Comparing measures of cognitive bias relating to eating behavior

Emmanuel M. Pothos, Raff Calitri, Katy Tapper

Department of Psychology

Swansea University

Jeffrey M. Brunstrom, & Peter J. Rogers

Department of Experimental Psychology

University of Bristol

<u>Research Article</u>

Running head: cognitive biases for food stimuli; **word count**: 5,669

Please address correspondence regarding this article to Emmanuel Pothos, Department of Psychology, Swansea University, Swansea SA2 8PP, UK. Electronic mail may be sent at e.m.pothos@swansea.ac.uk.

Abstract

Consumption of and/or abstinence from substances with a high reward value (e.g., heroin, marijuana, alcohol, nicotine, certain foods) are associated with cognitive biases for information related to the substance. Such cognitive biases are important since they may contribute to difficulties in controlling intake of the substance. We examine cognitive biases for stimuli related to food. For the first time, we concurrently employ and compare five conceptually distinct measures of cognitive bias (dot probe, emotional Stroop, recognition, EAST, explicit attitudes). Contrary to expectations from current theory, the relation between the cognitive measures was weak and evident only in certain subsets of the population sample, as defined by gender and emotional-, restrained-, and external-eating characteristics of our participants. We discuss some methodological implications of our findings.

Consuming and/or abstaining from a substance with a high reward value is often associated with cognitive biases for information relating to this substance. This has been observed with a wide range of substances (for example, alcohol, heroin, marijuana, nicotine, and certain foods—the latter is the focus of the present study; Cox et al., 2002; Field, 2006; Franken et al., 2000; Stacy, 1997), other types of addictive behavior (e.g., gambling; Boyer & Dickerson, 2003), as well as other types of psychological problems (e.g., excessive anxiety; Mogg & Bradley,1998; for reviews see Cox, Fadardi, & Pothos, 2006, Hogarth & Duka, 2006, and Williams, Mathews, & C. MacLeod, 1996). Substance abuse is the most common behavior with which cognitive biases have been observed and relevant theory is often formulated in terms of substance abuse.

Cognitive biases have been demonstrated with several paradigms. In an alcohol version of the Stroop task an *attentional* bias for alcohol related information is evident in that alcohol abusers require more time to name the ink color of alcohol words compared to neutral ones (e.g., Cox et al., 2006). Memory biases have been observed in recall/ recognition tasks, whereby (e.g.) alcohol-related words will be produced more readily in stem completion tasks and will be recalled better by heavy drinkers, compared to neutral words (Jones & Schulze, 1999; Stacy, 1997). Multidimensional scaling analyses show a *categorization* bias, in that alcohol-related concepts are closer to positive alcohol expectancies for alcohol abusers, but not for light drinkers (Rather et al., 1992). Pothos and Cox (2002) observed a *learning* bias. Alcohol abusers were impaired in an Artificial Grammar Learning task (Pothos, 2007) when the stimuli were instantiated as sequences of drinks, but not when they were instantiated as sequences of neutral words.

An underlying *assumption* into research on cognitive biases related to substance abuse is that different cognitive tasks simply measure different aspects of the same cognitive process. However, there have been no direct, comprehensive comparisons between cognitive biases. The objective of this study is to address this shortcoming and so examine the above assumption. If different cognitive biases reflect the same process, one would predict that they would correlate with each other for the same individual.

Some comparative work has already been carried out. Field, Mogg, and Bradley (2004) reported a correlation between attentional bias as measured by the dot probe and gaze shifts measured by eye tracking, in the context of nicotine deprivation (see also, Mogg, Field, & Bradley, 2005; Mogg et al., 2003). More relevant to the present research, both Egloff and Hock (2003) and Duka and Townshend (2004) found the expected association between emotional Stroop and dot probe, the former with threat-related stimuli and the latter using an alcohol pre-load with social drinkers (although note that Duka & Townshend examined Stroop results using error rate, as opposed to response latencies, which is the more common dependent variable). This research is certainly suggestive, but it offers limited guidance in anticipating the conclusions of a more comprehensive comparison between measures of cognitive bias.

It is impractical to carry out a comparison of different measures of cognitive bias across the full range of substances/ behaviors for which cognitive biases have been observed. We chose to focus on food-related stimuli. Thus, our conclusions do not necessarily generalize to cognitive biases for other types of stimuli (e.g., alcohol). Nonetheless, existing theory regarding cognitive biases is not specific to a particular type of substance/ behavior (e.g., Cox & Klinger, 2004; Hogarth et al., 2006;

Robinson & Berridge, 1993; Tiffany, 1990). In other words, the theory for why alcohol abusers display a cognitive bias for alcohol-related information is identical to the corresponding theory for, e.g., heroin users. Accordingly, we can tentatively be optimistic regarding the generalizability of our conclusions.

Methodologically, the emphasis on food stimuli is convenient. Cognitive biases relating to food stimuli are ubiquitous. For example, there is evidence that such biases arise for people who are particularly sensitive to the appetitive qualities of food (external eaters; Drayna, 2005; Franken & Muris, 2005; Robinson & Berridge, 1993) and people who try to restrict their food intake (restrained eaters; Tapper et al., in press). However, note that it has been difficult to find correlations between the food version of the Stroop task and body mass index (BMI), which is a measure of how overweight a person is (e.g., Boon, Vogelzang, & Jansen, 2000; Francis, Stewart, & Hounsell, 1997; Green & Rogers, 1993; Overduin, Jansen, & Louwerse, 1995).

Some researchers have already indirectly challenged the assumption that different cognitive biases correspond to the same psychological process. Field (2006) suggested that the short and long exposures in the dot probe task (see later) reflect differential effects of craving. Moreover, Hogarth and colleagues demonstrated that in some cases a cognitive bias is associated with performance (it results from an expectation of the drug; Hogarth et al., 2007a), in other cases with learning (participants attended more to an uncertain predictor of an outcome, and cognitive biases were observed only for participants who were aware of a contingency between predictors and outcomes; Hogarth et al., 2007b). Promising as these results are, they are not a substitute for a direct comparison between measures of cognitive bias.

