Appetite 191 (2023) 107065

Contents lists available at ScienceDirect

Appetite

journal homepage: www.elsevier.com/locate/appet

No effect of hunger on attentional capture by food cues: Two replication studies

Courtney Neal [a], Gillian V. Pepper [b], Caroline Allen [c], Daniel Nettle [d,e,*]

[a] Population Health Sciences Institute, Newcastle University, Newcastle, UK
[b] Psychology Department, Northumbria University at Newcastle, Newcastle, UK
[c] School of Psychology, Newcastle University, UK
[d] Institut Jean Nicod, Département D'études Cognitives, École Normale Supérieure, Université PSL, EHESS, CNRS, Paris, France
[e] Department of Social Work, Education and Community Wellbeing, Northumbria University at Newcastle, Newcastle, UK

ARTICLE INFO

Keywords:
Food cues
Hunger
Attentional blink
Attentional bias
Replication

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, Pastorino, & Zald, 2010. Appetite, 54, 579–582]. 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.

1. 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), and 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

* Corresponding author. Department of Social Work, Education and Community Wellbeing, Coach Lane Campus, Northumbria University at Newcastle, Newcastle, NE7 7TR, UK.
E-mail address: daniel.nettle@northumbria.ac.uk (D. Nettle).

https://doi.org/10.1016/j.appet.2023.107065
Received 11 May 2023; Received in revised form 15 September 2023; Accepted 27 September 2023
Available online 28 September 2023
0195-6663/

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., 2022 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 6-h 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 Methods, section 2.1). 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.

2. Methods

2.1. 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 Results, section 3.1).

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 the Supplementary Materials.

2.2. 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 Bayes factor of $<1/10$ or >10 for the critical state by image category interaction in lag2 trials (see Methods, section 2.6) 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).

2.3. 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.

2.4. 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 6 h 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., 2007) 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 6 h 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., 2022). 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 min. 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.

2.5. Measures

2.5.1. 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 (Fig. 1). The task consisted of one block of 16 practice trials and six blocks of 32 real trials. There were 1-min 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 5 s 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 5 s. 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 they 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.

2.5.2. 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

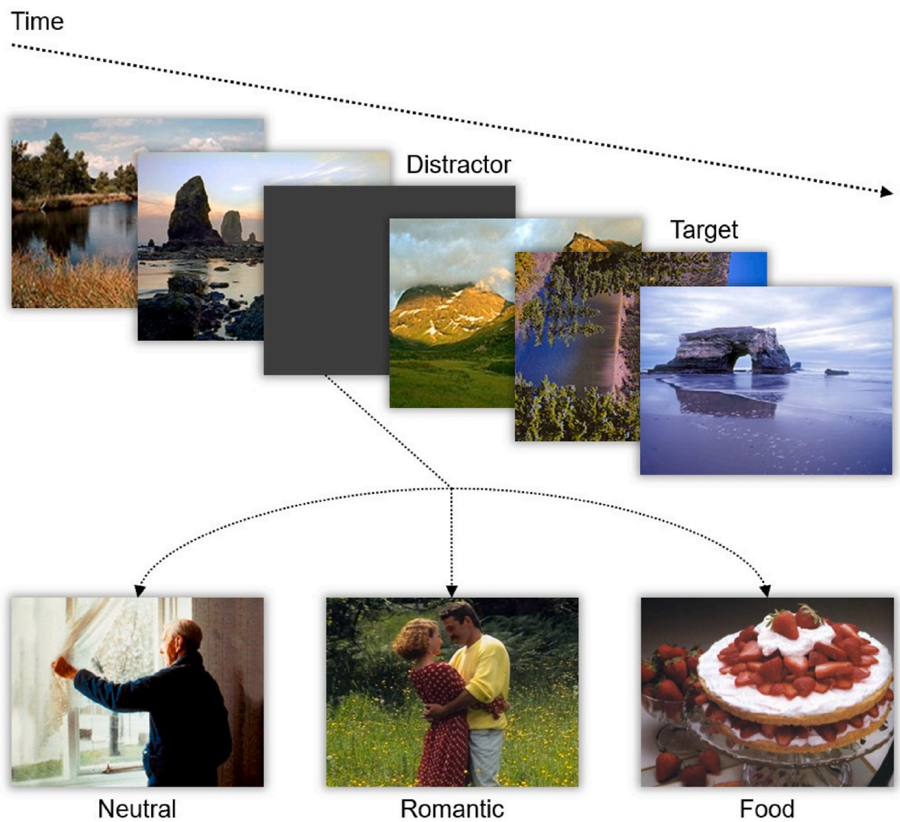

Fig. 1. Representation of part of a single EBA task trial.
Note. In half of the trials with targets, the distractor image was shown eight places before the target (lag8) rather than two (lag2) as shown.

anchored at 0 (not at all hungry) and 7 (extremely hungry). We used their responses as a manipulation check.

