. https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/measurement/

USDA Economic Research Services. (2023b). *USDA ERS - Survey Tools*. https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/survey-tools/#six

Valdez, P., Reilly, T., & Waterhouse, J. (2008). Rhythms of mental performance. *Mind, Brain, and Education*, *2*(1), 7–16. https://doi.org/10.1111/j.1751-228X.2008.00023.x

Villaseñor, J. J., Sklenar, A. M., Frankenstein, A. N., Levy, P. U., McCurdy, M. P., & Leshikar, E. D. (2021). Value-directed memory effects on item and context memory. *Memory & Cognition*, *49*(6), 1082–1100. https://doi.org/10.3758/s13421-021-01153-6

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

Walsh, H., & Simmonds, E. (2024). *Cheapest supermarkets 2024*. Which? https://www.which.co.uk/reviews/supermarkets/article/supermarket-price-comparison-aPpYp9j1MFin

Wang, J.-Y., Wang, Q.-W., Yang, X.-Y., Yang, W., Li, D.-R., Jin, J.-Y., Zhang, H.-C., & Zhang, X.-F. (2023). GLP−1 receptor agonists for the treatment of obesity: role as a promising approach. *Frontiers in Endocrinology*, *14*, 1085799. https://doi.org/10.3389/fendo.2023.1085799

Werthmann, J., Jansen, A., & Roefs, A. (2015). Worry or craving? A selective review of evidence for food-related attention biases in obese individuals, eating-disorder patients, restrained eaters and healthy samples. *Proceedings of the Nutrition Society*, *74*(2), 99–114. https://doi.org/10.1017/S0029665114001451

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical evidence in experimental psychology: an empirical comparison using 855 t tests. *Perspectives on Psychological Science*, *6*(3), 291–298. https://doi.org/10.1177/1745691611406923

Whitley, H. P., Trujillo, J. M., & Neumiller, J. J. (2023). Special report: potential strategies for addressing GLP-1 and dual GLP-1/GIP receptor agonist shortages. *Clinical Diabetes*, *41*(3), 467–473. https://doi.org/10.2337/cd23-0023

WHO. (2022). *WHO European Regional Obesity Report 2022*. WHO Regional Office for Europe. https://iris.who.int/bitstream/handle/10665/353747/9789289057738-eng.pdf

Wilding, J. P. H., Batterham, R. L., Davies, M., Van Gaal, L. F., Kandler, K., Konakli, K., Lingvay, I., McGowan, B. M., Oral, T. K., Rosenstock, J., Wadden, T. A., Wharton, S., Yokote, K., Kushner, R. F., & Step, 1 Study Group. (2022). Weight regain and cardiometabolic effects after withdrawal of semaglutide: The STEP 1 trial extension. *Diabetes, Obesity and Metabolism*, *24*(8), 1553–1564. https://doi.org/10.1111/dom.14725

Williams, E. P., Mesidor, M., Winters, K., Dubbert, P. M., & Wyatt, S. B. (2015). Overweight and obesity: prevalence, consequences, and causes of a growing public health problem. *Current Obesity Reports*, *4*(3), 363–370. https://doi.org/10.1007/s13679-015-0169-4

Witter, M. S., & Swaddle, J. P. (1997). Mass regulation in juvenile Starlings: response to change in food availability depends on initial body mass. *Functional Ecology*, *11*(1), 11–15. https://doi.org/10.1046/j.1365-2435.1997.00041.x

Wrzus, C., & Neubauer, A. B. (2022). Ecological Momentary Assessment: A Meta-Analysis on Designs, Samples, and Compliance Across Research Fields. *Assessment*, *30*(3), 825–846. https://doi.org/10.1177/10731911211067538

Yonelinas, A. P. (2001). Components of episodic memory: the contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1413), 1363–1374. https://doi.org/10.1098/rstb.2001.0939

Zafon, C. (2007). Oscillations in total body fat content through life: an evolutionary perspective. *Obesity Reviews*, *8*(6), 525–530. https://doi.org/10.1111/j.1467-789X.2007.00377.x

Appendices

A Chapter 2 Supplementary Materials

A.1 Summary of Deviations From Pre-registered Protocol and Analysis

A summary of our deviations from our preregistered protocol and analyses.

Pre-registered action	Deviation
Recruit participants from parent populations of schools where over 40% of the school’s pupils have been eligible for free school meals (FSM) at any time during the past six years.	Significant recruitment challenges meant that we could not reach our target sample size via schools alone, so we also used target social media advertising to recruit participants.
Match FI individuals to FS individuals based on their age, BMI and the school they were recruited from after participants expressed taking part in the study.	It was not possible or practical to achieve this matching while also reaching our target sample size. Therefore, we dropped our matching procedure to prioritise our recruitment aims.
Use Bayesian statistical methods to analyse the data.	We used equivalent frequentist methods instead for ease of implementation and understanding.