Finally, research into cognitive biases is important for (at least) two reasons. First, there has been a promise of practical application. The alcohol Stroop task appears to have diagnostic value, since alcohol Stroop interference can predict whether a person is a light, medium, or heavy drinker (Cox et al., 2007). Moreover, alcohol Stroop interference can predict changes in number of drinking days for excessive drinkers, six months after testing (Cox, Pothos, & Hosier, 2007). Such results have led some researchers to pursue cognitive-style interventions for substance abuse (Wiers et al., 2006). Second, research into cognitive biases can inform cognitive science theory. Some explanations of cognitive biases postulate that for, e.g., an alcohol abuser there are automatic links between alcohol information, aspects of her environment, and emotions related to alcohol consumption (Cox et al., 2006; McKenna & Sharma, 1995; Peretti, 1998; Stetter et al., 1995; Tiffany, 1990). The development of such links is a challenge for current theories of automaticity (e.g., Logan, 1998; Tiffany, 1990; Tzelgov, 1997).

Cognitive bias measures

In a food version of the Stroop task participants see either food-related words or neutral words and they have to name the ink color of each word. The rationale is that the greater the attentional bias for food-related information, the longer it will take participants to disengage their attention from a food word and identify its ink color, relative to the time required for neutral words.

An alternative task for measuring attentional biases is the dot probe task (C. MacLeod, Mathews, & Tata, 1986). In each trial a participant sees a food-related and a neutral word. Then, the words disappear and a dot appears at the position of one of them. Participants have to identify the location of the dot as quickly as possible. For each possible pair of words, in one trial the dot replaces the food word and in a second trial it replaces the neutral word. If a participant is attending to the food word then she should be quicker to identify the location of the dot when it appears in the same location as the food word, compared to when it appears in the neutral word location. A possible problem with this interpretation is that a participant attends to a food word initially, but away from it at longer exposure times (Mogg et al., 2003; Rafal, Davies, & Lauder, 2006). A way to address this problem is by employing different exposure times, one that reflects initial attentional orientation (in the present study 500 ms) and one which corresponds to later processing (in the present study 1250 ms).

Memory biases were assessed with a recognition task. After the food Stroop and the dot probe tasks, participants were asked to recognize the food and the neutral words (twenty each), which were the stimuli in these tasks. A memory bias for food-related information would be indicated in higher recognition ability for the actual food words used, relative to distractor ones.

A cognitive bias for food-related information might also be shown in terms of a positive attitude towards food. We used the Extrinsic Affective Simon Task (EAST; De Houwer, 2003) to examine this issue. The EAST allows measurement of positive versus negative attitudes for a particular conceptual theme, such as ‘food’. It works by training participants to associate one key press response with ‘good’ stimuli and another with ‘bad’ ones. Subsequently, a comparison is made between responses to a food word with the ‘good’ key and with the ‘bad’ key; responses will be faster with the ‘good’ key when the individual holds a positive attitude towards the word and vice versa. The EAST is a simplified version of the better known Implicit Association Test (IAT; e.g., Nosek, 2005). The EAST and the IAT have been advocated as implicit measures of attitudes, in the sense that sometimes the EAST may reveal an attitude which participants deny having when questioned explicitly (cf., Pothos, 2007). Note, finally, that interpreting EAST results assumes that there is a meaningful variation in attitudes towards food, from positive to negative; in idiosyncratic samples, where (e.g.) all participants are highly positive, it is unclear how interpretable EAST results are. Finally, we also considered a direct measure of attitudes towards food (a four-item attitude questionnaire).

Measures of eating behavior

Our aim is to examine the relation between measures of cognitive bias for food stimuli, not the characteristics which may be affecting food intake (i.e., eating behavior). Nonetheless, it is important to consider some determinants of eating behavior, since we want to avoid a situation where all our participants would be too uniform in their eating behavior. Unless there is variability in eating behavior (and so in the corresponding cognitive biases), we might fail to find correlations between cognitive measures because the range of values in each measure is too narrow.

An intuitive measure of eating behavior is the BMI. High BMI can be associated with food cognitive biases, e.g., because it is sometimes a result of overeating, which might arise because of the higher incentive salience of food-related cues (cf. Brunstrom et al., in press;). The Dutch Eating Behavior Questionnaire (DEBQ; Van Strien, Frijters, Bergers & Defares, 1986; Van Strien, 1997) is a widely used measure of three kinds of eating behavior, external eating, restraint eating, and emotional eating. Emotional eating is a measure of how likely an individual is to overeat as a coping mechanism for stress, anxiety, and other similar, negative emotions (cf., Pecina, Schulkin, & Berridge, 2006). DEBQ constructs have been identified as moderators in the relationship between eating behavior and factors thought to affect eating behavior (such as stress; O’Connor et al., 2008). Equally, DEBQ indices have been used to predict cognitive biases relating to food information

(Tapper et al., in press). Finally, we used a hunger scale, since hunger might enhance cognitive biases for food-related stimuli (Field, 2006; Hogarth et al., 2007a).

Pilot investigation

We collected data from 25 participants to identify stimuli suitable for the three word categories required for the cognitive measures: healthy foods (10 words), unhealthy foods (10 words), and office words (for the neutral category). Using the MRC Psycholinguistic Database, potential word pairs were identified according to matching length, Kucera-Francis written frequency, and number of syllables. Participants were presented with the chosen words and were asked to rate them according to whether they belonged to the categories of “food” and “things found in the office”. Participants also rated the familiarity of the words, and the extent to which each food word represented a healthy or unhealthy food (no differences between the categories of interest). Finally, the valence of the office words were measured to ensure that these were neutral (this was the case). Note that we identified matching controls for the healthy and unhealthy food words separately (there is a subset of control words that is matched to the healthy food words, and the remaining control words were matched to the unhealthy food words). Appendix 1 lists the words.

Experimental investigation

Participants

One hundred and fifty one first year Swansea University students (88 female, 63 male), who identified themselves as living in shared student accommodation, native English speakers, and not color-blind, were recruited to take part in this research for a small payment (£10). We avoided participants still living with their parents, since such individuals would have less control over their diet.

