

Food's visually perceived fat content affects discrimination speed in an orthogonal spatial task

Vanessa Harrar · Ulrike Toepel · Micah M. Murray · Charles Spence

Received: 31 May 2011 / Accepted: 2 August 2011 / Published online: 20 August 2011
© Springer-Verlag 2011

Abstract Choosing what to eat is a complex activity for humans. Determining a food's pleasantness requires us to combine information about what is available at a given time with knowledge of the food's palatability, texture, fat content, and other nutritional information. It has been suggested that humans may have an implicit knowledge of a food's fat content based on its appearance; Toepel et al. (Neuroimage 44:967–974, 2009) reported visual-evoked potential modulations after participants viewed images of high-energy, high-fat food (HF), as compared to viewing low-fat food (LF). In the present study, we investigated whether there are any immediate behavioural consequences of these modulations for human performance. HF, LF, or non-food (NF) images were used to exogenously direct participants' attention to either the left or the right. Next, participants made speeded elevation discrimination responses (up vs. down) to visual targets presented either

above or below the midline (and at one of three stimulus onset asynchronies: 150, 300, or 450 ms). Participants responded significantly more rapidly following the presentation of a HF image than following the presentation of either LF or NF images, despite the fact that the identity of the images was entirely task-irrelevant. Similar results were found when comparing response speeds following images of high-carbohydrate (HC) food items to low-carbohydrate (LC) food items. These results support the view that people rapidly process (i.e. within a few hundred milliseconds) the fat/carbohydrate/energy value or, perhaps more generally, the pleasantness of food. Potentially as a result of HF/HC food items being more pleasant and thus having a higher incentive value, it seems as though seeing these foods results in a response readiness, or an overall alerting effect, in the human brain.

Keywords Food · Fat content · Caloric content · Response time · Visual selective attention · Exogenous cuing · Pleasantness · Spatial discrimination · Carbohydrate

V. Harrar (✉) · C. Spence
Department of Experimental Psychology,
University of Oxford, South Parks Road,
Oxford OX1 3UD, United Kingdom
e-mail: vanessa.harrar@psy.ox.ac.uk

U. Toepel · M. M. Murray
Radiology Department, Centre Hospitalier Universitaire Vaudois
and University of Lausanne, Lausanne, Switzerland

U. Toepel · M. M. Murray
Neuropsychology and Neurorehabilitation Service,
Department of Clinical Neurosciences,
Centre Hospitalier Universitaire Vaudois and University
of Lausanne, Lausanne, Switzerland

M. M. Murray
EEG Brain Mapping Core,
Center for Biomedical Imaging of Lausanne and Geneva,
Lausanne, Switzerland

Introduction

The ability to visually recognise food is critical to our survival, so much so that it has been suggested that colour vision may have evolved in order to facilitate finding food (such as ripe red fruit) amongst a visually dense background of foliage (e.g. Sumner and Mollon 2000). While the ability to find provisions was clearly critical at some earlier point in human evolution, many now have the luxury of choosing what to eat and when. When choosing what to eat, the pleasantness, and presumably also the nutritional content of the available food might be initially assessed purely through visual inspection. Thus, vision, or

the allocation of visual attention to food, still seems important.

Nummenmaa et al. (2011) reported that visual selective attention is preferentially directed towards food items. These researchers showed that participants searching for visual food or non-food targets detected and responded to food items more quickly than to non-food items. The direction of visual attention towards food appears to be particularly affected by a person's hunger state. For example, Piech et al. (2010) reported that food images attracted the attention of hungry participants more than that of participants who were not hungry. They demonstrated poorer performance when detecting visual targets (rotated landscape images amongst a series of upright images) when the targets were presented 200 ms after a food image, whereas attention capture effects of non-food images (romantic or neutral in nature) did not depend on the participants' hunger level. Similarly, hunger has been shown to increase visual selective attention to food-related words (Mogg et al. 1998), to improve people's memory for visually presented food items (Morris and Dolan 2001) and to limit people's ability to shift their visual attention (Piech et al. 2009). Similarly, di Pellegrino et al. (2011) have demonstrated that attention effects towards normally pleasant food could be decreased if that food was consumed to satiety, at which point it was no longer rated as being particularly "pleasant". Since, in their study, pleasantness ratings were significantly correlated with the attentional bias, they concluded that the allocation of visual attention is related to an individual's food preferences at a given moment in time. If the pleasantness of a food predicts the attentional bias, then, presumably, foods that are more often rated as pleasant should consistently produce a larger attentional bias as compared to foods that are not.