A.2 The Perceived Stress Scale Measure Included at Each Momentary Assessment

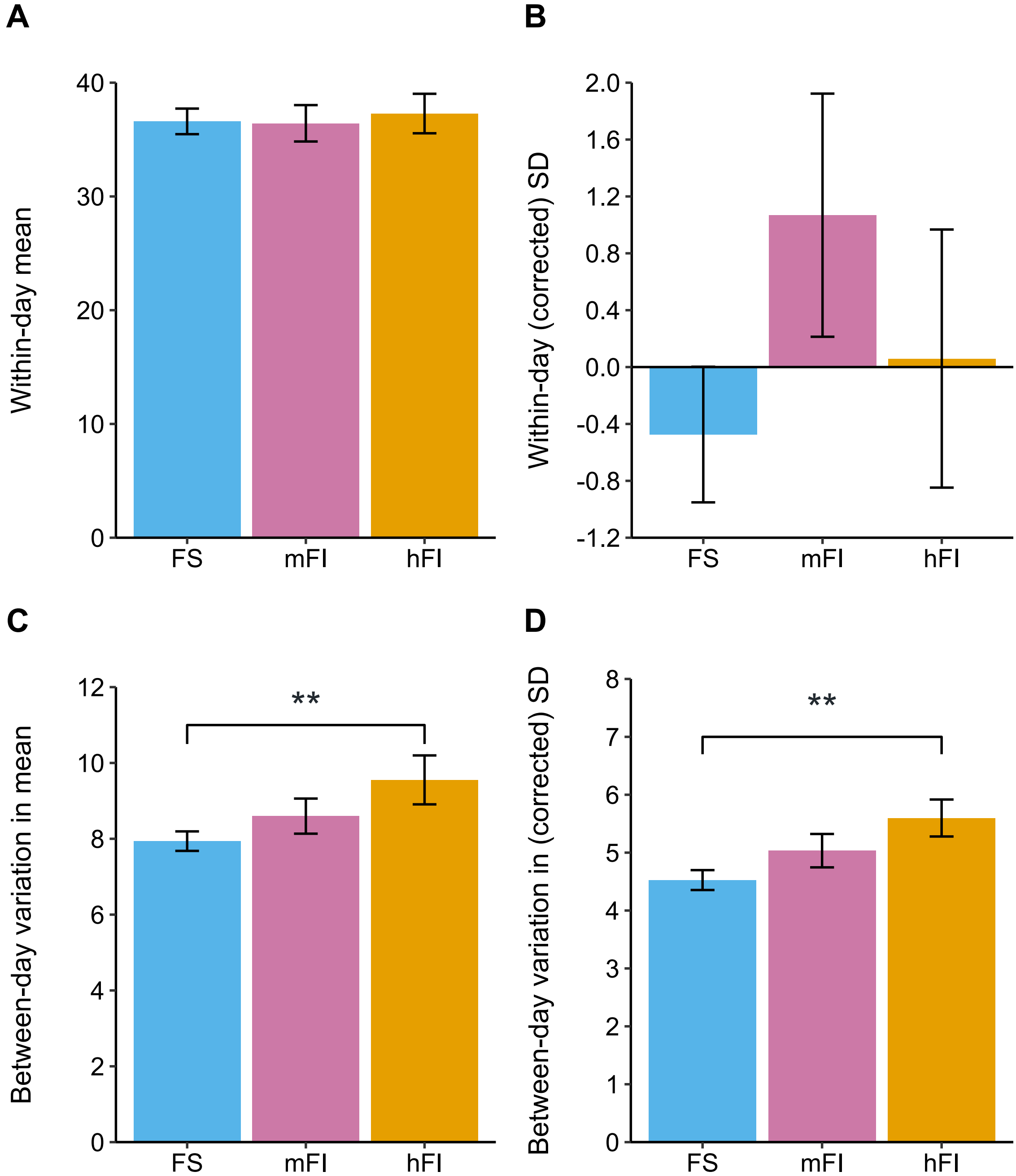
During the EMA period, we asked participants: “How stressed do you feel right now?”, “Do you feel that you can cope with things right now?” and “Do you feel that you’re on top of things right now?”. They responded using a VAS anchored at 0 (“Not at all”) and 100 (“Very much”). These questions constitute the Perceived Stress Scale adapted for momentary use by Reichenberger et al. (2020).

A.3 Results from Principal Component Analysis of Daily Momentary Hunger Scale Summary Statistics

In our pilot dataset, we derived several statistics summarising daily hunger from our dataset and found strong correlations between variables. Therefore, we pre-registered to repeat this analysis with our full dataset to ensure we use appropriate hunger measures.

To do this, we ran a principal component analysis with seven daily hunger summary statistics: mean, median, area under curve (AUC), standard deviation, range, minimum and maximum. This resulted in two principal components (PC). Daily average hunger (PC1) was strongly loaded by the mean (0.96), median (0.94), and AUC (0.95). Variation in daily hunger (PC2) was strongly loaded by the SD (0.97) and the range (0.99). Because of these strong associations, we opted to use daily mean hunger rating and daily SD of hunger rating instead of PC1 and PC2, respectively, as these were the most intuitive options for ease of interpreting results.

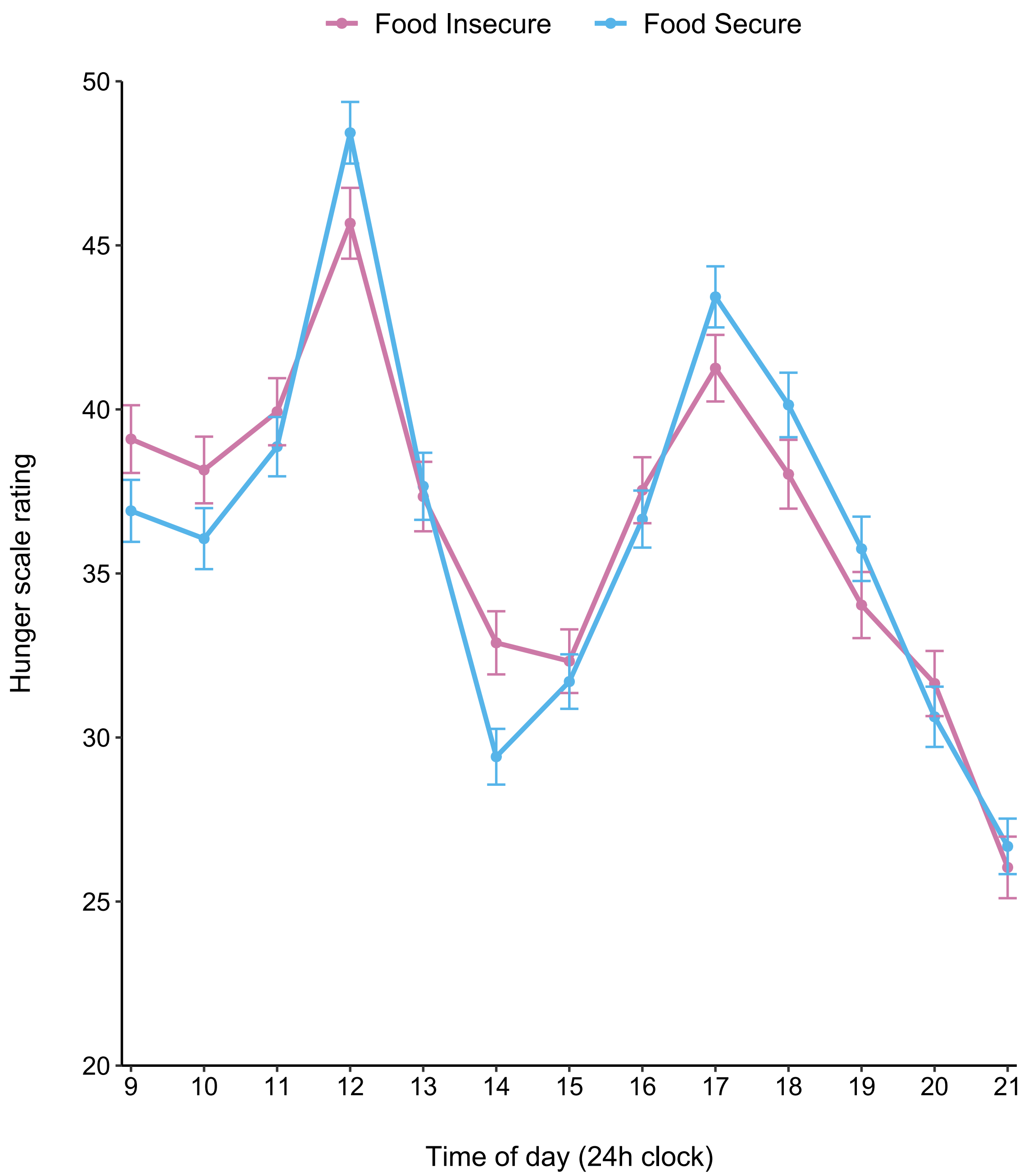
A.4 Within- and Between-Day Hunger in High and Moderate FI and FS