Methods and procedure

The main components of the study were three experimental tasks (food Stroop, dot probe, recognition), two measures of attitudes towards food (the EAST and a brief questionnaire), the DEBQ, and the hunger scale (some additional measures were included). The visual display and response recording for the dot probe, Stroop, EAST and word recognition tasks were controlled by Cedrus SuperLab Pro (version 2.04) software, running on a fast PC. The monitor had a 1280 x 1024 resolution at 85 Hz refresh rate and was placed on a stand so that the bottom of the screen was 24.5cm from the tabletop. Participants responded using a Cedrus RB-730 7-button response box, to ensure ms accuracy.

Dot probe

The dot probe procedure requires participants to identify the location of a dot (the probe) across a series of trials. Each trial ran as follows: a fixation point ('+') appeared in the centre of the screen for 500 ms. Next, the word pair was presented. Each word in the stimulus pair was placed on the vertical axis approximately 15 mm from the screen centre, so that one word was above/below the other. Words were presented in white, bold, size 40, Arial, lowercase font on a black background. Characters were approximately 5 to 7 mm tall. The word pair remained on screen for either 500ms or 1250ms depending on experimental block (see below). After it disappeared, a dot (the probe; '.') was presented in one of the word locations. The dot remained on screen until a response. There was a 1000 ms inter-trial interval.

Two sets of dot probe trials were created, one in which the dot probe replaced the words quickly, 500 ms (initial orientation) and another in which the dot probe replaced the words after 1250 ms (sustained attention). In each word pair, control words were matched with food words according to length, number of syllables, written frequency, and familiarity. Trials where a food word and the probe appeared in the same location are referred to as *congruent*. Trials where the food word and the probe appeared in the opposite location are referred to as *incongruent*. There was a block of eight practice trials and a block of four buffer trials before the first experimental block of trials began. The practice and buffer trials comprised ‘number’ words (e.g., one, six, ninety, thirty). The objective of the practice and buffer trials here and elsewhere was to reduce practice effects on reaction times. The first experimental block assessed initial orientation and the second sustained attention. Before the second experimental block, there was another block of four buffer trials.

In each experimental block the location of the dot probe relative to the food words was fully counter-balanced. Also, all food and office words appeared in both the top half and bottom half of the screen. Therefore, each of the 20 food words (10 healthy, 10 unhealthy) appeared 4 times (congruent top, congruent bottom, incongruent top, incongruent bottom), resulting in 80 experimental trials. Word pair order was randomized. Finally, two dot probe tasks were constructed. In the first, participants had to press the left key on the response box when the dot appeared in the upper location (and right key when the dot appeared in the lower location). In the second, the response keys were switched. Participants were randomly assigned to one version of the task.

Food Stroop task

Participants were required to identify the color of the word that appeared on the computer screen by pressing the corresponding button on the response box (red, blue, green, yellow). Each trial ran as follows: an initial fixation point (‘+’) appeared in the centre of the screen for 500 ms. Next, the word was presented in the centre of the screen and remained there until participants made their response. There was a 1000 ms inter-trial interval. Screen background, font style and size was the same as the dot probe task. The ink colors used were red, blue, green, yellow. Participants first completed an initial practice phase consisting of two blocks of 12 trials. The stimuli in the practice phase were 12 number words (e.g., one, thirty, ninety). After the practice block, but before the experimental block, there was also a buffer block of trials, containing four trials of number words. The experimental block consisted of 160 trials where each of the 20 food words (10 healthy and 10 unhealthy) and 20 office words were presented in each of the four colors. The food and office words were blocked such that participants responded to all the food words first or all the office words first, following the recommendation of Cox et al. (2006). For each participant, the order of these blocks was randomized, as was the order of words within block.

Recognition

Participants were told that they would be presented with 80 words; 40 of these words would be from the previous computer tasks and 40 would be new words (Appendix 1). Each word was presented in the centre of the screen until participants decided whether they had previously encountered the word. The screen background, font style and size were the same as the dot probe task. There was a 1000ms inter-trial interval.

EAST

The EAST protocol followed De Houwer (2003). Participants pressed keys on the response box in reaction to words presented on a computer screen. Each key was associated with both a valence and a color, so that one key corresponded to Good or Blue and another key corresponded to Bad or Green. Food words were presented only as colored words. Faster responding was expected when the word’s emotional valence (good or bad) matched the key used for responding to the word's color. For example, suppose a participant has a positive attitude towards food and that she sees a food word in blue. Then, faster responding is anticipated since the response Blue is associated (within the task) with the valence Good. It is assumed that a stimulus partly activates both response dimensions, despite the (sometimes) explicit instructions to focus on color. The association between valence and a response key was reinforced through dozens of practice trials to appropriately valenced, but colorless, stimuli.

Screen background, font style and size were the same as for the dot probe task. Each trial ran as follows: a fixation point ('+') appeared in the screen centre for 500 ms. Next, the stimulus word (white, green or blue) was presented and remained visible until a response was made; for words presented in white the appropriate response was valence (i.e. either good or bad), for words presented in color (the food and office words), the appropriate response was color (green or blue). Errors were indicated by a red cross appearing underneath the word; the cross was visible until a correct response was given. The inter-trial interval was 1000 ms. 'Good' and 'Bad' words were unrelated to the food words (e.g., kindness, rainbow, prison, loneliness).

In one initial practice block participants responded to white words (20 trials in total, of which 10 included clearly positive words and 10 clearly negative words) and in a second practice block participants responded to colored words (20 trials in total, 10 blue words and 10 green words). Subsequently, participants went through 10 buffer trials (two positive white words, two negative white words, three words presented once in blue and once in green). The colored words in the practice and buffer blocks comprised number words (e.g., one, thirteen, sixty). Positive and negative valenced words were selected from Bellezza, Greenwald, and Banaji (1986).

The two experimental blocks followed. Each consisted of 140 trials (randomly ordered), of which 40 were white words (20 positive and 20 negative, from Bellezza et al., 1986) and were presented once. The 20 food words (10 healthy and 10 unhealthy) and 20 office control words were presented once in each color (80 trials in total). Critical trials were those where the (colored) food words were paired with the 'good' key (faster response for positive attitude) and the 'bad' key (faster response for negative attitude). Two versions of the EAST were constructed. For one, the right hand response button was used to categorize positive words and the left hand button negative words (vice versa for the other version).

Attitude towards food

Participants indicated on a 7-point scale the extent to which they believed food was: enjoyable/ unenjoyable, pleasant/ unpleasant, satisfying/ unsatisfying, interesting/ boring. Higher scores indicated a more favorable attitude.