Here we investigate the more stable allocation of visual attention towards food items—allocations that vary with the type of food objectively categorised rather than subjectively rated at a given moment. The question here is whether there is some characteristic in the food (e.g. fat or carbohydrate content) that captures attention. Perhaps high-fat foods, generally rated as more pleasant, might capture a person's visual attention more effectively than low-fat foods, regardless of individual differences in satiety or hunger. Toepel et al. (2009) revealed modulations in visual-evoked potentials (VEPs) after participants viewed images of high-fat (HF) as compared to low-fat foods (LF). Toepel and her colleagues thus suggested that humans have an implicit knowledge of the fat content of foods. In the present study, an exogenous spatial cuing paradigm was used in order to test for behavioural/attentional effects of the differential brain activation related to the fat content in foods.

In a typical exogenous cuing paradigm, a cue (here a picture of a food item) is presented with equal probability from either side of fixation. Next, a target is presented either near the cue or on the opposite side, and participants are instructed to respond to the target as rapidly and accurately as possible. Because the cue is task-irrelevant, both spatially and semantically, it does not provide any information that would facilitate responding to the target; i.e. the target should be localised at the same speed regardless of its spatial proximity to the cue. However, the cue appears to exogenously shift a participant's attention to that spatial location, and a decreased reaction time (RT) is typically observed when targets are presented on the same side as the cue (the spatial cuing effect), even when the cue and target are presented in different sensory modalities (see Spence 2010 for a review).

Using the exogenous spatial cuing paradigm, we hypothesised that if HF foods are more attended to than LF foods, perhaps as a result of being more pleasant, one might expect to find an increased shift of spatial attention (and consequently a larger spatial cuing effect) when images of HF foods capture a person's spatial attention. Alternatively, an image of a food item might instead have a non-spatial attention effect (such as seen following the presentation of an alerting cue, e.g. Posner 1978), which speeds up performance regardless of the location of the target—that is, regardless of whether it appears on the cued side or not.

Methods

Participants

Twenty-five participants were tested, 20 women and 5 men with a median age of 27 years. The experiment was approved by the Central University Research Ethics Committee of the University of Oxford. Participants gave their informed consent before taking part in the study. The experiment lasted for approximately 20 min, and participants were compensated for taking part with a £5 (UK Sterling) gift voucher.

Procedure

Participants sat in a sound-attenuated booth and were presented with images on a 19" colour CRT monitor, with a 60 Hz refresh-rate, placed 54 cm in front of the participants' eyes. Each participant rested his or her head on a chin rest. A fixation cross (.05°) appeared at the start of the trial and remained on the screen until the end of the trial. The participants were instructed to maintain fixation on that position even when an image was flashed in the

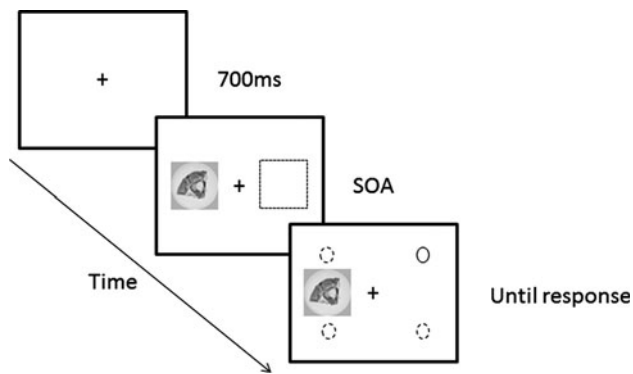


Fig. 1 Methods. The *first frame* shows the fixation cross, which was seen by participants for 700 ms. The *second frame* shows the visual cue (a slice of pizza) appearing to the left of the fixation cross—a *dashed rectangle* shows the other possible location for the visual cue. The *third frame* shows a visual target (not drawn to scale) presented in the *top right corner* (the other three possible locations for the visual target are depicted by *dashed circles*). The condition shown in the figure is a non-cued trial with a high-fat food image. The visual cues, although *greyscale* here, were in colour for the experiment