Within-day (A, B) and between-day (C, D) measures of hunger in high FI (hFI), moderate FI (mFI) and FS groups.

Note. Error bars indicate one standard error of the mean. **significance of *p* <.01.

A.5 Mean Hunger in FI and FS Groups At Each Assessment



Mean hunger scale rating in FI and FS groups at each momentary assessment.

Note. Error bars indicate one standard error of the mean for momentary assessments. N = 22,179.

B Chapter 3 Supplementary Materials

B.1 Dietary Restraint and Accuracy in Lag2 Food Trials in the Hungry Condition

Relationship between dietary restraint and accuracy in lag2 trials with food distractor images in the hungry condition (E1, E2, PPZ)***.***

	Bayesian analyses		*Frequentist analyses*	
PPZ	**E1**	**E2**	**E1**	**E2**
Mean dietary restraint scale score (SD) and classification				
Medium restraint 12.9 (4.9)	**High restraint** 24.0 (8.5)	**High restraint** 24.2 (7.8)	-	-
Linear regression output				
Not significant *p* = .192 *r* = -.29	**Evidence for null** BF = 0.029 median = 0.08, 89% CI [-0.48, 0.66], pd = 62.18%	**Evidence for null** BF = 0.032 median = 0.11, 89% CI [-0.50, 0.69], pd = 64.95%	**Not significant** *p* = .76 *r* = .08	**Not significant** *p* = .69 *r* = .12

Note. The dependent variable is the score on the dietary restraint scale as used in Herman and Polivy (1975) and Herman *et al*. (1978).

B.2 Structure of Models In Chapter 3

Structure of models used to analyse practice effect, all trials, lag2 trials and lag8 trials.

Analysis	**Model**
Practice effect	accuracy (all trials) ~ session + (1|participant)
All trials	accuracy (all trials) ~ lag*state*category*sequence + (1|participant)
Lag2 trials	accuracy (lag2 trials) ~ state*category*sequence + (1|participant)
Lag8 trials	accuracy (lag8 trials) ~ state*category*sequence + (1|participant)

B.3 Frequentist Replication of Results

Frequentist replication of results presented in the main text (E1, E2, PPZ).

PPZ	**E1**	**E2**
*Hunger manipulation check**		
Significant $p < .0005$ $t(22) = 10.9$	**Significant** $p < .0001$ $t(28) = 26.18$	**Significant** $p < .0001$ $t(27) = 17.73$
*Blink of attention check***		
Significant $p < .0005$ $F(1, 21) = 130.1$	**Significant** $p < .0001$ $F(1, 27) = 23.61$	**Significant** $p < .0001$ $F(1, 26) = 36.94$
*Practice effect check***		
Significant $p = .001$ $F(1, 22) = 15.3$	**Significant** $p < .01$ $F(1, 28) = 8.18$	**Significant** $p < .01$ $F(1, 27) = 12.27$
*Replication of main findings (P1)***		
Significant $p = .03$ $F(2, 21) = 3.8$	Not **significant** $p = .81$ $F(2, 54) = 0.21$	**Not significant** $p = .88$ $F(2, 52) = 0.13$
*Replication of main findings (P2)***		
Not **significant** *Not reported*	**Not significant** $p = .06$ $F(2, 54) = 3.01$	**Not significant** $p = .53$ $F(2, 52) = 0.65$

Note. The dependent variables are *hunger rating (0-7) and **accuracy (%).

B.4 Accuracy Differences Between Lags for Each Distractor Category

All trials: Bayesian and frequentist t-tests assessing accuracy differences between lags for each distractor category (E1, E2, PPZ).

Original study	***Bayesian analyses***		***Frequentist analyses***	
PPZ	**E1**	**E2**	**E1**	**E2**
Neutral lag2 vs. neutral lag8				
Significant All *p*s < .0005 All *t*s > 4.5	**Evidence for effect** BF > 100 difference: median = 6.88, 89% CI [4.23, 9.60], pd = 99.98%	**Evidence for effect** BF = 22.00 difference: median = 6.08, 89% CI [2.88, 8.91], pd = 99.88%	**Significant** $p < .001$ $t(28) = -4.47$	**Significant** $p < .01$ $t(27) = -3.50$
Romantic lag2 vs. romantic lag8				
Significant As above	**Evidence for effect** BF > 100 difference: median = 7.52, 89% CI [4.39, 10.19], pd = 100%	**Evidence for effect** BF > 100 difference: median = 6.60, 89% CI [4.27, 8.95], pd = 100%	**Significant** $p < .0001$ $t(28) = -4.55$	**Not significant** $p < .0001$ $t(27) = -4.82$
Food lag2 vs. food lag8				
Significant As above	**Evidence for effect** BF = 14.80 difference: median = 5.11, 89% CI [2.41, 7.98], pd = 99.78%	**Evidence for effect** BF > 100 difference: median = 7.64, 89% CI [4.78, 10.37], pd = 100%	**Significant** $p < .01$ $t(28) = -3.32$	**Significant** $p < .0001$ $t(27) = -4.76$

Note. The dependent variable is accuracy (%).

B.5 Lag2 Trials: Effects of Category, State, and State-Sequence and Category-State-Sequence Interactions

Lag2 trials: Bayesian and frequentist analyses of the main effects of category and state, and state-sequence and category-state-sequence interaction effects, on accuracy (E1, E2, PPZ).