Eating behavior measures.

Participants had their weight and height measured (without shoes) for BMI. The DEBQ consists of 33 items, 13 of which assess emotional eating, 10 external eating, and 10 restrained eating. Each question was answered on a 5-point scale: never; rarely; sometimes; often; very often. DEBQ indices were computed as prescribed in van Strien et al. (1996). For all indices values ranged between one and five, with higher values indicating higher levels of restrained/ external/ emotional eating. The Grand (1968) hunger scale was used to assess hunger at the time of testing. Participants were asked how hungry they were at the time of testing and how much of their favourite food they could eat at that time. Responses were indicated using visual analogue scales anchored by 'not at all hungry' / 'extremely hungry' and 'none at all' / 'as much as I could get' respectively.

Other procedural details

The dot probe and emotional Stroop were completed first, with order of completion being counterbalanced. The next task was the EAST, after which the recognition task followed. The recognition task was placed last, since each of the other cognitive measures could effectively count as additional exposure for the recognition task. The two attention tasks were first since Stroop interference has been shown to increase with prior exposure to the same or related words. Eating behavior measures were last.

Participants were tested individually. They were told that they would complete a number of tasks related to food preoccupation and consumption. After completion of the EAST, participants were given a 10-minute break, during which they were allowed to read magazines (none contained food or office articles/pictures). Subsequently, they completed the recognition task and the rest of the assessments.

Preliminary analyses and computation of measures

General descriptives

Of the 151 participants who took part in the study, 88 were female. The average age of the participants was 18.7 years (sd 0.78 years). Mean age was nearly identical for males and females.

Data cleaning

Outliers in reaction times were defined as any reaction time over 3.5 standard deviations in each reaction time variable. In other words, to compute, e.g., interference on the food Stroop task for a participant, her reaction times for the 80 food words would be averaged and subtracted from the average for the 80 neutral words. Repeating this procedure for each participant, we obtained 151 interference values. It is with respect to these values (and analogous values for the dot probe, EAST task etc.) that outliers were computed. Approximately 3% of the values in each variable were removed.

Computing food Stroop scores

For all the measures we computed scores for healthy and unhealthy foods separately, as well as an aggregate measure.

For each participant, we computed the difference between average reaction time for trials with healthy/ unhealthy food words minus average reaction time for trials with neutral words, to obtain a measure of healthy and unhealthy food interference respectively. Higher (more positive) scores indicated more attentional bias toward healthy or unhealthy foods. An aggregate food Stroop interference score was computed, by averaging the healthy and unhealthy interference scores.

Computing dot probe scores

An attentional bias score for healthy and unhealthy foods was computed by subtracting the mean response times for congruent trials from the mean response times for incongruent corresponding trials. For example, ‘spinach’ was paired with ‘stapler’. For all spinach-stapler trials, we computed the difference between the average reaction time of the trials when the dot replaced stapler minus the average reaction time of the trials when the dot replaced spinach. Thus, positive scores indicated attention toward food cues respectively, with higher scores reflecting a stronger bias for food-related information. The healthy and unhealthy food indices were averaged to create an aggregate index of attentional bias toward food.

Computing EAST scores

EAST scores were computed for the healthy and unhealthy foods by subtracting the mean response time for trials with food words paired with the positive key from the mean response latency for food words paired with the negative key. Higher, positive scores reflected stronger, positive evaluation. Three EAST indices were computed, one for healthy foods, one for unhealthy ones, and an aggregate one.

A problem with the EAST scores computed in this way is that they do not take into account individual variation in a general propensity to consider things as ‘good’ or ‘bad’. Such a general propensity would be evident in positive scores for office words, since such words are supposed to be emotionally neutral (this was confirmed via the valence ratings in the pilot study). Therefore, by analogy with how emotional Stroop scores are computed, we specified the healthy foods EAST measure as the difference between healthy foods EAST scores minus corresponding office words EAST scores, and so on for the unhealthy foods EAST measure and the aggregate foods EAST measure. (Recall that 10 office words were specifically matched to the 10 healthy words, and 10 different office words to the unhealthy ones.)

Computing recognition scores

The number of correctly identified previously seen words (true positives) and incorrectly identified unseen words (false positives) were counted, for the healthy (maximum of 10), unhealthy foods (maximum of 10), and all foods. A recognition score was computed as (true positives + 10 – false positives)/ 2, for the healthy and unhealthy food words; for all the food words the accuracy score was given by (true positives + 20 – false positives)/ 2. By analogy with the way Stroop interference scores are specified, we computed a food memory bias as food word recognition accuracy minus office word recognition accuracy. Note that when computing the unhealthy and healthy word memory indices, we considered the recognition accuracy only for the office words matched to the respective unhealthy and healthy food words. Three memory bias variables were computed (healthy foods, unhealthy foods, aggregate). A positive, high memory bias indicates recognition of food words over and above the neutral ones. The maximum value for the healthy/ unhealthy memory bias variables is 10 (corresponding to perfect recognition of food words, no recognition accuracy of office words). The maximum value for the aggregate memory bias is 20. The range of the three memory bias variables was: -3 to 4 (healthy foods), -2.5 to 4 (unhealthy foods), and -4 to 5.5 (aggregate; negative values indicate that the neutral words were remembered better). Thus, there are no concerns about ceiling values in the recognition task.

Computing scores for attitude towards food

Each participant provided four answers to indicate his/her attitudes towards food, such that higher numerical values indicated a more positive attitude. We created a single index of attitude towards food by averaging the answers to the four questions.

Computing the indices of eating behavior

BMI is computed as the weight of an individual in Kg, divided by his height in m^2. The DEBQ allows the computation of three indices (external, restraint, and emotional eating), which vary between 1 and 5. The answers to the two ‘hunger’ questions were coded onto a 0 to 10 scale and correlated highly with each other (r = .641, p < .0005). We combined these values into a single hunger index by averaging them.

Results

Assessing eating behavior.

We examined whether our participants demonstrated variability in their eating behavior. Figure 1 shows the distribution of BMI scores. Most of the participants are slightly underweight (BMI<20) or normal weight (20<BMI<25), but there were some overweight participants as well (BMI>25; note that Figures 1, 2, 3 include information about the mean and standard deviation of the corresponding variables).