periphery. After 700 ms, the visual cue (one of three types of images: HF, LF, or NF) was presented randomly on either the left or the right of fixation. The square image was 8.5° wide, presented at an eccentricity of 13.3° . After one of three stimulus onset asynchronies (SOAs: 150, 300, or 450 ms), an empty circle $.05^\circ$ diameter (the visual target) was presented either above or below the visual cue, or in the mirror location on the opposite side of fixation (at an eccentricity of 12.5° ; see Fig. 1). The participants were instructed to respond as rapidly as possible by pressing one of two keys on the number pad of a standard keyboard. Their task was to report the position of the visual target (up/down). Between responses, the participants were instructed to hold their index finger on a middle resting key (#5) which had a little bump as a tactile cue. Whenever the target was presented above the midline, participants responded “UP” using the #8 key (directly above the resting #5 key). If the target was presented below the midline, they responded “DOWN” using the #2 key (directly below the resting #5 key). The participants were given 2 s to respond; otherwise, the trial was terminated, and no response was recorded for that trial. The presentation of stimuli and the recording of responses were controlled by E-Prime (Psychology Software Tools Inc., Pittsburgh, USA; <http://www.pstnet.com/eprime>). Initially, there were 8 trials (4-UP and 4-DOWN; 3HF, 3LF, and 2NF images) as practice. Each of these trials was followed by a feedback screen that also displays the participants’ RT. In the data collection phase, there were 72 conditions (3 image types \times 3 SOAs \times 2 image locations \times 4 target locations) per block. There were 10 blocks of trials, between which participants could take breaks, which took about 20 min to complete in total.

Image parameters

There were three image types: HF, LF, and NF. The latter category consisted of pictures of kitchen utensils (see Fig. 2a). The food images were subdivided into HF and LF based on their fat content, using the Swiss nutritional database (released by the Swiss Federal Office of Health and the Swiss Federal Institute of Technology Zurich), and, in cases where the items were not listed, the nutrition database of the U.S. Department of Agriculture (<http://www.nal.usda.gov/fnic>). Both the HF and the LF image classes comprised 50% savoury foods and 50% sweet foods. All of the photographs measured 150×150 pixels, which corresponded to 8.5° visual angle on the computer monitor and were taken using an identical background and from an identical bird’s-eye-view. The images were controlled for low-level physical stimulus features (see Knebel et al. 2008 for details on the procedure). First, spatial frequency differences between the three image sets were minimised by selecting a subset of images, from a larger image database, which yielded the highest spatial frequency similarity (in terms of spectral power value across image space). Visual inspection of histograms of the spatial spectral power reveals the highly similar spatial frequency pattern between, in particular, the HF and LF images (see Fig. 2b). In addition, the luminance of each image was individually adjusted to a mean value calculated with reference to the entire image set.

There were 24 images in each of the three categories (see Fig. 2a for examples). The fat content of the LF images ranged from .1 to 5.0 g/100 g (mean \pm sem = $1.0 \pm .3$ g) and for the HF group from 11.0 to 58.0 g/100 g (mean \pm sem = 23.3 ± 2.3 g). Respective mean (\pm sem) energetic values for the food classes were 154.1 ± 27.9 kcal/100 g for LF images and 369.6 ± 29.3 kcal/100 g for HF images. The foods contained in the HF and LF images differed significantly in terms of their fat content per 100 g ($t_{46} = 9.5$, $P < .001$) and energy value ($t_{46} = 5.3$, $P < .001$).