Original study	***Bayesian analyses***		***Frequentist analyses***	
PPZ	**E1**	**E2**	**E1**	**E2**
Main effect of category				
Significant *p* = .001 *F*(1, 21) = 11.4	**Evidence for effect** BF = 66.77 <u>romantic</u>: median = 4.91, 89% CI [2.81, 6.98], pd = 99.98% <u>food</u>: median = 4.47, 89% CI [2.32, 6.53], pd = 99.92%	**Evidence for effect** BF = 37.76 <u>romantic</u>: median = 6.14, 89% CI [3.52, 8.60], pd = 99.98% <u>food</u>: median = 3.50, 89% CI [0.83, 5.87], pd = 98.35%	**Significant** *p* < .001 *F*(2, 54) = 9.56	**Significant** *p* < .01 *F*(2, 52) = 5.38
Main effect of state				
Significant *p* = .009 *F*(1, 21) = 8.3	**Inconclusive** BF = 2.52 <u>sated</u>: median = 2.70, 89% CI [0.93, 4.37], pd = 99.08%	**Inconclusive** BF = 0.16 <u>sated</u>: median = - 0.52, 89% CI [-2.72, 1.55], pd = 65.72%	**Significant** *p* = .02 *F*(1, 27) = 5.96	**Not significant** *p* = .39 *F*(1, 26) = 0.77
*State-sequence interaction**				
Significant *p* < .0005 *F*(1, 21) = 31.2	**Inconclusive** BF = 1.45 <u>satedfirst:sated</u>: median = -4.02, 89% CI [-7.09, -0.47], pd = 97.32%	**Inconclusive** BF = 4.89 <u>satedfirst:sated</u>: median = -6.18, 89% CI [-10.35, -2.24], pd = 99.22%	**Not significant** *p* = .07 *F*(1, 27) = 3.52	**Significant** *p* < .01 *F*(1, 26) = 8.32
Category-state-sequence interaction				
Significant *p* = .011 *F*(2, 21) = 5.0	**Inconclusive** BF = 0.81 <u>satedfirst:sated:romantic</u>: median = 6.77, 89% CI [0.40, 13.69], pd = 94.35% <u>satedfirst:sated:food</u>: median = -0.18, 89% CI [-6.77, 6.47], pd = 51.92%	**Inconclusive** BF = 0.24 <u>satedfirst:sated:romantic</u>: median = -6.18, 89% CI [-10.35, -2.24], pd = 99.22% <u>satedfirst:sated:food</u>: median = -0.08, 89% CI [-7.60, 7.81], pd = 50.68%	**Not significant** *p* = .11 *F*(2, 54) = 2.29	**Not significant** *p* = .93 *F*(2, 52) = 0.07

Note. The dependent variable is accuracy (%). *"Sequence" describes whether a participant did their hungry or sated session first. PPZ suggested that the presence and extent of a practice effect between sessions was dependent on which session a participant completed first.

B.6 Lag2 Trials: Accuracy Differences Between Distracter Categories

Lag2 trials: Bayesian and frequentist t-tests assessing accuracy differences between distractor categories (E1, E2, PPZ).

Original study	***Bayesian analyses***		***Frequentist analyses***	
PPZ	**E1**	**E2**	**E1**	**E2**
*Neutral vs. romantic**				
Significant $p < .0005$ $t(22) = 4.5$	**Evidence for effect** BF > 100 difference: median = 4.77, 89% CI [1.78, 6.48], pd = 99.98%	**Evidence for effect** BF = 35.72 difference: median = 5.86, 89% CI [2.97, 8.42], pd = 99.90%	**Significant** $p < .0001$ $t(28) = -4.59$	**Significant** $p < .0001$ $t(27) = -3.71$
Neutral vs. food				
Not reported	**Inconclusive** BF = 9.27 difference: median = 4.29, 89% CI [1.78, 6.48], pd = 99.80%	**Inconclusive** BF = 0.62 difference: median = 3.25, 89% CI [0.05,6.96], pd = 94.30%	**Significant** $p < .01$ $t(28) = -3.10$	**Not significant** $p = .123$ $t(27) = -1.59$
Romantic vs. food				
Significant $p = .009$ $t(22) = 3.9$	**Inconclusive** BF = 0.21 difference: median = 0.42, 89% CI [-1.82, 2.25], pd = 63.18 %	**Inconclusive** BF = 0.70 difference: median = 2.36, 89% CI [0.17, 5.12], pd = 94.10%	**Not significant** $p = .745$ $t(28) = -0.33$	**Not significant** $p = .103$ $t(27) = -1.67$

Note. The dependent variable is accuracy (%). *In PPZ, accuracy was greater in lag2 trials with neutral distractors than with romantic distractors. In E1 and E2, this effect was in the opposite direction.

B.7 Lag8 Trials: Effects of Category, State, and State-Sequence and Category-State-Sequence Interactions

Lag8 trials: Bayesian and frequentist analyses of main effects of category and state, and a state-sequence interaction effect, on accuracy (E1, E2, PPZ).

Original study	***Bayesian analyses***		***Frequentist analyses***	
PPZ	**E1**	**E2**	**E1**	**E2**
Main effect of category				
Significant $p < .0005$ $F(2, 21) = 34.7$	**Evidence for effect** BF > 1000 romantic: median = 5.61, 89% CI [3.73, 7.36], pd = 100% food: median = 2.76, 89% CI [1.06, 4.55], pd = 99.12%	**Evidence for effect** BF > 1000 romantic: median = 6.70, 89% CI [5.01, 8.34], pd = 100% food: median = 5.13, 89% CI [3.49, 6.84], pd = 100%	**Significant** $p < .0001$ $F(2, 54) = 14.40$	**Significant** $p < .0001$ $F(2, 52) = 24.13$
Main effect of state				
Not significant	**Inconclusive** BF = 0.13 sated: median = 0.80, 89% CI [-0.70, 2.31], pd = 80.08%	**Inconclusive** BF = 0.17 sated: median = 1.01, 89% CI [-0.50, 2.52], pd = 85.85%	**Not significant** $p = .41$ $F(1, 27) = 0.71$	**Not significant** $p = .55$ $F(1, 26) = 0.38$
*State-sequence interaction**				
Significant $p = .026$ $F(1, 21) = 5.7$	**Evidence for effect** BF = 58.10 satedfirst:sated: median = -6.10, 89% CI [-9.02, -3.34], pd = 99.92%	**Inconclusive** BF = 7.33 satedfirst:sated: median = -4.81, 89% CI [-7.50, -1.89], pd = 99.60%	**Significant** $p < .01$ $F(1, 27) = 8.67$	**Significant** $p < .05$ $F(1, 26) = 5.73$

Note. The dependent variable is accuracy (%). *“Sequence” describes whether a participant did their hungry or sated session first. PPZ suggested that the presence and extent of a practice effect between sessions was dependent on which session a participant completed first.

