A novel tool to predict food intake: The Visual Meal Creator

Adrian Holliday[a], Chris Batey[a], Dr Frank F. Eves, Dr Andrew K Blannin[a]

[a]School of Sport, Exercise & Rehabilitation Sciences,

University of Birmingham,

Edgbaston,

Birmingham.

B15 2TT

Corresponding Author:

Dr Andrew Blannin

School of Sport, Exercise & Rehabilitation Sciences,

University of Birmingham,

Edgbaston,

Birmingham.

B15 2TT

a.k.blannin@bham.ac.uk

Requests for reprints should be addressed to:

Dr Andrew Blannin, School of Sport & Exercise Sciences, University of Birmingham, Edgbaston,

Birmingham. B15 2TT

a.k.blannin@bham.ac.uk

ABSTRACT

Background: Subjective appetite is commonly measured using an abstract visual analogue scale (VAS) technique, that provides no direct information about desired portion size or food choice, which are both key components of eating behaviour. The purpose of the current investigation was to develop and validate a user-friendly tool – the Visual Meal Creator (VIMEC) - that would allow for independent, repeated measures of subjective appetite and provide a prediction of food intake.

Method: Participants experienced dietary control over a 5-hour period to manipulate hunger state on three occasions (small breakfast (SB) vs. large breakfast (LB) vs. large breakfast + snacks (LB+S)). Appetite measures were obtained every 60 minutes using the VIMEC and VAS. At 4.5 hours, participants were presented with an *ad libitum* test meal, from which energy intake (EI) was measured. The efficacy of the VIMEC was assessed by its ability to detect expected patterns of appetite and its strength as a predictor of energy intake. Day-to-day and test-retest reproducibility were assessed.

Results: Between- and within-condition differences in VAS and VIMEC scores were significantly correlated with one another throughout. Between- and within-condition changes in appetite scores obtained with the VIMEC exhibited a stronger correlation with EI at the test meal than those obtained with VAS. Pearson correlation coefficients for within-condition comparisons were 0.951, 0.914 and 0.875 (all $p < 0.001$) for SB, LB and LB+S respectively. Correlation coefficients for between-condition differences in VIMEC and EI were 0.273, 0.904 ($p < 0.001$) and 0.575 ($p < 0.05$) for SB – LB+S, SB – LB and LB – LB+S respectively. The VIMEC exhibited a similar degree of reproducibility to VAS.

Conclusion: The VIMEC appears to be a stronger predictor of energy intake and may prove to be a more preferable measure of subjective appetite than VAS.

KEYWORDS: Appetite, eating behaviour, food photography, energy intake

BACKGROUND

Methodological issues associated with measures of appetite persist. Self-report questionnaires and scales are commonly used for the measure of subjective appetite, with the visual analogue scale method (VAS) being the most prevalent within appetite research [1]. Such measures are inexpensive and both quick and simple to administer. While validity is not easily assessed, it is generally considered that the VAS is a valid measure of subjective appetite [1-3], demonstrating sensitivity to manipulation, especially when used to address within-subject comparisons [4]. The VAS's ability to predict eating behaviour is less clear. While some studies have demonstrated a significant correlation between VAS scores and subsequent aspects of eating behaviour [4, 5], others have shown a lack of a relationship [5, 6]. Test-retest reproducibility has been shown to be good [3, 4], but day-to-day reproducibility is considerably weaker [4, 6]. There are limitations associated with the use of the VAS method; notably the abstract nature of the question and line format and the difficulty in conceptualising the constructs of "hunger" and "fullness," and the lack of any indication of desired portion size, food choice or food-specific desires. Hence, important aspects of eating behaviour are neither assessed nor predicted when using this method of measurement.

Subjective appetite measures are usually used in conjunction with more objective measures, typically in the form of food intake measures. Commonly, *ad libitum* energy intake is measured from buffet-style or constant composition test meals. The *ad libitum* food intake method has been shown to exhibit a high degree of day-to-day reproducibility, both when presented as a buffet [7] and when the meal is of a constant composition [8]. While allowing for a valid quantitative measure of objective appetite, food intake in test meals can be influenced by a number of external factors, such as the amount of food presented [9, 10], the variety of foods available [11] and the perceived palatability of the food [12]. In addition, buffet-style presentation and a laboratory setting are not habitual eating environments for the majority of people and may influence intake [1, 11, 13]. Such external cues are potent stimuli for appetite regulation and can override physiological determinants of hunger. A key limitation of the *ad libitum* intake method is that it does not allow for independent, repeated measures within a short space of time, in contrast to VAS, which can be repeated frequently to track acute chances in appetite. Any intake measure will have a large impact upon subsequent measures and, while total or mean intake

values can be calculated over a study period, each separate intake or eating episode will not be independent from previous measures. From a practical viewpoint, the *ad libitum* intake method can be expensive, time-consuming and labour-intensive to administer and can result in large amounts of food wastage.

The aim of the current study was two-fold. Firstly, to develop a novel tool for the measure of subjective appetite that could potentially rectify the short-comings of the VAS and *ad libitum* intake methods of assessment. We aimed to provide a less abstract subjective measure than VAS, using a portion selection method, while also allowing for indicative measures of food choice and enabling independent, repeated measures in a cost-effective and time-efficient manner. Secondly, we aimed to address the validity and reproducibility of the tool – the Visual Meal Creator (VIMEC) – relative to both the VAS and *ad libitum* intake methods.