Importantly, the foods contained in the HF and LF images were controlled for in terms of carbohydrate content; i.e. the two groups were not significantly different in this regard ($t_{46} = .4$, $P < .688$). Given this, and given the likelihood that any effects related to fat content might be related to pleasantness or energy content as the underlying factor, an additional post hoc analysis (suggested by a referee) was conducted comparing images of high-carbohydrate foods (HC) to images containing low-carbohydrate foods (LC). The division of HC and LC was done by splitting the total number of food images (48) into two groups based on the carbohydrate content. Thus, the carbohydrate content of the LC foods ranged from 0 to

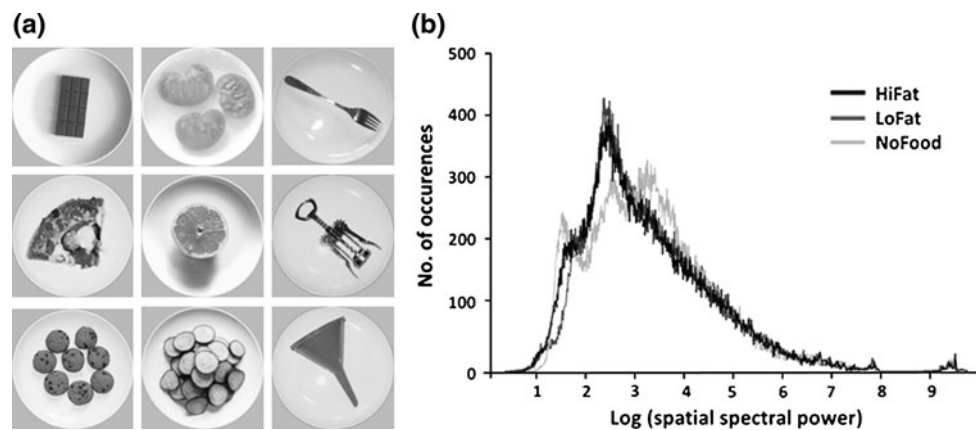


Fig. 2 The three image types. **a** A subset of the three types of images used in the experiment are shown here: high-fat food (*left column*), low-fat food (*middle column*), and non-food items (*right column*). The images, although *greyscale* here, were in colour for the experiment.

18.0 g/100 g (30.6 ± 7.2) and the HC group ranged from 18.2 to 93.5 g/100 g (34.2 ± 5.4).

Results

Participants responded erroneously on less than 1% of trials overall. Outliers were removed on an individual basis, using the Z-score >3 rule, which resulted in the removal of 295 responses (1.6%). RTs for those trials in which the participants responded correctly were averaged across the 10 blocks, left and right image locations, and upper and lower target locations. Thus, there were three independent measures of interest (3 image types, 3 SOAs, and 2 spatial cuing: cued or uncued), which were assessed using a repeated measures analysis of variance (ANOVA). Sex was also analysed as a between subject variable in the ANOVA, but there was no significant difference ($F_{1,23} < 1$), and therefore, it did not enter any further analysis. The results revealed a significant main effect of spatial cuing ($F_{1,24} = 30.43$, $P < .001$, $\eta_p^2 = .559$) with participants responding more rapidly to cued targets (those presented on the same side as the visual cue) than to uncued targets (those presented on opposite the visual cue, see Fig. 3a). There was also a significant main effect of SOA (Greenhouse–Geisser correction for sphericity, $F_{2,48} = 203.95$, $P < .001$, $\eta_p^2 = .895$; see Fig. 3b). Pairwise comparisons revealed significant differences between all three delays ($P < .001$ for all three comparisons) such that as the SOA increased, RTs decreased—with faster responses being observed as the onset of the target was delayed with respect to the onset of the cue. There was also a significant interaction between SOA and spatial cuing ($F_{2,48} = 3.97$, $P = .025$, $\eta_p^2 = .142$).