B.8 Lag8 Trials: Accuracy differences between distracter categories

Lag8 trials: Bayesian and frequentist t-tests assessing accuracy differences between distractor categories (E1, E2, PPZ).

Original study	***Bayesian analyses***		***Frequentist analyses***	
PPZ	**E1**	**E2**	**E1**	**E2**
*Neutral vs. romantic**				
Significant *p* < .0005 *t*s not reported	**Evidence for effect** BF > 100 <u>difference</u>: median = 5.39, 89% CI [3.59, 7.49], pd = 99.98%	**Evidence for effect** BF > 1000 <u>difference</u>: median = 6.58, 89% CI [5.04, 8.22], pd = 100%	**Significant** $p < .0001$ $t(28) = -4.73$	**Significant** $p < .0001$ $t(27) = -7.31$
Neutral vs. food				
Not reported Described as being 'at an equivalent level'.	**Inconclusive** BF = 3.57 <u>difference</u>: median = 2.25, 89% CI [0.81, 4.21], pd = 99.05%	**Evidence for effect** BF > 100 <u>difference</u>: median = 4.29, 89% CI [2.88,6.82], pd = 100%	**Significant** $p < .05$ $t(28) = -2.64$	**Significant** $p < .001$ $t(27) = -4.37$
Romantic vs. food				
Significant *p* < .0005 *t*s not reported	**Evidence for effect** BF = 15 <u>difference</u>: median = 2.64, 89% CI [1.39, 4.13], pd = 99.75%	**Inconclusive** BF = 0.70 <u>difference</u>: median = 1.43, 89% CI [0.03, 2.92], pd = 94.12%	**Significant** $p < .01$ $t(28) = -3.34$	**Not significant** $p = .103$ $t(27) = -1.69$

Note. The dependent variable is accuracy (%). *In PPZ, accuracy was greater in lag8 trials with neutral distractors than with romantic distractors. In E1 and E2, this effect was in the same direction.

B.9 All Trials: Lag-Category Interaction Effect on Accuracy

All trials: Bayesian and frequentist analyses of the lag-category interaction effect on accuracy (E1, E2, PPZ).

Original study	***Bayesian analyses***		***Frequentist analyses***	
PPZ	**E1**	**E2**	**E1**	**E2**
Lag-category interaction				
Significant *p* < .0005 F(2, 21) = 38.1	**Evidence for null** BF = 0.07 lag8:romantic: median = 0.95, 89% CI [-1.98, 4.06], pd = 68.92% lag8:food: median = -1.48, 89% CI [-4.70, 1.40], pd = 77.78%	**Evidence for null** BF = 0.05 lag8:romantic: median = 0.70, 89% CI [-2.45, 3.80], pd = 63.55% lag8:food: median = 1.77, 89% CI [-1.45, 4.93],pd = 80.25%	**Not significant** *p* = .29 F(2, 54) = 1.27	**Not significant** *p* = .71 F(2, 52) = 0.35

Note. The dependent variable is accuracy (%).

B.10 Lag2 Trials: Effect of Food Type and Food Type-State Interaction on Accuracy

Lag2 trials: Bayesian and frequentist analyses of the main effect of food type (sweet or savoury) and the interaction effect of food type-state on accuracy (E1, E2).

Bayesian analyses		**Frequentist analyses**	
E1	**E2**	**E1**	**E2**
Main effect of food type			
Evidence for null BF = 0.10 sweet: median = -0.06, 89% CI [-0.78, 0.65], pd = 55.50%	**Inconclusive** BF = 0.24 sweet: median = 0.45, 89% CI [-0.76, 0.72], pd = 50.35%	**Not significant** *p* = .43 F(1, 28) = 0.64	**Not significant** *p* = .94 F(1, 28) = 0.01
Interaction effect food type:state			
Evidence for null BF = 0.01 sweet:sated: median = 0.13, 89% CI [-1.02, 1.00], pd = 50.15%	**Evidence for null** BF = 0.02 sweet:sated: median = -0.05, 89% CI [-1.09, 1.02], pd = 53.10%	**Not significant** *p* = .58 F(1, 28) = 0.31	**Not significant** *p* = .22 F(1, 28) = 1.60

Note. The dependent variable is accuracy (%).

B.11 Lag2 Trials: Effect of Gender and of Gender-Category-State Interaction on Accuracy

Lag2 trials: Bayesian and frequentist analyses of the main effect of gender and the interaction effect of gender-category-state on accuracy (E1, E2).

Bayesian analyses		***Frequentist analyses***	
E1	**E2**	**E1**	**E2**
Main effect of gender			
Inconclusive BF = 0.97 women: median = 0.61, 89% CI [-1.60, 2.94], pd = 66.80%	**Inconclusive** BF = 0.96 women: median = 0.66, 89% CI [-1.74, 3.00], pd = 67.45%	**Not significant** $p = .52$ $F(1, 25) = 0.42$	**Not significant** $p = .20$ $F(1, 23) = 1.78$
Interaction effect of gender:category:state			
Inconclusive BF = 0.97 women:romantic:sated: median = -0.39, 89% CI [-3.05, 2.22], pd = 59.67% women:food:sated: median = 0.29, 89% CI [-2.36, 2.95], pd = 56.95%	**Inconclusive** BF = 0.84 women:romantic:sated: median = -2.07, 89% CI [-5.19, 1.16], pd = 84.88% women:food:sated: median = -1.55, 89% CI [-4.71. 1.63], pd = 77.78%	**Not significant** $p = .57$ $F(2, 50) = 0.57$	**Not significant** $p = .73$ $F(2, 46) = 0.47$

Note. The dependent variable is accuracy (%). The data of one non-binary participant has been dropped from the E2 analyses presented here for simplicity in interpreting results.

B.12 Lag2 Trials: Effect of Age and Age-Category-State Interaction on Accuracy

Lag2 trials: Bayesian and frequentist analyses of the main effect of age and the interaction effect of age-category-state on accuracy (E1, E2).