Figure 2 shows that all DEBQ indices are roughly normally distributed and cover the available range of values (1 to 5). A similar conclusion is forthcoming from Figure 3, which shows the distribution of hunger scores. Overall, we examined five determinants of eating behavior and found variability in all five of them. Accordingly, there is no sense in which our participants were too uniform in their eating behavior to possibly confound a study of corresponding cognitive biases.

------------------------------Figures 1, 2, 3------------------------------

Relation between cognitive measures

We examined the relation between initial orientation dot probe, sustained attention dot probe, Stroop, memory biases, and attitude towards food. Tables 1, 2, and 3 present the correlation tables for the versions of the tasks with healthy words, unhealthy words, and all words together (note that the 'attitude towards food' measure does not have healthy/unhealthy versions). The overwhelming impression is that there is very little relation between the cognitive measures. The majority of correlations are not only not significant, but they are very close to zero as well. The handful of correlations that are significant (two out of 45) should be approached with caution, since, given the number of associations we examined, some might be significant simply by chance.

------------------------------Tables 1, 2, 3, 4------------------------------

We next examined the possible moderating influence of gender and the DEBQ indices (cf. O'Connor et al., 2008). It is possible that relations between cognitive biases for food stimuli are evidenced only for high emotional/external/restrained eaters. We trichotomized each of the DEBQ variables into three groups (low, medium, high), so that approximately the same number of cases were assigned to each group. We then examined the relation between the cognitive measures only for the participants in the 'high' subset of each index. The results are shown in Table 4.

Finally, we examined the extent to which the healthy and unhealthy versions of each task correlated with each other. Everything else being equal, we would expect such correlations, so that their presence (or not) would be somewhat suggestive of the validity of the study. There was a correlation between the healthy and unhealthy initial dot probe tasks ($r = .375$, $p < .0005$), the sustained dot probe tasks ($r = .211$, $p = .011$), the Stroop tasks ($r = .609$, $p < .0005$), and the recognition tasks ($r = .243$, $p = .003$).

Discussion

We set out to examine the relation between different measures of cognitive bias related to food information. We included two measures of attention (Stroop, dot probe), a measure of implicit cognition (EAST), a memory measure (a recognition task), and a measure of explicit attitudes. There is a very strong expectation in the literature that these measures would highly correlate with each other. The overall conclusion of our study is that this expectation is wrong; it cannot be assumed that different cognitive bias measures correspond to the same psychological process. Note again that this conclusion applies to cognitive biases related to food stimuli only.

In defense of our conclusion, note first that our sample was approximately 150 participants, exceeding the sample size of many studies on cognitive biases. Even if a relation between cognitive biases could be established with more participants, it is arguable as to how valuable such a relation would be. Second, it is possible that our results were simply too noisy to enable us to identify significant correlations between the cognitive measures. However, this is unlikely to be true, since we did discover several, expected significant correlations (e.g., between the versions of the tasks with healthy and unhealthy words). Third, in analyses we report elsewhere, we did identify expected correlations between cognitive biases and DEBQ indices and BMI and DEBQ indices (cf. Tapper et al., in press); such results also argue against a concern that our data was too noisy to allow an observation of associations between the different cognitive bias measures. Finally, we checked that our population sample displayed considerable variability in terms of their eating behavior. If our participants were too uniform, there would have been no variability in the cognitive measures as well—under such circumstances, we would also have found no relation between the cognitive measures, but not for any interesting reason.

Prescriptively, the measure that appears to have led to the most significant correlations is the dot probe. Moreover, we obtained more significant correlations by considering the healthy versions of the cognitive tasks separately from the unhealthy versions. Accordingly, it appears appropriate to consider healthy food and unhealthy food versions of cognitive tasks separately. Finally, the DEBQ provided a very useful characterization of individual differences in our sample, and it allowed the identification of a range of relations between the cognitive bias measures, not evident otherwise. Specifically, for several subsets of our participants, there was a relation between the initial dot probe and the EAST, and sustained dot probe and the EAST. Why dot probe results correlated primarily with the EAST is an intriguing issue which requires further research. Moreover, there were some significant correlations with Stroop interference and the initial dot probe. This is also an interesting finding, which offers the promise to resolve the debate of what is the exact cognitive process responsible for emotional Stroop interference (cf. Cox et al., 2007, Field, 2006).

In sum, our comparative approach has challenged a widespread assumption in the study of cognitive biases for food-related stimuli and has led to some intriguing questions for further research. We hope that in future studies we might be able to extend these analyses with other relevant types of behavior, such as alcohol abuse.

Acknowledgements

This research was partly supported from grant ReF05/1/010 from the Wales Office of Research and Development (WORD). We would like to thank Miles Cox, Dora Duka, Javad Fadardi, Bob Rafal, and Michelle Taylor for their helpful suggestions.

References

Bellezza, F. S., Greenwald, A. G., & Banaji, M. R. (1986), Words high and low in pleasantness as rated by male and female college students. *Behavior Research Methods, Instruments, and Computers*, *18*, 299-303.

Boon, B., Vogelzang, L. & Jansen, A. (2000). Do restrained eaters show attention toward or away from food, shape and weight stimuli? *European Eating Disorders Review*, *8*, 51-58.

Boyer, M., & Dickerson, M. (2003). Attentional bias and addictive behaviour: Automaticity in a gambling-specific modified Stroop task. *Addiction, 98*(1), 61-70.

Brunstrom, J. M., Rogers, P. J., Pothos, E. M., Calitri, R., & Tapper, K. (in press). Estimating everyday portion size using a ‘method of constant stimuli’: In a student sample, portion size is predicted by gender, dietary behaviour, and hunger, but not BMI. *Appetite*.

Cox, W. M., & Klinger, E. (2004). A motivational model of alcohol use: Determinants of use and change. In W. M. Cox & E. Klinger (Eds.), *Handbook of motivational counseling: Concepts, approaches, and assessment* (pp. 121-138). Chichester, United Kingdom: Wiley.

Cox, W. M., Fadardi, J. S., & Pothos, E. M. (2006). The addiction-Stroop test: Theoretical considerations and procedural recommendations, *Psychological Bulletin*, *132*, 443-476.