Of particular interest, there was a significant main effect of image type ($F_{2,48} = 4.59$, $P = .015$, $\eta_p^2 = .160$; see

b Histograms of the spatial spectral power for high-fat food images (*black*), low-fat food images (*dark grey*), and non-food images (*light grey*) demonstrating highly similar spatial frequency across images sets

Fig. 3c). Pairwise comparisons revealed that RTs following HF images were significantly faster than RTs following LF images ($P = .021$) or NF images ($P = .021$). The difference between RTs following LF and NF images was not significant ($P = .771$). Finally, the ANOVA did not reveal a significant interaction between image type and spatial cuing ($F_{2,48} < 1$, power = .083); that is, while participants' responses to targets presented on the same side as the visual cue were faster than targets on the opposite side, they were not additionally faster if the cue was a HF image compared to a LF or NF image.

A post hoc analysis was also conducted using the images divided into HC and LC groups. As in the HF/LF analysis, there were 18 conditions made up of three independent measures of interest (3 image types, 3 SOAs, and 2 spatial cuing: cued or uncued) assessed with a repeated measures ANOVA.¹ The pattern of results for the carbohydrate analysis was the same as the fat content analysis: a significant main effect of spatial cuing ($F_{1,24} = 29.16$, $P < .001$, $\eta_p^2 = .549$); a significant main effect of SOA (Greenhouse–Geisser correction for sphericity, $F_{2,48} = 198.6$, $P < .001$, $\eta_p^2 = .892$); a significant interaction effect between SOA and spatial cuing ($F_{2,48} = 4.05$, $P = .024$, $\eta_p^2 = .144$); and, of particular interest, a significant main effect of image type ($F_{2,48} = 5.76$, $P = .006$, $\eta_p^2 = .193$) with no significant interaction between image type and spatial cuing ($F_{2,48} = 1.65$, power = .331). Pairwise comparisons revealed, as before, that RTs following HC images were significantly faster than RTs following LC images ($P = .007$) or NF images ($P = .008$). The

¹ Given that this analysis was post hoc, the number of stimuli for each participant was not the same for each condition. For the fat analysis, there were 40 RTs that were averaged for each of the 18 conditions, while in the carbohydrate content analysis, there were 29–50 RTs averaged for each condition.

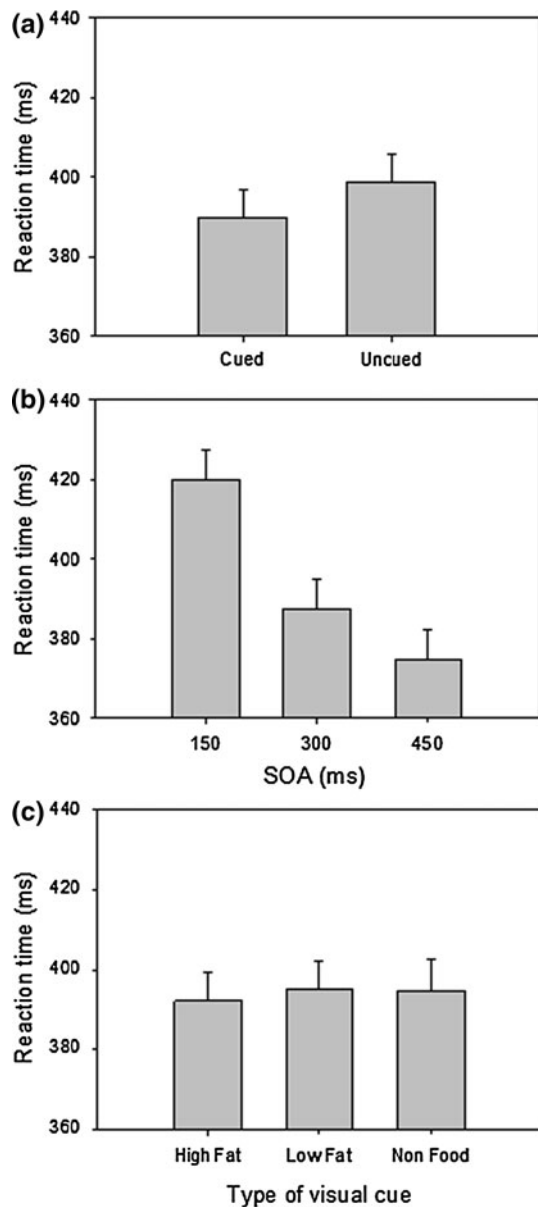


Fig. 3 Reaction time (RT) results. Means and standard errors of the mean are shown for each of the relevant conditions. **a** Cued images were reacted to faster than uncued images ($P < .001$). **b** As the time between the visual cue and the visual target increased, reactions to the target were significantly faster ($P < .001$). **c** RTs to targets following high-fat food images were significantly faster than low-fat food images ($P = .021$) and non-food images ($P = .021$)

difference in RTs following LC and NF images was not significant ($P = .674$).