MATERIALS AND METHODS

Participants: Twelve recreationally active participants (8 female, 4 male; mean age 23 ± 2 years; mean body mass 70.4 ± 17.3 kg; mean BMI 22.8 ± 3.6 kg•m^{-2}) were recruited from the School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham. Those suffering from illness such as cold or flu, those taking medication that was likely to affect appetite or that needed to be taken with food more frequently than once a day, those with food allergies and those suffering from diabetes were excluded from taking part. Ethical approval was obtained from the Ethics Committee of the University of Birmingham.

Study design: The validity of any form of rating scale is not easily addressed. To attempt this, we used the three assessments as highlighted by Stubbs *et al* [2]. These are a) the apparent validity of the measure in terms of its ability to predict the behaviour which is being assessed, which was assessed by comparing the VIMEC score with an *ad libitum* test meal energy intake; b) the change in rating score under conditions where it should change if sensitive, with changes compared with those seen with a valid, commonly-used technique for the measure of subjective appetite – the visual analogue scale (VAS) test and c) the reproducibility of the measures, which was assessed by comparing day-to-day measures and short-term test-retest measures.

A within-subject, randomised crossover study design was utilised. Participants were randomly assigned to each of the three experimental conditions; small breakfast (SB), large breakfast (LB) and large breakfast with snacks (LB+S). These feeding conditions were used to manipulate hunger state.

Procedure & protocol: Participants arrived at the Exercise Metabolism Laboratory within the School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham between 07.00 and 09.00, after a ten-hour overnight fast. Upon arrival at the laboratory for the first time, participants were provided with further verbal information regarding the nature of the study and given the opportunity to ask any questions regarding their participation. A written consent form was then signed. Health questionnaires were completed and breakfast food selections were made. Participants were then randomly allocated to one of the three trial conditions. The study protocol for each condition is shown in diagrammatical form in **figure 1**. Participants remained sedentary throughout the trial period.

Prior to the first trial, participants were provided with a food diary and instructed on how to complete it. They were asked to complete the diary on the day before their first trial and instructed to replicate this diet on the day prior to the following two trials.

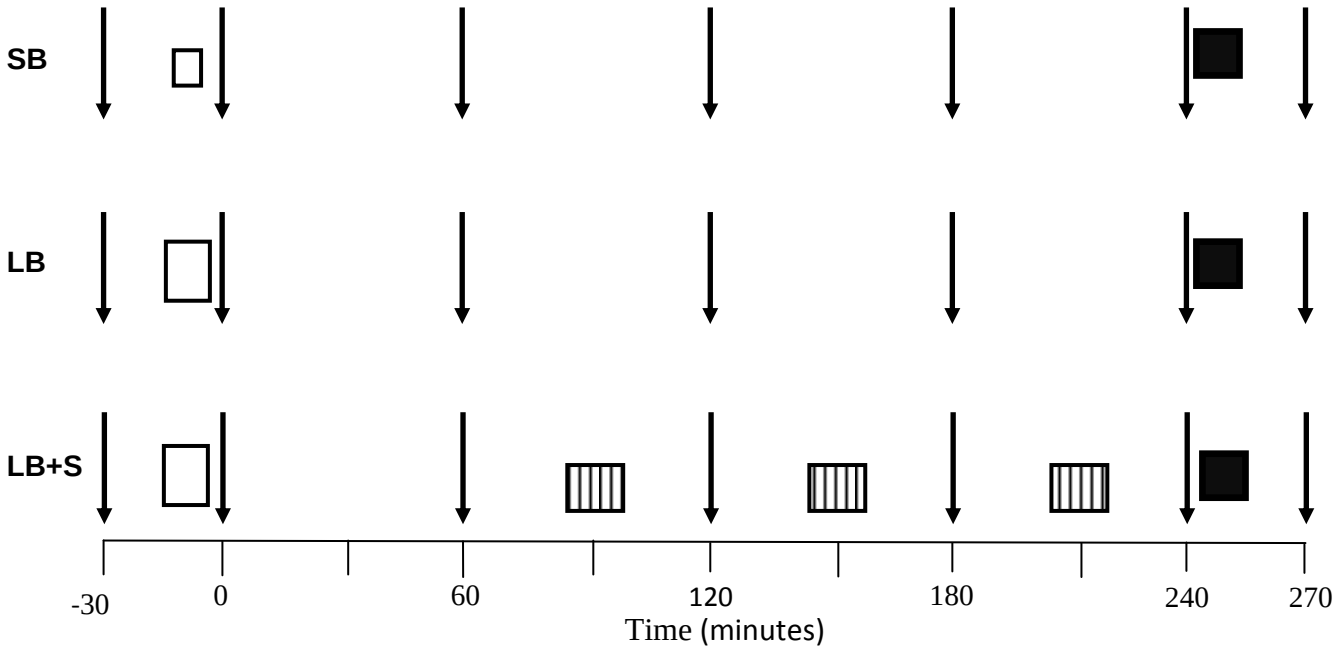


Figure 1. Study protocol. Arrow = appetite measure (VAS and VIMEC). Small clear rectangle = small breakfast. Large clear rectangle = large breakfast. Striped rectangle = snack. Black square = *ad libitum* lunch meal