Cox, W. M., Pothos, E. M., & Hosier, S. G. (2007). Cognitive-Motivational Predictors of Excessive Drinkers' Success in Changing. *Psychopharmacology, 192,* 499-510

Cox, W. M., Hogan, L. M., Kristian, M. R., & Race, J. H. (2002). Alcohol attentional bias as a predictor of alcohol abusers' treatment outcome. *Drug and Alcohol Dependence, 68*(3), 237-243.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing. A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671-684.

De Houwer, J. (2003). The extrinsic affective simon task. *Experimental Psychology*, *50*, 77-85.

Drayna, D. (2005). Human taste genetics. *Annual Review of Genomics and Human Genetics*, *6,* 217-235.

Duka, T. & Townshend, J. M. (2004). The priming effect of alcohol pre-load on attentional bias to alcohol-related stimuli. *Psychopharmacology*, *176,* 353-361.

Egloff, B. & Hock, M. (2003). Assessing attention allocation toward threat-related stimuli: a comparison of the emotional Stroop task and the attentional probe task. *Personality and Individual Differences*, *35*, 475-483.

Field, M. (2006). Attentional biases in drug abuse and addiction: cognitive mechanisms, causes, consequences, and implications. In: M. R. Munafo & I. P.

Albery (Eds.), *Cognition and Addiction*. pp. 73- 99. Oxford University Press, Oxford.

Field, M., Mogg, K., & Bradley, B. P. (2004). Eye movements to smoking-related cues: effects of nicotine deprivation. *Psychopharmacology, 173,* 116-123.

Francis, J.A, Stewart, S.H., & Hounsell, S. (1997). Dietary restraint and the selective processing of forbidden and nonforbidden food words. *Cognitive Therapy and Research, 21*, 633-646.

Franken, I.H.A. & Muris, P. (2005). Individual differences in reward sensitivity are related to food craving and relative body weight in healthy women. *Appetite, 45,* 198-201.

Franken, I. H., Kroon, L. Y., Wiers, R. W., & Jansen, A. (2000). Selective cognitive processing of drug cues in heroin dependence. *Journal of Psychopharmacology, 14*, 395-400.

Green, M.W. & Rogers, P.J. (1993). Selective attention to food and body shape words in dieters and restrained nondieters. *International Journal of Eating Disorders, 14*, 515-517.

Hogarth, L. & Duka, T. (2006). Human nicotine conditioning requires explicit contingency knowledge: is addictive behaviour cognitively mediated? *Psychopharmacology, 184,* 553-566.

Hogarth, L., Dickinson, A., Hutton, S. B., Elbers, N., &Duka, T. (2006). Drug expectancy is necessary for stimulus control of human attention, instrumental drug-seeking behaviour and subjective pleasure. *Psychopharmacology, 185,* 495-504.

Hogarth, L., Dickinson, A., Wright, A., Kouvaraki, M., & Duka, T. (2007a). The role of drug expectancy in the control of human drug seeking. *Journal of Experimental Psychology: Animal Behaviour Processes*, *33*, 484-496.

Hogarth, L., Dickinson, A., Austin, A., Brown, C., & Duka, T. (2007b). Attention and expectation in human predictive learning: the role of uncertainly. *The Quarterly Journal of Experimental Psychology*.

Johansson, L., Ghaderi, A., & Andersson, G. (2004). The role of sensitivity to external food cues in attentional allocation to food words on dot probe and Strop tasks. *Eating Behaviors*, *5*, 261-271.

Jones, B. T. & Schulze, D. (2000). Alcohol-related words of positive affect are more accessible in social drinkers' memory than other words when sip-primed by alcohol. *Addiction Research, 8,* 221-232.

Logan, G. D. (1998). What is learned during automatization? II. Obligatory encoding of spatial location. *Journal of Experimental Psychology: Human Perception and Performance*. *24*(6), 1720-1736.

MacLeod, C., Mathews, A, & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*, 15-20.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163-203.

McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: an investigation of the emotional Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1595-1607.

Mogg, K. & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behavior Research and Therapy*, *36*, 809-848.

Mogg, K., Bradley, B., & Williams, R. (1995). Attentional bias in anxiety and depression: the role of awareness. *British Journal of Clinical Psychology*, *34*, 17-36.

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? *Behavior Research and Therapy*, *36*, 227-237.

Mogg, K., Bradley, B. P., Field, M., & De Houwer, J. (2003). Eye movements to smoking-related pictures in smokers: relationship between attentional biases and implicit and explicit measures of stimulus valence. *Addiction*, *98,* 825-836.

Mogg, K., Bradley, B. P., Field, M., & De Houwer, J. (2003). Eye movements to smoking-related pictures in smokers: relationship between attentional biases and implicit and explicit measures of stimulus valence. *Addiction*, *98*. 825-836.

Mogg, K., Field, M., & Bradley, B. P. (2005). Attentional and approach biases for smoking cues in smokers: an investigation of competing theoretical views of addiction. *Psychopharmacology*, *180*, 333-341.

Mogg, K., Bradley, B. P., Field, M., & De Houwer, J. (2003). Eye movements to smoking-related pictures in smokers: relationship between attentional biases and implicit and explicit measures of stimulus valence. *Addiction*, *98*, 825-836.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*, 565-584.

O'Connor, D. B., Jones, F., Conner, M., McMillan, B., & Ferguson, E. (2008). Effects of daily hassles and eating style on eating behavior. *Health Psychology, 27*, 20, 31.

Overduin, J., Jansen, A., & Louwerse, E. (1995). Stroop interference and food intake. *International Journal of Eating Disorders*, *18*, 277-285.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 324-354.

Pecina, S., Schulkin, J., & Berridge, K. C. (2006). Nucleus accumbens corticotropin-releasing factor increases cue-triggered motivation for sucrose reward: paradoxical positive incentive effects in stress? *BMC Biology*, *4*, 1-16.

Peretti, C. S. (1998). Anxiété et troubles cognitifs [Anxiety and cognition disorders]. *Encephale*, *24*, 256-259.

Posner, M. I., Rafal., R. D., Choate, L., & Vaughn, J. (1985). Inhibition of return: neural basis and function. *Cognitive Neuropsychology*, *2*, 211-228.