Bayesian analyses		***Frequentist analyses***	
E1	**E2**	**E1**	**E2**
Main effect of age			
Inconclusive BF = 0.86 age: median = 0.22, 89% CI [0.00, 0.45], pd = 94.58%	**Inconclusive** BF = 0.87 age: median = -0.02, 89% CI [-0.08, 0.04], pd = 69.60%	**Not significant** *p* = .10 F(1, 25) = 2.86	**Significant** *p* = .01 F(1, 24) = 7.91
Interaction effect of age:category:state			
Inconclusive BF = 1.35 age:romantic:sated: median = -0.004, 89% CI [-0.26, 0.23], pd = 50.98% age:food:sated: median = 0.05, 89% CI [-0.19, 0.29], pd = 63.48%	**Inconclusive** BF = 0.76 age:romantic:sated: median = 0.12, 89% CI [0.03, 0.20], pd = 98.68% age:food:sated: median = 0.12, 89% CI [0.03, 0.20], pd = 98.45%	**Not significant** *p* = .77 F(2, 50) = 0.27	**Significant*** *p* = .02 F(2, 48) = 4.14

Note. The dependent variable is accuracy (%). *Frequentist t-test results showed no difference in performance in hungry food trials compared to sated food trials in younger participants (*t*(27.88) = -0.06, *p* = .95, 95% CI [-2.18, 2.05]) or older participants (*t*(23.36) = 0.27, *p* = .79, 95% CI [-1.68, 2.18]).

C Chapter 4 Supplementary Materials

C.1 S1 Methods

Participants

We calculated that 85 participants would be required to achieve 80% power, with a medium effect size (d = 0.5) and an alpha level of 0.05. We used a medium effect size as we could not find a comparable study with published effect sizes. Consequently, we recruited 100 participants (women = 56, men = 44; ages 18 – 59 years, M = 33.1, SD = 10.4) using Prolific (an online participant recruitment platform) to ensure a minimum sample size of 85 would be achieved even after removing participants who reported being interrupted during the task. All participants were currently residing in the UK.

To achieve a wide range of self-reported momentary hunger ratings across our sample, we recruited participants in two waves on the same day (11:10 and 14:00 on 10th August 2020). The waves were timed to be either side of a typical British lunchtime (midday) to capture individuals who had not yet eaten or had recently finished lunch. We removed four participants who indicated they had been interrupted during the study and five whose Qualtrics data indicated they had seen some stimuli longer than they should have during encoding.

Procedure

We advertised our study on Prolific and informed participants that they should be over 18 years old, live in the UK, and have access to the internet on a tablet or a computer to participate. We instructed them that they would complete a memory task and be asked for additional information about themselves. Participants recruited in both waves completed the same task and questions.

After we recruited participants on Prolific, we automatically redirected them to Qualtrics. Participants then provided their informed consent, age, and gender before beginning the image recognition task (see “Image recognition task”). This was immediately followed by the price recall task (see “Price recall task”). Before starting the tasks, we instructed participants to ensure they were in a quiet place with no distractions, give their full attention to the screen, and not write anything down.

Upon completing the tasks, we asked participants if they had been interrupted during the tasks, when they had last eaten food, and how hungry they felt (see “Self-reported hunger rating”). After that, we presented debrief materials to participants on Qualtrics and their payments were approved and received through Prolific. Participants received 1.45 GBP for taking part in the study. This was based on the study taking approximately ten minutes to complete and the 2020/21 UK minimum wage.

Measures

Image recognition task The task started with a serial visual presentation of 15 images of food items and 15 images of non-food items, with a price above each item (see Figure C1). Each image was shown for three seconds, and the order was randomised. We instructed participants to pay close attention and remember what they saw.

Once the presentation was completed, we presented participants with a grid containing 60 words with checkboxes. Of these words, 30 were the names of items shown in the presentation (15 food and 15 non-food), and 30 were items not shown (15 food items and 15 non-food items). We instructed participants to select the 30 items they had been shown and to select 30 items exactly, even if they could not remember all 30 items they had been shown.

Price recall task Following the image recognition task, we presented participants with a list of all images from the serial visual presentation above. Participants tried to recall the price of each image and entered it into a textbox next to the image. The order of the images was randomised between participants, and there were no time restrictions.

Self-reported hunger rating Participants answered the question “*How hungry do you feel?*” using a visual analogue scale (VAS) anchored at 0 (“*not feeling hungry at all*”), 5 (“*feeling moderately hungry*”), and 10 (“*feeling extremely hungry*”).

Figure C1. An example of the serial visual presentation of images and prices in S1.

Note. An image of a food or non-food item was shown below a price, which was always the same font and size.

C.2 S2 Methods

Participants

As previously mentioned, we deviated from our pre-registered analysis plan for S2. However, we retained our pre-registered flexible Bayesian stopping rule for recruitment instead of an a priori power analysis to determine sample size. We required a minimum sample size of 50 and a Bayes factor of < 1/10 or > 10 for several pre-registered predictions. After 50 participants, these criteria were not met, and we continued to test participants. We stopped data collection after our next check as these criteria were met. We recruited 103 participants (women = 71, men = 31, non-binary = 1; ages 18 – 67 years, M = 29.6, SD = 11.7) from undergraduate populations, university-maintained participant mailing lists, and social media. Because of the inclusion of gender as a covariate in our models, we removed data from the single participant who identified as non-binary so that gender categories were balanced in sample size.

Procedure

Recruitment and data collection occurred between 6th December 2021 and 24th February 2022. Before signing up, we informed participants that they should be over 18 years old, have normal or corrected-to-normal vision, and not have a medical condition that would be exacerbated by fasting. We also told them that they would need to be able to attend Newcastle University for one 30-minute session on a weekday between 2 p.m. and 5 p.m.

After signing up and providing consent, we allocated participants to a hungry or a sated condition on an alternate sign-up basis. We instructed all participants to eat breakfast before 9 a.m. on the day of their participation to ensure that they were awake at similar times to minimise the potential impacts of circadian rhythm on cognition.

We asked participants in the sated condition to eat and drink as usual and to eat in the hour preceding their study session. We asked participants in the hungry condition to refrain from eating after 9 a.m. until after they had completed their session. We informed participants that they could consume water, tea, and coffee, but instructed them to avoid calorie-dense drinks or drinks containing a lot of sugar or milk. At 6 p.m. the evening before the study, we sent participants an email reminding them of these instructions.

When participants arrived on site, we asked whether they had complied with the fasting or eating requirements and reminded them what the study would entail. We then showed the participant to a quiet laboratory with a PC readied with the study.