Breakfast meals and snack: The small breakfast meal consisted of a 25g cereal bar (oat and raisin, Sainsbury's) with 200ml of pure orange or apple juice (Sainsbury's), exhibiting the following characteristics: ~140 kcal, 27g carbohydrate, 2.3g fat, 1.8g protein, 1.1g fibre. The large breakfast consisted of cereal (80g of Original Swiss-style Alpen, or 55g of Kellogg's Bran Flakes); 125-150ml of semi-skimmed milk (Sainsbury's); 2 slices of toast (Kingsmill 50/50 thick slice, ~88g); 16 g of margarine (Flora light) and 30g of jam (strawberry, Sainsbury's) with 200ml of pure orange or apple juice (Sainsbury's). A choice of two cereals, with similar energy density and macronutrient content was provided to allow for individual preferences and dislikes. However, when Bran Flakes were selected, a banana was added to the meal in order for energy content to be similar between the two options, accounting for the smaller portion of Bran Flakes. The same cereal was consumed for both large

breakfast conditions. The large breakfast meal (Alpen cereal chosen) typically exhibited the following characteristics: ~763 kcal, 133.4g carbohydrate, 15.2g fat, 22.5g protein, 10.6g fibre. The meals were consumed within 15 minutes.

Three snack items were administered at 1.5, 2.5 and 3.5 hours in the LB+S condition. These were, in order, a 50g flapjack bar (Sainsbury's, 223 kcal); a 25g cereal bar (oat and raisin, Sainsbury's, 98 kcal) and a ~152g, medium sized banana (Sainsbury's, ~98 kcal). This provided an additional 419 kcal, 67.1g carbohydrate, 13.9g fat, 5.4g protein, 6.6g fibre, resulting in a total intake in the LB+S of approximately 1182 kcal, 200.6g carbohydrate, 29.2g fat, 27.9g protein, 17.2g fibre. This compared with a total energy intake of ~763 kcal in the LB condition and ~140 kcal in the SB condition.

Measures: Subjective appetite was measured using the VIMEC and the widely used visual analogue scale technique (VAS). The VIMEC is a computer programme test, designed and developed in the School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, in which the participant is asked to construct a computerised visual meal from an extensive menu, represented by a library of food images. The participant is asked to select the foods that they would opt to consume, should they eat a meal or snack at this moment in time. Selecting no food is an option available. The participant is presented with a screen exhibiting the food items available (see **figure 2a**). The participant is free to select up to a maximum of four "main meal" items (from a selection of 17), which can be displayed on the meal plate, along with any number of "snack or dessert" items, which are selected individually and displayed separately. Once selections are made, the participant is then presented with a screen consisting of a meal plate on which their selected food items appear. The portion size of each item can then be manipulated individually using sliding bar scales (see **figure 2b**). The number of images for each food item varies, depending on the nature of the item, typical portion sizes and the number of food items selected. Typically, however, this number ranges from 10-40 images per food item, allowing for a high resolution. This process is then repeated separately for any "snack or dessert" items selected. Typically, this task took between 30 seconds and 2 minutes to complete.

Once the computerised meal was fully constructed, the meal was saved. The results were analysed and the investigator was able to see which portion size was selected and what number photograph this selection corresponded to. All food images were of a known weight and the food

209 characteristics of each food item were recorded (energy density, macronutrient content). Hence, it was
210 then possible to calculate the nutritional content of the meal.
211 Subjective appetite was also assessed using the 4-question, 150mm-line VAS test for subjective
212 appetite, addressing "hunger", "fullness", "desire to eat" and "expected food intake" [14]. A composite
213 VAS test score was calculated (hunger score + desire score + expected intake score + (150-fullness
214 score)). This single score was used for the ease of data analysis and presentation. With the original 6
215 question VAS technique of Hill & Blundell [14], the scores for each question co-vary to a large extent
216 [2] and the first principle component of the questions is the mean value of the scores [15]. The two
217 appetite measures, VIMEC and VAS, were completed in a counterbalanced order to partition out effects
218 of order.
219 Energy intake was measured with the use of an *ad libitum* test meal. The content of this test
220 meal was dependant on the food choices made by the participant when using the VIMEC. The food
221 items selected during the measure obtained 60 minutes before the test meal (t=180) were presented for
222 the test meal. At the appetite measure obtained immediately prior to the test meal, the participant was
223 asked not to select any new, additional food items. Participants were accompanied to the Research
224 Kitchen within the School of Sport, Exercise and Rehabilitation Sciences, where they were provided
225 with a dinner plate and a bowl at a table. The food items of the test meal were presented buffet-style on
226 a separate work surface, and of portion-size similar to that of the largest portion available on the
227 VIMEC tool. Participants were instructed to serve the food that they desired to eat from the buffet on to
228 the plate or into the bowl and return to the table to eat. They were informed that they could return for
229 further servings and that more of each food item was available. They were instructed to eat until they
230 felt satisfyingly full. Covertly, each food item presented was weighed prior to the meal commencing
231 and again at the cessation of eating, with the difference between the two indicating the amount
232 consumed. Subtracted from this was food left remaining on the plate or in the bowl, which was also
233 weighed after the meal. Energy density of all food was known, allowing for the calculation of energy
234 intake.
235 The reproducibility of VIMEC and VAS was assessed by comparing day-to-day and test-retest
236 reproducibility. Day-to-day comparisons were made between the first measure obtained, prior to the

breakfast meal, for each condition. A second comparison was made between the second, post-breakfast measures obtained in the LB and LB+S conditions, as the same breakfast was consumed in each condition. One appetite measure was randomly selected for each participant for a retest measure. In this instance, participants were asked to repeat the measure within 2-3 minutes of the initial measure. These comparisons were made for both the VAS and VIMEC techniques, hence allowing for between-measure comparisons, as well as within-measure comparisons.