2.5.3. *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.

2.5.4. *Dietary restraint*

Participants completed the dietary restraint scale (Herman & Polivy, 1975; Herman, Polivy, Pliner, Threlkeld & Munic, 1978) after the EBA task in their second session. We scored participants using the methods of Herman and 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 the Supplementary Materials (Table S1).

2.6. *Data analysis*

Data were analysed and visualised in R (R Core Development Team, 2020). 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 Table S2 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 Table S3.

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., 2019).

3. 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 Tables S4–9.

3.1. Hunger manipulation check

Participants had higher hunger ratings in their hungry session than in their sated session (BFs >1000; Table 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%).

3.2. Blink of attention check

Accuracy was higher in lag8 trials, compared to lag2 trials (Table 1; Fig. 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 (Table S4) provided evidence for accuracy differences in lag2 and lag8 trials for each distractor category, in line with PPZ (Fig. 2).

3.3. Practice effect check

There was evidence to support a practice effect in E1 and E2 (Table 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., 2019).

3.4. 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 2; Fig. 2). Thus, P1 was not supported. In E2, there was no evidence of a state by image category interaction effect in lag8 trials (Table 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 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.

Table 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 trials.

Study	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 (11.07)	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.

3.5. 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 (Table S10). 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 (Table S11). 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 (Table S12). However, supplementary frequentist analyses suggested a main effect of age and an interaction effect of age by state by image category in E2 (Table S12). 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 (Table S12) were not significant, suggesting that state did not impact performance on food trials regardless of age.

4. 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 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 (Table S1). 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.