Before beginning their first task on the PC, all participants provided their informed consent, age and gender. Upon completing the tasks, we asked participants when they had last eaten food, how hungry they felt (see “Self-reported hunger rating”), and to confirm whether they had fasted between 9 a.m. and their session. Finally, we presented debrief materials to participants on-screen. They were then paid in vouchers (or with course credits if they were psychology undergraduates) immediately after the session ended.

The task order was counterbalanced: half of the study population completed an image recognition task (see “Image recognition task”), followed by a one-minute break, and then a price recall task (see “Price recall task”). The remaining half completed the price recall task, followed by a one-minute break, and then the image recognition task. The counterbalancing of the task sequence was distributed evenly between the two conditions.

Measures

Image recognition task The task started with a serial visual presentation of 30 images of food items and 30 images of non-food items (see Figure C2). Each image was shown immediately after the other for one second, and the order was randomised. We instructed participants to pay close attention and remember what they saw.

Immediately after the image sequence, participants completed a simple reaction time task on the PC and a verbal forward-digit span task with the researcher (see Appendix C.10for protocol). We included these as distractor tasks to increase the difficulty of the image recognition task by engaging participants in a different activity between the encoding and response sections of the image recognition task. These were timed to ensure all participants experienced a similar delay.

Following the distractor tasks, participants completed the response section of the image recognition task. Participants were shown a sequence of 120 images: 30 food and 30 non-food images, which were "old", and 30 food images and 30 non-food images that were "new". The order of the images was randomised. Each image appeared for five seconds. During those five seconds, participants indicated whether they had seen the image before using the "y" (yes, old) and "n" (no, new) keys.

Price recall task The task started with a serial visual presentation of 15 images of food items and 15 images of non-food items alongside a price in GBP (Figure C2). Each image was shown for four seconds (as in Morris and Dolan (2001)), and the order of images was randomised. We instructed participants to remember as many prices as possible. There was no overlap in image use between the image recognition and price recall tasks.

We rounded prices to the nearest 0.10 GBP, which was not a whole pound (e.g., 3 GBP) or 0.50 GBP (e.g., 3.50 GBP). For example, £2.03 would be rounded up to £2.10, but £2.13 would be rounded down to £2.10. We did this to reduce variability in the memorability of prices, as in S1, we found evidence that whole number prices were more memorable. Furthermore, all items in the same category (food or non-food) were priced differently from one another to minimise potential advantages for item prices through association. To achieve this, we occasionally had to round prices up by 0.20 GBP rather than 0.10 GBP. We selected items and prices across the two categories so that the mean price and variation in price were as similar as possible between food (M = 1.38 GBP, SD = 1.01 GBP) and non-food items (M = 1.32 GBP, SD = 1.03 GBP).

After seeing the sequence of images and prices, participants had a two-minute break and a countdown timer was shown onscreen for the duration. Once this was complete, we presented the 30 images individually without their prices. Participants had to recall the price from memory and type it into the textbox below each image (see Figure C2). Each image was shown for 10 seconds (a countdown timer was shown beneath the textbox), and participants had to type their responses within that time. We limited the time, so participants progressed through the task at a similar rate.

Self-reported hunger rating We asked all participants how hungry they felt on a visual analogue scale anchored at 0 (not at all hungry) and 7 (extremely hungry), as previously used in Piech et al. (2010).

Figure C2. An example of how images were presented during the (A) serial visual presentation of the image recognition task, (B) serial visual presentation of the price recall task, and (C) response section of the price recall task in S2.

***Note*.** (**A**) This is also how stimuli were presented in the response section of the task. (**B**) The image was always shown below a price with the same font and size.

C.3 Rationale for Deviating From S2 Pre-registration

We deviated from our pre-registered analysis for S2 as the pre-registered plans were not ideally suited to our data. There were two main differences.

1. For the image recognition tasks, we pre-registered that we would use the number of hits plus correct rejections for our measure of task accuracy and the number of hits plus false alarms for our measure of bias. Instead, we used the SDT measures d' and c for the reasons described in the main text.

2. We changed the structure of our models. Our outcome variables were based only on food trials in the pre-registered analysis. We then planned to include the equivalent outcome variables for all trials as a predictor in the model to control for individual differences in performance. However, we recognised that this would result in an overlap of the same data being used as outcomes and predictors in the same model. Thus, we moved away from the approach to the model structures described in the main text.

C.4 Descriptions and Interpretation of the Bayesian Indices pd and ROPE

Descriptions of the Bayesian indices pd and ROPE and how to interpret their values, as described by Makowski et al. (2019a) and Makowski et al. (2019b).

Parameter	Description	Interpretation
pd	The probability of direction (pd) is "…an index of **effect existence**, ranging from 50% to 100%, representing the certainty with which an effect goes in a particular direction (i.e., is positive or negative)" (Makowski et al. 2019a; Makowski et al. 2019b). The pd often has a direct association with the frequentist *p* value.	The higher the value, the greater the evidence for the effect's existence. E.g. a two-tailed *p* value of .05 would correspond to a pd value of 97.5%.
ROPE	ROPE-based indices describe whether an effect is negligible or too small to be of any practical equivalence; they speak to **effect significance** regarding whether the effect is large enough to be notable in the real world. The full region of practical equivalence (ROPE) is the proportion of the posterior distribution that lies within the ROPE. For linear models, Kruschke (2014) suggested that ROPE be set as a range from [-0.1*SD_y, 0.1*SD_y] and, for logistic models, [-0.18*SD_y, 0.18*SD_y], where SD_y is the standard deviation of the effect of interest. These values are used as 0.1 and 0.18 are negligible effect sizes for linear and logistic models, respectively, as determined by Cohen (2013).	The lower the value, the more likely the effect is to have practical significance. E.g. for the full ROPE (100%), the null hypothesis is rejected if the percentage of the posterior distribution within the ROPE is less than 2.5% or accepted if it is greater than 97.5%.

C.5 Effects of Age and Gender on d', Food Response Bias (S1), and c (S2)

Bayesian linear mixed model results of effects of age and gender on d' and food response bias (S1)/c (S2) in the recognition task.

		Parameter	**Median diff.**	**89% CI**		**pd** (%)	**ROPE** (%)
				LL	*UL*		
d'	**S1**	Age	0.00	-0.01	0.02	68.1	100.0
		Gender (*Women*)	0.19	-0.09	0.48	86.0	21.9
	S2	Age	0.00	-0.01	0.01	61.3	100.0
		Gender (*Women*)	0.14	-0.07	0.36	84.9	22.4
Bias (*food*)	**S1**	Age	0.00	-0.03	0.02	60.3	100.0
		Gender (*Women*)	0.02	-0.41	0.46	52.8	36.0
c	**S2**	Age	-0.01	-0.01	0.00	98.1	100.0
		Gender (*Women*)	-0.07	-0.18	0.05	83.3	25.9

Note. *N* = 91 (S1) and *N* = 100 (S2). Median diff. = median difference; CI = credible interval; *LL* = lower limit; *UL* = upper limit; pd = probability of direction; ROPE = region of practical equivalence.