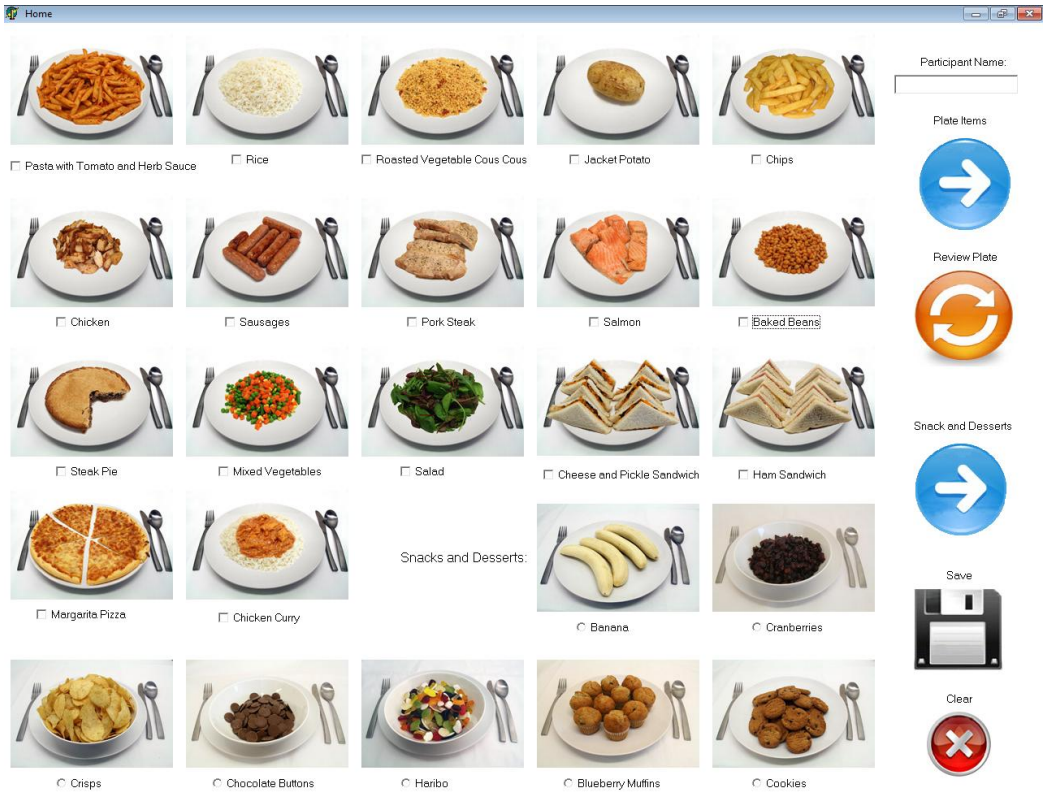
Statistical analysis: The mean energy intake values of the test meal for each condition were compared using a one-way, repeated measures ANOVA. To test for sensitivity to change in appetite, appetite scores from the VIMEC and the VAS were both assessed using a 3 (condition: SB, LB, LB+S) x 7 (time: -30, 0, 60, 120, 180, 240, 300) factorial, repeated-measures ANOVA. Significant main effects and interactions from ANOVA were further assessed by pairwise comparisons using Bonferroni post-hoc analysis. VIMEC appetite scores were also compared with VAS test scores, using Pearson product moment correlation analysis, for all measures obtained within each condition, separately. This was also conducted for between-condition, within-subject comparison, by assessing percentage difference between the conditions (SB – LB+S, SB – LB and LB – LB+S). This approach allows for comparisons of the ability to detect inter-subject changes in appetite.

To assess the ability of the VIMEC to predict between-subject differences in energy intake, appetite scores obtained immediately prior to the test meal were compared with energy intake at the test meal .To assess the ability of the VIMEC to predict within-subject differences in energy intake, between-condition percentage difference (SB – LB+S, SB – LB and LB – LB+S) for energy intake, VAS score and VIMEC score was calculated and these differences were compared using correlation analysis. Differences in correlation coefficients were assessed using t-tests for non-independent correlation coefficients.

Day-to-day measures were compared using a one-way, repeated measures ANOVA (pre-breakfast measures, SB vs. LB vs. LB+S) and a paired samples t-test (post-breakfast measures, LB vs. LB+S). Test-retest measures were compared using a paired samples t-test. The coefficient of variation was calculated for all reproducibility measures, with these coefficient of variation values for the VIMEC and VAS methods compared using paired samples t-tests. A statistical significance level of $p <$

0.05 was used throughout. All statistical analysis was carried out using the SPSS software programme (SPSS inc. Chicago, Illinois, USA).