Pothos, E. M. (2007). Theories of Artificial Grammar Learning. *Psychological Bulletin*, *133*, 227-244.

Pothos, E. M., & Cox, W. M. (2002). Cognitive bias for alcohol-related information in inferential processes. *Drug and Alcohol Dependence*, *66*, 235-241.

Rafal, R., Davies, J., & Lauder, J. (2006). Inhibitory tagging at subsequently fixated locations: Generation of “inhibition of return” without saccade inhibition. *Visual Cognition*, *13*, 308-323.

Rather, B. C., Goldman, M. S., Roehrich, L., Brannick, M. (1992). Empirical modeling of an alcohol expectancy memory network using multidimensional scaling. *Journal of Abnormal Psychology*, *101*, 173-183.

Robinson, T. E. & Berridge, K. C. (1993). The neural basis of craving: An incentive-sensitization theory of addiction. *Brain Research Reviews*, *18*, 247-291.

Stacy, A. W. (1997). Memory activation and expectancy as prospective predictors of alcohol and marijuana use. *Journal of Abnormal Psychology*, *106*, 61-73.

Stetter, F., Ackerman, C., Bizer, A., Straube, E. R., & Mann, K. (1995). Effects of disease-related cues in alcoholic inpatients: Results of a controlled "alcohol Stroop" study. *Alcoholism: Clinical and Experimental Research*, *19,* 593-599.

Tapper, K., Pothos, E. M., Fadardi, J. S., & Ziori, E. (in press). Restraint, disinhibition and food-related processing bias. *Appetite*.

Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behavior: Role of automatic and nonautomatic processes. *Psychological Review*, *97*, 147-168.

Tipper, S. P. (1985). The Negative Priming Effect :Inhibitory Priming by Ignored Objects. *The Quarterly Journal of Experimental Psychology*, *37A*, 571-590.

Tipper, S. P. & Cranston, M. (1985). Selective attention and priming: Inhibitory and facilitatory effects of ignored primes. *The Quarterly Journal of Experimental Psychology*, *37A*, 591-611.

Tzelgov, J. (1997). Specifying the relations between automaticity and conscious: A theoretical note. *Consciousness and Cognition*, *6*, 441-451.

Van Strien T. (1997). The concurrent validity of a classification of dieters with low versus high susceptibility toward failure of restraint. *Addictive Behaviors*, *22*, 587-597.

Van Strien, T., Frijters, J.E.R., Bergers, G.P.A. & Defares, P.B. (1986). The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional and external eating behavior. *International Journal of Eating Disorders*, *5*, 295-315.

Waters, H., & Green, M. W. (2003). A demonstration of attentional bias, using a novel dual task paradigm, towards clinically salient material in recovering alcohol abuse patients? *Psychological Medicine*, *33*, 491-498.

Wiers, R. W., Cox, W. M., Field, M., Fadardi, J. S., Palfai, T. P., Schoenmakers, T., & Stacy, A. W. (2006). The search for new ways to change alcohol-related cognitions in heavy drinkers. *Alcoholism: Clinical and Experimental Research*, *30*, 320-331.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*, 3-24.

Appendix 1

The words used in the cognitive measures (Stroop, dot probe, EAST, and the recognition task). Category H indicates an intended healthy food and U an intended unhealthy one. Office words had to be matched both as a group and individually to the food words.

Category	Food Words	Control (Office) Words
H	Spinach	Stapler
H	Apples	Pencil
H	Salad	Paper
H	Lentils	Scanner
H	Vegetables	Calculator
H	Fish	Lamp
H	Carrot	Webcam
H	Peas	Pins
H	Broccoli	Calendar
H	Grapes	Stamps
U	Cake	Desk
U	Sugar	Ruler
U	Doughnut	Scissors
U	Hamburger	Envelopes
U	Chips	Files
U	Caramel	Monitor
U	Crisps	Blinds
U	Kebab	Memos
U	Pizza	Table
U	Cream	Books

Tables

Tables 1. The relationship between different measures of cognitive bias, for the healthy versions of each task. All correlations are marked as either at the .05 level (*) or at the .01 level or better (**).

	Dot probe (initial)	Dot probe (sustained)	Stroop	EAST	Recognition	Attitudes
Dot probe (initial)		.046	.07	-.20*	-.08	.07
Dot probe (sustained)			-.071	0	.064	.04
Stoop				.043	0	.05
EAST					-.033	-.048
Recognition						.002

Tables 2. The relationship between different measures of cognitive bias, for the unhealthy versions of each task. All correlations are marked as either at the .05 level (*) or at the .01 level or better (**).

	Dot probe (initial)	Dot probe (sustained)	Stroop	EAST	Recognition	Attitudes
Dot probe (initial)		.147	-.12	.036	.02	.047
Dot probe (sustained)			.037	.092	.074	-.011
Stoop				-.053	.021	- .016
EAST					.071	.17*
Recognition						-.022

Tables 3. The relationship between different measures of cognitive bias, for the aggregate versions of each measure (that is, the ones including both healthy and unhealthy words). All correlations are marked as either at the .05 level (*) or at the .01 level or better (**).

	Dot probe (initial)	Dot probe (sustained)	Stroop	EAST	Recognition	Attitudes
Dot probe (initial)		.013	.03	-.095	-.064	.073
Dot probe (sustained)			-.033	.083	.042	.016
Stoop				.059	-.014	.021
EAST					.067	.08
Recognition						-.014

Table 4. Correlations between cognitive bias measures for different subsets of our participants. All correlations are marked as either at the .05 level (*) or at the .01 level or better (**). The correlation value is shown below the measures for which a significant correlation was identified.

	Healthy versions	Unhealthy versions	Aggregate versions
Males		sustained dot probe/ EAST	
		r = .27*	
Females	initial dot probe/ EAST		
	r = -.22*		
High emotional eaters	sustained dot probe/ EAST	initial dot probe/ Stroop	Stroop/ recall
	r = .31*	*r* = -.34*	*r* = .35*
High external eaters	initial dot probe/ EAST	initial dot probe/ Stroop	
	r = -.45**	*r* = -.35*	
High restraint eaters	sustained dot probe/ EAST	EAST/ attitudes	sustained dot probe/ EAST
	r = .31*	*r* = .38**	*r* = .33*

Figure 1. The range of BMI values for the participants in the study.