C.6 Effects of Age, Gender and Hunger – When Considering Calorie Content – on Hit Rate and Correct Food Price Recall

Bayesian logistic regression results of effects of age and gender and hunger rating(S1)/state (S2) – when considering calorie content – on whether a participant responded correctly in recognition task trials with “old” food images (hit rate) and whether a participant correctly recalled a food price in the price recall task.

		Parameter	**Median** (odds ratio)	**89% CI** *LL*	*UL*	**pd** (%)	**ROPE** (%)
Hit (yes or no)	**S1**	Age	1.01	0.98	1.04	61.7	100.0
		Gender (*Women*)	1.49	0.82	2.75	85.6	22.7
		Hunger rating	0.92	0.81	1.04	85.1	90.1
	S2	Age	1.01	0.99	1.02	84.5	100.0
		Gender (*Women*)	1.36	0.97	1.91	93.1	26.0
		State (*Sated*)	1.34	0.98	1.85	93.7	26.9
Food price recall (correct or not)	**S1**	Age	1.03	1.01	1.05	97.4	100.0
		Gender (*Women*)	0.88	0.56	1.35	68.7	43.8
		Hunger rating	1.06	0.97	1.15	84.3	99.0
	S2	Age	0.99	0.98	1.01	76.2	100.0
		Gender (*Women*)	0.77	0.57	1.05	91.2	31.7
		State (*Sated*)	1.12	0.83	1.48	73.4	59.1

Note. *N* = 91 (S1) and *N* = 100 (S2). Median diff. = median difference; CI = credible interval; *LL* = lower limit; *UL* = upper limit; pd = probability of direction; ROPE = region of practical equivalence.

C.7 Effects of the Hunger, Item Type, Age, and Gender on Price Recall Score and Accuracy

Bayesian linear mixed model results of the main effects of the hunger variable (hunger rating (S1) or state (S2)), item type, age, and gender, and a hunger variable X item type interaction effect on the score (the total number of prices remembered) and accuracy in the price recall task.

		Parameter	**Median diff.**	**89% CI**		**pd** (%)	**ROPE** (%)
				LL	*UL*		
No. of prices recalled	**S1**	Age	0.04	-0.01	0.09	92.1	100.0
		Gender (*Women*)	-0.55	-1.45	0.39	81.6	25.3
	S2	Age	0.00	-0.03	0.04	51.9	100.0
		Gender (*Women*)	-0.62	-1.53	0.25	87.6	21.7
Price recall accuracy	**S1**	Age	-0.01	-0.02	0.00	91.9	100.0
		Gender (*Women*)	0.06	-0.13	0.25	70.3	37.2
	S2	Age	0.00	-0.01	0.00	92.8	100.0
		Gender (*Women*)	0.08	-0.01	0.16	91.8	18.5

Note. *N* = 91 (S1) and *N* = 102 (S2). Median diff. = median difference; CI = credible interval; *LL* = lower limit; *UL* = upper limit; pd = probability of direction; ROPE = region of practical equivalence.

C.8 Effects of Age, Gender, and Item Type –When Considering "Whole Number" Pricing – on the Proportion of Prices Correctly Recalled in S1

Bayesian linear mixed model results of the effects of age, gender, and item type – when taking "whole number" pricing into consideration - on the proportion of prices correctly recalled by a participant in S1.

Parameter	**Median diff.**	**89% CI**		**pd** (%)	**ROPE**(%)
		LL	*UL*		
Age	1.00	1.00	1.01	86.6	100.0
Gender (*Women*)	0.99	0.92	1.06	59.8	49.3
Item type (*Non-Food*)	1.41	1.35	1.48	100.0	0.0

Note. *N* = 91. Median diff. = median difference; CI = credible interval; *LL* = lower limit; *UL* = upper limit; pd = probability of direction; ROPE = region of practical equivalence.

C.9 Effects of Age, Gender, and Hunger – When Considering Calorie Content - on Price Recall Accuracy

Bayesian linear mixed model results of the effects of age, gender, and a hunger variable (hunger rating (S1) or state (S2)) – when considering calorie content - on participants' price recall accuracy.

	Parameter	**Median diff.**	**89% CI**		**pd** (%)	**ROPE** (%)
			LL	*UL*		
S1	Age	-0.01	-0.01	0.00	98.7	100.0
	Gender (*Women*)	-0.01	-0.12	0.11	54.8	73.3
	Hunger rating	0.00	-0.03	0.02	62.8	100.0
S2	Age	0.00	0.00	0.01	76.5	100.0
	Gender (*Women*)	0.08	-0.01	0.17	93.1	50.7
	State (*Sated*)	-0.05	-0.12	0.04	81.9	76.5

Note. *N* = 91 (S1) and *N* = 102 (S2). Median diff. = median difference; CI = credible interval; *LL* = lower limit; *UL* = upper limit; pd = probability of direction; ROPE = region of practical equivalence.

C.10 Protocol for Distractor Tasks in the S2 Image Recognition Task

The simple reaction time task comprised eight practice trials and 20 real trials. Participants had to press the spacebar as quickly as possible when a black cross appeared in the white box that was continuously present on their black screen. We calculated individual baseline simple reaction times by averaging response times on the real trials. This was displayed to participants in milliseconds at the end of the reaction time task.

The verbal forward-digit span task immediately followed this. A three-minute timer appeared on the screen, alongside instructions for the participant to get the researcher’s attention. The researcher entered the room and explained what the participant needed to do in the task. When the participant was ready, the researcher read randomly generated sequences of numbers to the participant, and the participant repeated them back to the best of their ability. The first sequence had three digits. Each time the participant repeated the sequence correctly, the sequence length in the subsequent trial increased by one digit. This continued until the participant made a mistake. Then, the participant had one more opportunity to repeat a new sequence correctly at that sequence length. If they were unsuccessful again, the task would be over. The participant’s digit span was the maximum number of digits in a sequence they could correctly repeat.

Next, the researcher instructed the participant that a button to continue would appear once the three-minute timer had ended, and they should press it. This timer ensured that all participants experienced the same delay before the response section of the task.

THE END