(a)



(b)

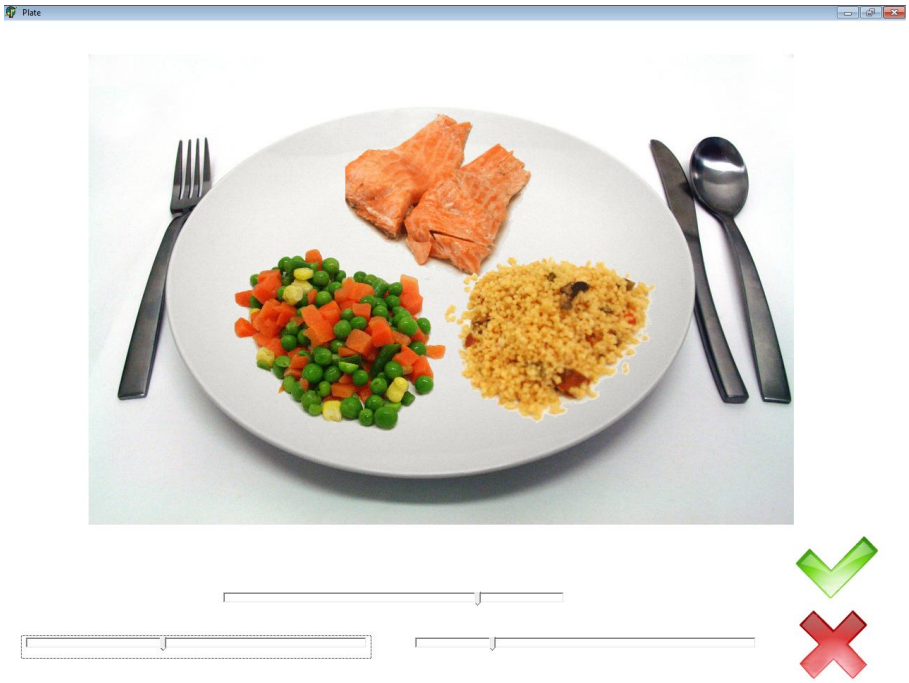


Figure 2. The Visual Meal Creator. The menu screen (a) and an example meal (b). Portion size of each item in the meal can be manipulated using the sliding bar scales.

RESULTS

Energy intake at the test meal: Mean energy intake values at the test meal, for each of the three trial conditions are shown in **figure 3**. A significant condition effect was observed for mean energy intake ($F(2,22) = 8.253$, $p = 0.002$). Pairwise comparisons demonstrated that the mean intake in the LB+S (404 ± 255 kcal) was significantly lower than both mean LB intake (675 ± 313 kcal, $p = 0.003$) and SB intake (786 ± 519 kcal, $p = 0.02$), which did not differ.

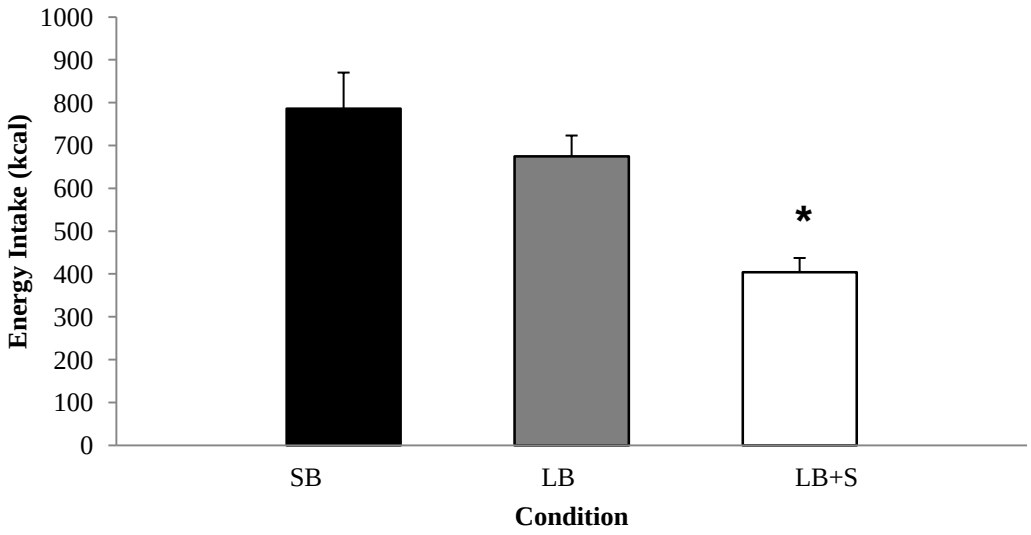


Figure 3. Mean energy intake values (± SEM) for the SB, LB and LB+S conditions.
* = significantly different to LB and SB.