Discussion

Consistent with previous findings using an orthogonal spatial cuing paradigm (cf. Spence et al. 2004), the results of the present study highlighted the existence of a small but

significant exogenous spatial cuing effect ($M = 10$ ms). The largest spatial cuing effect was observed at the shortest SOA of 150 ms (mean cuing effect of 13 ms). As the SOA increased, the spatial advantage decreased presumably because inhibition of return (IOR) was competing with (and potentially masking) the cuing effect (Klein 2000). The novel effect observed in the present study was the decrease in RTs (i.e. enhanced behavioural responding) in response to visual targets presented shortly after images of high-fat or high-carbohydrates food—even though the images of the food were completely irrelevant to the participants' visual elevation discrimination task. The results reported here are important as they suggest completely unintuitive differences in the ways in which images are processed even when the content of the images is irrelevant to the task, as in an orthogonal visual spatial cuing task.

We hypothesised that HF images might be more salient than LF or NF images. That is, that HF images would stand out more than other images, attract participants' spatial attention more than other images, and cause attention to be focused in the direction of the HF image. If HF images were more salient, then we would expect to find a larger spatial cuing effect for HF images as compared to the spatial cuing effect of either LF or NF images; targets presented on the same side as a HF image should be responded to even faster than targets presented on the same side as LF or NF images. However, the lack of an interaction effect between the type of image and the cuing effect means that there was no additional enhancement for cued trials in which HF (or HC) images were presented. This result suggests that the advantage for the HF food images does not appear to be related to spatial attention per se; instead, the HF (and HC) food advantage appears to result in a more general speeding up of human information processing.

This general performance increase is similar to, and could potentially be explained by, Pavlovian-instrumental transfer (PIT). In PIT, an unrelated stimulus (such as food) has been shown to increase the vigour of an action following the stimulus even though the action and stimulus are completely unrelated (cf. Talmi et al. 2008). The increase in activity following the stimulus is related to the Pavlovian-type (associative) learning, pre-experimentally, of the motivational nature of the stimulus or food. From the results presented here, it appears as though HF/HC foods have a larger PIT, are simply more motivating, than LF/LC foods.

Are previous results relating attention allocation to food also caused by non-spatially specific increases in activity, such as those seen in PIT? Attentional differences in response to the presentation of images of food items have already been reported. For example, di Pellegrino et al. (2011) used a visual probe task and reported a larger bias in

attention to foods that were rated as more pleasant. They found that after consuming the food, pleasantness ratings decreased, and RTs to probes/targets placed at the same location as the image of the food were slower. However, as they concede in their discussion, their experimental design was not able to reveal whether the attentional effects that they reported were spatially specific or not.

The design of the present study, however, can distinguish between these two hypotheses. The behavioural enhancement seen in response to the presentation of HF food images is not, strictly speaking, an attentional effect—or at least not a spatially specific one. Nummenmaa et al. (2011) reported that images of food items attracted attention more than non-food items, but failed to report a spatial attention effect when comparing the effects of presenting pleasant as compared to bland-tasting foods. Nummenmaa et al.'s lack of any spatial attentional difference between pleasant and neutral foods, as shown here when comparing HF and LF foods, might be related to the fact that all of the food stimuli used in these various studies were actually fairly pleasant.