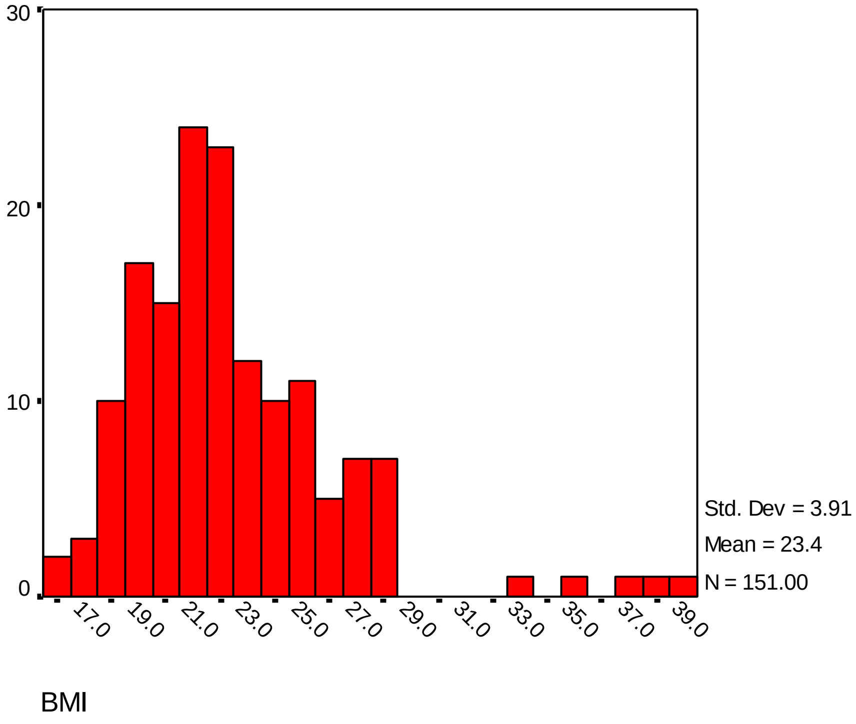
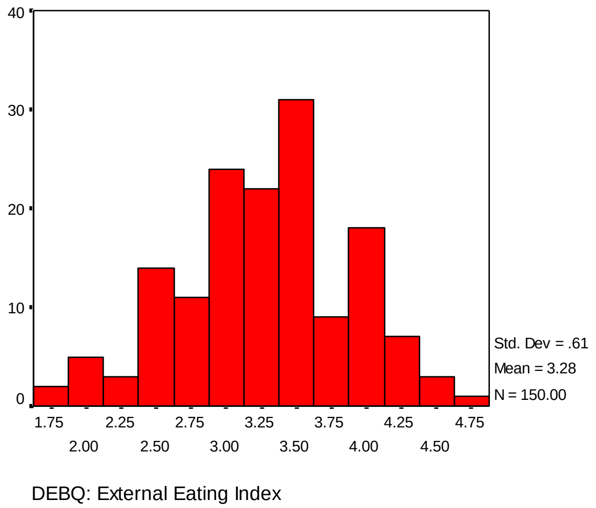
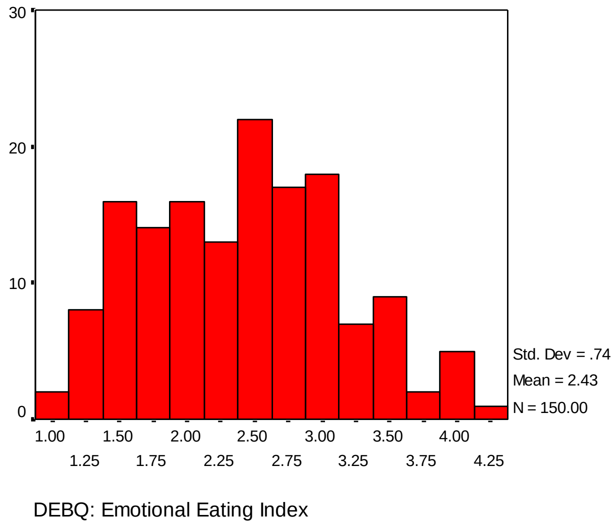


Figure 2. External, restraint, and emotional eating for the participants in our study. Least and greatest possible values for each DEBQ scales are 1 and 5 respectively. For men, the mean external, restraint, and emotional eating values are 2.64, 1.84, and 2.06, respectively. For women, the analogous values are 2.68, 2.49, and 2.06.




30
20
10
0
Std. Dev = .74
Mean = 2.43
N = 150.00
1.00
1.50
2.00
2.50
3.00
3.50
4.00
1.25
1.75
2.25
2.75
3.25
3.75
4.25
DEBQ: Emotional Eating Index

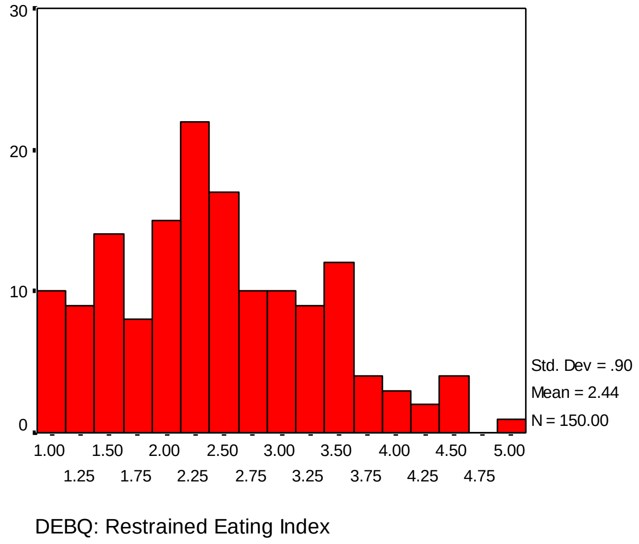

30
20
10
0
1.00 1.50 2.00 2.50 3.00 3.50 4.00 4.50 5.00
1.25 1.75 2.25 2.75 3.25 3.75 4.25 4.75
Std. Dev = .90
Mean = 2.44
N = 150.00
DEBQ: Restrained Eating Index

Figure 3. The distribution of aggregate hunger index values for the participants in the study.