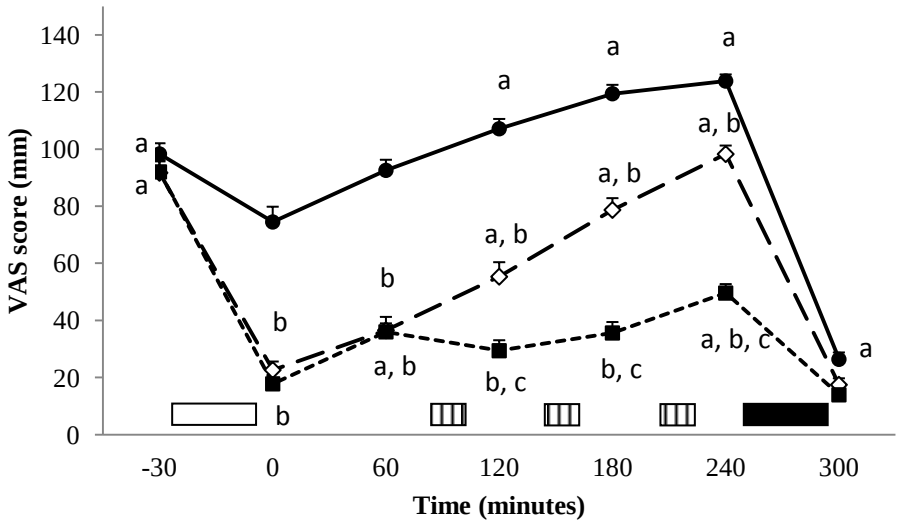
Subjective appetite scores

VAS: Changes in appetite scores, obtained with the VAS measure, over the trial periods, for each of the three trial conditions are shown in **figure 4a**. A factorial, repeated measures ANOVA demonstrated a significant condition x time interaction effect ($F(12,132) = 21.039$, $p < 0.001$). Post-hoc pairwise comparisons showed significant within- and between-subject differences, as illustrated in **figure 4a**.

VIMEC: Changes in appetite scores, obtained with the VIMEC, over the trial periods, for each of the three trial conditions are shown in **figure 4b**. A factorial, repeated measures ANOVA demonstrated a significant condition x time interaction effect ($F(12,132) = 6.973$, $p < 0.001$). Pairwise

comparisons highlighted significant within- and between condition differences. These are shown in **figure 4b**.

(a)



(b)

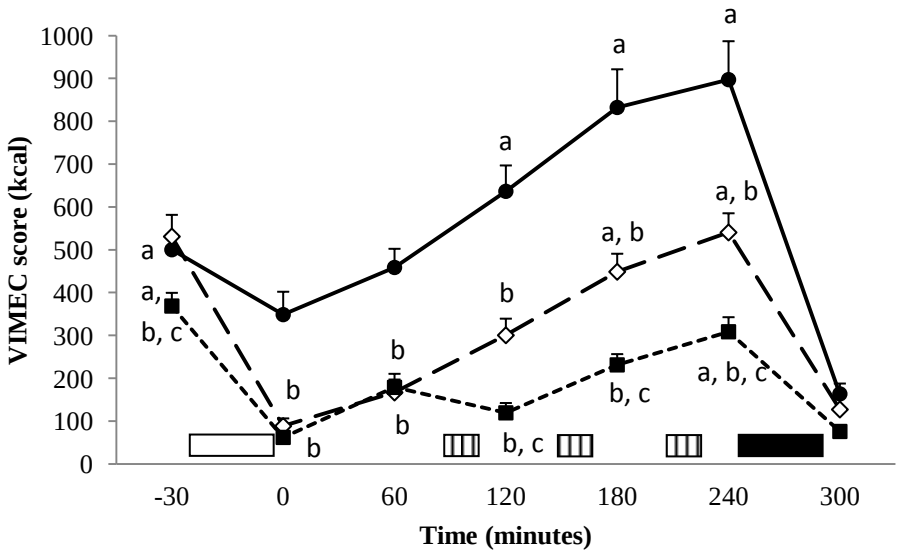


Figure 4. Appetite profiles for the SB, LB and LB+S conditions for (a) VAS and (b) VIMEC methods. Values are means ± SEM. SB (●), LB (◊) and LB+S (■) conditions. Solid line indicates SB, dashed line indicates LB, dotted line indicates LB+S. Hollow rectangle = breakfast meal. Vertical lined rectangles = snacks. Solid black rectangle = *ad libitum* lunch meal. a = within-condition effect, significantly different to t=0. b = between-condition effect, significantly different to SB. c = between-condition effect, significantly different to LB.

Correlation between subjective appetite scores obtained with VAS and VIMEC

Between-subject, within-condition correlations for VAS scores and VIMEC scores were of moderate-strength to strong and statistically significant in each condition (SB, $r = 0.656$, $p < 0.001$; LB, $r = 0.813$, $p < 0.001$; LB+S, $r = 0.673$, $p < 0.001$).

Within-subject, between-condition correlations for percentage difference in VAS and VIMEC scores were also statistically significant, demonstrating moderate-strength correlation (SB – LB+S, $r = 0.570$; SB – LB, $r = 0.526$; LB – LB+S, $r = 0.503$, all $p < 0.001$).

Ability to predict between-subject differences in energy intake

VAS: Correlation between VAS scores immediately prior to the lunch test meal and EI at the test meal for each of the three trial conditions are shown in **table 1**. Pearson correlation coefficients were significant for LB but not for SB or LB+S.