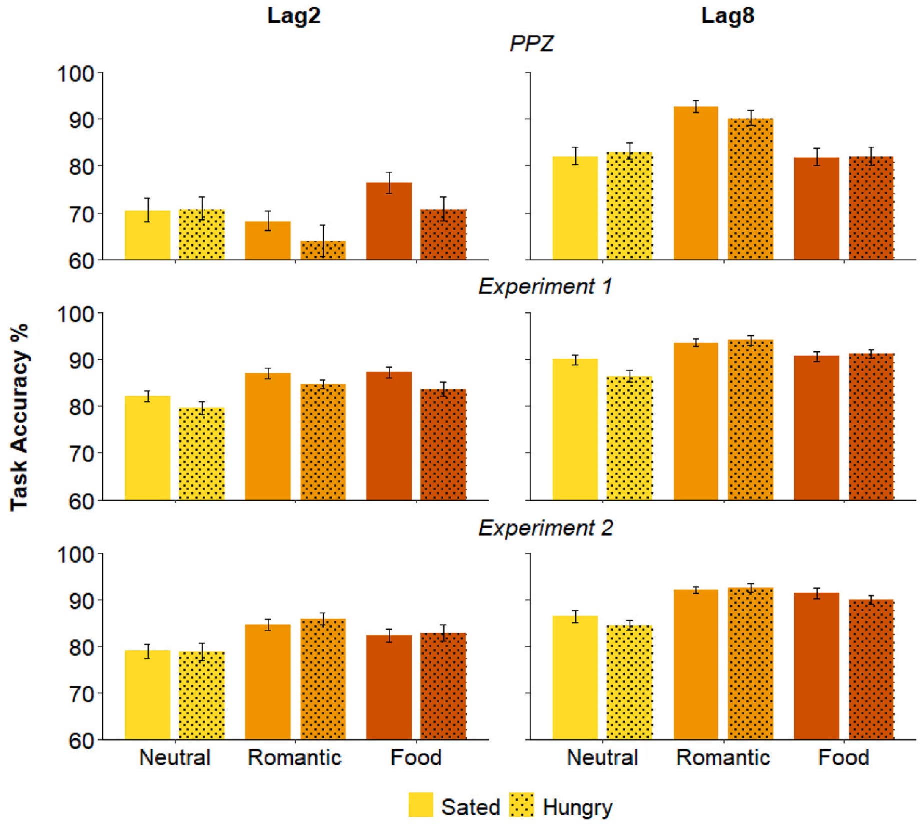

Fig. 2. Accuracy in PPZ, and E1 and E2, separated by lag, image category, and state.
Note. Trials are grouped by image distractor category, lag, and state. Lower task performance was hypothesised to indicate a greater attentional blink effect. The PPZ plots have been reproduced using Graph Data Extractor (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 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	Evidence for null	Evidence for null
$p = .03$	BF = 0.06	BF = 0.06
$F(2, 21) = 3.80$	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	Inconclusive	Evidence for null*
Not reported	BF = 0.26	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.

4.1. 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 in 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.

4.2. 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.

4.3. 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.

4.4. 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 $\geq$ 80% and 20 USD if they scored $\geq$ 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 $\geq$ 80% or $\geq$ 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.

4.5. Differences in accuracy across image categories and image sets

We found notable differences in the main effects of image category on accuracy (Tables S4–8) 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 were 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 (Table S10). 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.

4.6. 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. (2022) 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.

4.7. 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.

7. author contributions

CN: Conceptualisation, methodology, software, formal analysis, investigation, data curation, writing – original draft, writing – review and editing, visualisation, project administration, funding acquisition. GVP and CA: Conceptualisation, methodology, writing – review and editing. DN: Conceptualisation, methodology, writing – review and editing, funding acquisition. All authors (CN, GVP, CA, and DN) approved the final version of the manuscript.

Funding declaration

This work was supported by the Economic and Social Research Council funded NINE Doctoral Training Partnership [grant number ES/P000762/1]. The funders were not involved in any stage of the study.

Ethics declaration

The Newcastle University Faculty of Medical Science Research Ethics Committee (reference 8999/2020) granted ethical approval for both studies. Informed consent was obtained from all participants.

Declaration of competing interest

Declarations of interest: none.

Data availability

The link to our code and data is in a data availability statement at the end of the manuscript.

6. Acknowledgements

The authors thank the original authors, Professor David H. Zald and Dr Richard M. Piech, for sharing study materials and additional details. The authors also thank Dr Oliver M. Shannon for his helpful feedback on the manuscript and NINEDTP for supporting this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.appet.2023.107065.

References

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

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

Brandt, M. J., Ijzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J., 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

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*, 1063–1073. https://doi.org/10.1038/ijo.2009.138

Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 91–115). New York: Guilford.

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

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*, 767–780. https://doi.org/10.1037/hea0000405

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

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*, Article 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*, Article 104986. https://doi.org/10.1016/j.appet.2020.104986

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

Herman, C. P., Polivy, J., Plliner, P., & Threlkeld, J. (1978). Distractibility in dieters and nondieter: An alternative view of "externality". *Journal of Personality and Social Psychology, 36*, 536–548. https://doi.org/10.1037//0022-3514.36.5.536

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

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. University of Florida. Technical Report A-8 https://www2.unifesp.br/dpsicobio/adap/instructions.pdf. (Accessed 27 April 2023).

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

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

McElreath, R. (2020). *Statistical rethinking: A bayesian course with examples in R and stan* (2nd ed.).

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

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

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*, 579–582. https://doi.org/10.1016/j.appet.2009.11.003

R Core Development Team. (2020). *[Computer software] R: A language and environment for statistical computing (4.01.* R Foundation for Statistical Computing.

Redlich, D., Memmert, D., & Kreitz, C. (2022). Does hunger promote the detection of foods? The effect of value on inattentional blindness. *Psychological Research, 86*, 98–109. https://doi.org/10.1007/s00426-021-01480-y

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, 1-20*. Advance online publication. https://doi.org/10.3758/s13414-023-02677-6

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

Stoet, G. (2010). PsyToolKit – a software package for programming psychological experiments using Linux. *Behaviour Research Methods, 42*, 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*, 24–31. https://doi.org/10.1177/0098628316677643

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

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*, 291–298. https://doi.org/10.1177/1745691611406923