There is consistent evidence that, in comparison to neutral stimuli, negative or unpleasant stimuli have a greater impact on the allocation of attention than positive stimuli (for a review, see Baumeister et al. 2001). That is, negative stimuli (such as scary faces, see West et al. 2009) appear to attract attentional resources more robustly than do positive stimuli. Thus, the present inconsistency might be due to the fact that the very few published papers on the topic have all used foods that are fairly positive, or at the very least neutral (Brignell et al. 2009; di Pellegrino et al. 2011; Nummenmaa et al. 2011). The use of more negative or aversive foods (for example, rotten or mouldy food) in future research would likely provide larger and more robust attention effects. Furthermore, these studies did not necessarily control for any differences between spatially specific attention effects and general increases in alertness. Future experiments should, therefore, use experimental paradigms that can distinguish between general alerting effects and/or spatial attention effects.

From the results reported here, it appears as though images of HF and HC food give rise to a non-spatially specific enhancement of behavioural responding. This enhancement is likely due to, or associated with, the underlying factor of pleasantness—HF and HC food items are generally more pleasant, which might give them a higher incentive value. This higher incentive value is associated with an overall increase in alertness in the brain. It would seem, therefore, that simply seeing the food actually has biological benefits that can affect the distribution of resources in the brain and the behavioural outcomes that follow.

Acknowledgments Vanessa Harrar holds a Mary Somerville Junior Research Fellowship. Ulrike Toepel and Micah Murray receive support from an interdisciplinary project awarded by the Faculty of Biology and Medicine at the University of Lausanne. Micah Murray receives support from the Swiss National Science Foundation (grant 310030B-133136). Thanks to Jean-Francois Knebel for providing help with the image selection and preparation.

References

- Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD (2001) Bad is stronger than good. *Rev Gen Psychol* 5:323–370
- Brignell C, Griffiths T, Bradley BP, Mogg K (2009) Attentional and approach biases for pictorial food cues. Influence of external eating. *Appetite* 52:299–306
- di Pellegrino G, Magarelli S, Mengarelli F (2011) Food pleasantness affects visual selective attention. *Q J Exp Psychol* 64:560–571
- Klein R (2000) Inhibition of return. *Trends Cogn Sci* 4:138–147
- Knebel JF, Toepel U, Hudry J, Le Coutre J, Murray MM (2008) Generating controlled image sets in cognitive neuroscience research. *Brain Topogr* 20:284–289
- Mogg K, Bradley BP, 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? *Behav Res Ther* 36:227–237
- Morris JS, Dolan RJ (2001) Involvement of human amygdala and orbitofrontal cortex in hunger-enhanced memory for food stimuli. *J Neurosci* 21:5304–5310
- Nummenmaa L, Hietanen JK, Calvo MG, Hyönä J (2011) Food catches the eye but not for everyone: a BMI-contingent attentional bias in rapid detection of nutrients. *PLoS ONE* 6(5):e19215
- Piech RM, Hampshire A, Owen AM, Parkinson JA (2009) Modulation of cognitive flexibility by hunger and desire. *Cogn Emot* 23:528–540
- Piech RM, Pastorino MT, Zald DH (2010) All I saw was the cake. Hunger effects on attentional capture by visual food cues. *Appetite* 54:579–582
- Posner MI (1978) Chronometric explorations of mind. Lawrence Erlbaum, Hillsdale
- Spence C (2010) Crossmodal spatial attention. *Ann N Y Acad Sci* 1191:182–200
- Spence C, MacDonald J, Driver J (2004) Exogenous spatial-cuing studies of human cross-modal attention and multisensory integration. In: Spence C, Driver J (eds) *Crossmodal space and crossmodal attention*. Oxford University Press, Oxford, pp 277–320
- Sumner P, Mollon JD (2000) Catarrhine photopigments are optimized for detecting targets against a foliage background. *J Exp Biol* 203:1963–1986
- Talmi D, Seymour B, Dayan P, Dolan RJ (2008) Human pavlovian-instrumental transfer. *J Neurosci* 28:360–368
- Toepel U, Knebel JF, Hudry J, le Coutre J, Murray MM (2009) The brain tracks the energetic value in food images. *Neuroimage* 44:967–974
- West G, Anderson A, Pratt J (2009) Motivationally significant stimuli show visual prior entry: direct evidence for attentional capture. *J Exp Psychol Hum Percept Perform* 35:1032–1042