VIMEC: Correlation between VIMEC scores immediately prior to the lunch test meal and EI at the test meal for each of the three trial conditions revealed significant correlation coefficients for all conditions (**table 3.1**).

T-tests for non-independent correlation coefficients revealed that, for each condition, the correlation coefficient for VIMEC vs. EI was significantly greater than for VAS vs. EI (all p values < 0.01).

	VAS vs. EI	**VIMEC vs. EI**
SB	0.548 (p=0.065)	0.951 (p<0.001) *
LB	0.632 (p=0.027)	0.914 (p<0.001) *
LB+S	0.401 (p=0.196)	0.875 (p<0.001) *

Table 1. Pearson product moment correlation coefficients for VAS scores vs. EI and VIMEC scores vs. EI. * = significantly greater than VAS vs. EI, $p < 0.01$. ** = significantly greater than VAS vs. EI, $p < 0.001$.

Ability to predict within-subject differences in energy intake

VAS: Percentage differences between each of the three conditions, for both EI and VAS were calculated. The correlation between percentage difference in EI and percentage difference in VAS scores proved to be weak (**table 2**).

VIMEC: Correlation between percentage differences in EI and percentage differences in VIMEC score immediately prior to the test meal across the three conditions proved strong (**table 2**). The correlation coefficients for two of the three comparisons (SB – LB and LB – LB+S) were statistically significant, exhibiting moderate-strength positive correlation ($r = 0.525$, $p = 0.04$) and very strong positive correlation ($r = 0.940$, $p < 0.001$) respectively.

Comparisons of EI vs. VAS correlation with EI vs. VIMEC correlations showed that the correlation between EI and VIMEC was significantly stronger for the SB – LB difference.

	SB - LB+S	**SB - LB**	**LB - LB+S**
EI vs. VAS	0.063 (p = 0.423)	-0.016 (p = 0.480)	0.011 (p = 0.193)
EI vs. VIMEC	0.273 (p = 0.195)	0.940 (p < 0.001)**	0.525 (p = 0.04)

Table 2. Product moment correlation coefficients for comparison of differences in EI with differences in VAS score and differences in VIMEC score between the three trial conditions. * = EI vs. VIMEC correlation significantly greater than EI vs. VAS correlation, $p < 0.05$, ** = EI vs. VIMEC correlation significantly greater than EI vs. VAS correlation, $p < 0.01$

Reliability and reproducibility of subjective appetite measures

Day-to-day measures: Comparisons of measures at t=-30 (baseline) showed that there were no significant differences between measures for VAS. There was a significant condition effect for VIMEC scores at baseline ($F(1) = 11.63$, $p = 0.006$), with post hoc analysis demonstrating that VIMEC scores were lower in the LB+S condition (369 ± 214 kcal), compared with both the SB (500 ± 251 kcal, $p = 0.017$) and LB (531 ± 351 kcal, $p = 0.047$) conditions. There were no differences between measures obtained at t=0 for the LB and LB+S conditions for either subjective appetite method. Between-measure comparisons of the coefficient of variation (CV) for t=-30 measures (SB vs. LB vs. LB+S) and CV for mean appetite scores for measures obtained at t=-30, t=0 and t=60 (LB vs. LB+S) were conducted. The mean CV value for VAS measures at t=-30 was significantly lower than that for

VIMEC measures (19.1 ± 11.7% vs. 32.2 ± 15.4%, p = 0.033). There was no significant difference in mean CV values for meaned VAS and VIMEC measures obtained at t=-30, t=0 and t=60 (23.3 ± 12.1% for VAS and 25.6 ± 21.4% for VIMEC, p = 0.754).

Test-retest measures: Paired sample T-tests comparing the test-retest scores showed that retest measures were similar to initial measures for both VAS and VIMEC methods. Mean CV values were small and did not differ between the two methods (6.0 ± 6.1% vs. 5.7 ± 6.2% for VAS and VIMEC respectively).

DISCUSSION

The aim of the current study was to assess the validity and reliability of the Visual Meal Creator (VIMEC) as a method for measuring subjective appetite. The VIMEC demonstrated the ability to detect expected changes in subjective appetite, as shown by the appetite profiles. By time point t=240, immediately prior to the lunch test meal, the appetite scores were significantly different between each trial condition. This was reflected by a significant trial condition effect for energy intake at the lunch test meal, although it should be noted that intakes in the SB and LB conditions, while differing by 14%, were not significantly different. The appetite profile for VIMEC measures was almost identical to the profile obtained from using the VAS method – a valid, reliable and highly-used method for the measure of subjective appetite. Between-subject, within-condition comparisons of VIMEC and VAS scores demonstrated significant, moderate-strength to strong correlation. Further, between-condition percentage difference for VIMEC and VAS scores demonstrated a moderate-strength relationship. While proving validity for such measures is difficult, this comparison suggests that the VIMEC was performing as intended: providing a quantitative measure of subjective appetite and detecting changes in subjective appetite after dietary manipulation.

The VIMEC showed potential as a predictor of eating behaviour, of which the lunch test meal energy intake acted as a proxy. Correlations between VIMEC scores immediately prior to the test meal and the energy intake values were very strong for each of the three conditions and compared favourably with those for VAS and energy intake, indicating that the VIMEC is a stronger predictor of between-subject differences in energy intake. The correlation for within-subject, between-condition differences in EI and differences in VIMEC scores immediately prior to the test meal was significant and of moderate-strength to strong in two of the three comparisons. This relationship was stronger than that of differences in VAS score and differences in EI for all three comparisons, proving significantly so in one of these cases.

Within appetite research, when the effect of an intervention upon appetite is under investigation, VAS is commonly used in conjunction with objective appetite measures, such as circulating levels of appetite-associated hormones or a measure of eating behaviour, such as *ad libitum* energy intake. In these instances, the correlation between VAS scores and these objective or

446 behavioural measures are rarely assessed, so direct evidence of VAS's strength as a predictor of eating
447 behaviour in such circumstances is not abundant. Nevertheless, it is generally considered that VAS
448 exhibits good predictive strength when more severe interventions are implemented (pharmacological),
449 but when more subtle interventions are in place, such as exercise, the reliability of VAS to predict
450 eating behaviour is poor [5, 16, 17]. The intervention in the current study was achieved by controlling
451 food intake at breakfast and for the following four hours until lunch in an attempt to manipulate
452 appetite. Under these circumstances, the VIMEC proved a strong predictor of eating behaviour. It
453 remains to be seen whether the VIMEC will prove a strong predictor of eating behaviour within
454 exercise intervention studies.

455 The correlation coefficients for between-subject, within-condition comparisons of VIMEC
456 score and EI in the present study were extremely high. It is possible that the study design contributed.
457 The food items selected at time point t=180 were the items that were presented at the buffet meal. This
458 measure was obtained 60 minutes prior to the lunch test meal, allowing sufficient time for food to be
459 prepared. At t=240, immediately prior to the meal, food item selection for the VIMEC was restricted to
460 those items selected at t=180. This ensured that the items selected here were those that the participant
461 would be presented with at the lunch test meal, allowing a strong comparison of the amount of each
462 item selected. It was possible that the number of food items selected (and hence made available at the
463 lunch test meal) could have constrained the subsequent energy intake. As a result, the magnitude of
464 correlation could have been artificially inflated, as food variety has been shown to influence energy
465 intake at a meal [11]. Therefore, partial correlations were calculated to remove the influence of the
466 number of food items on the energy intake of the test meal. These partial correlations differed
467 minimally from the original correlation coefficients (SB: 0.930 vs. 0.951; LB: 0.934 vs. 0.914; LB+S:
468 0.870 vs. 0.875). Hence, it would appear that the number of food items selected was not a strong
469 predictor of energy intake in this study and did not contribute to the very strong correlation observed
470 between VIMEC score and energy intake.

471 Stubbs (2000) highlighted the large between-subject variability in subjective appetite measures
472 when using VAS and recommended that the method was therefore more appropriate for within-subject
473 comparisons. Large between-subject variability is not uncommon with appetite measures, including *ad*

libitum test meal intakes [18] due to large biological variation in appetite, food preference and eating behaviour. The between-subject variability of the VIMEC scores immediately prior to the lunch test meal, was large (coefficient of variation (CV) values for 69%, 60% and 76% for SB, LB and LB+S respectively), although not vastly larger than the variability in the energy intake measures (CV values of 66%, 46% and 60% for SB, LB and LB+S respectively). Therefore, as with a number of other subjective appetite and eating behaviour measures, the VIMEC is likely to be best suited to within-subject comparisons and repeated-measure study designs

The VAS has previously been shown to exhibit good test-retest reproducibility [3, 4], but considerably poorer day-to-day reproducibility [3, 4, 6]. The results of the current study would suggest that the VIMEC exhibits a similar degree of test-retest reproducibility, with CV values very comparable to those observed with VAS. While the CV for day-to-day repeated measures at t=-30 was significantly higher than VAS, suggesting poorer day-to-day reproducibility when using the VIMEC, the mean VIMEC scores for measure t=-30, t=0 and t=60 for the LB and LB+S conditions were similar, suggesting a similar degree of reproducibility. One would perhaps have expected a greater degree of variation with the VIMEC, due to the option of choosing different food items of different energy densities. This large degree of choice, allowing for the selection of vastly different meal creations would lend itself to large variations in the measure. It should also be noted that, with any variation in day-to-day measure, it is difficult to disentangle the contribution of biological and methodological variation, especially when obtaining subjective measures, prone to variation [4].

There were, however, significant differences in baseline values for VIMEC, between the LB+S condition and LB condition and between LB+S and SB. This questions the day-to-day reproducibility. Inspection of the data would suggest that this was not driven by a single or small number of outliers. Other than the previously mentioned large degree of choice and consequent increased likelihood of variability, and the biological variation in appetite sensations from day-to-day, it is difficult to explain this observation. Food intake during the 24 hours prior to each trial was controlled by asking participants to record their dietary intake on the day prior to their first trial, then asking them to repeat this intake on the day before subsequent trials. It is possible that this was not well adhered to and that differences in dietary intake on the day prior to trials may have influenced baseline appetite measures.

When obtaining a subjective appetite measure using the VIMEC, there is an upper limit to the portion size available. This maximum portion is dependent on the food item and, for the main meal items, the number of food items selected. To alleviate this limitation, participants were informed that, should they desire more than the upper limit, they could save the current measure, clear the screen and complete a second measure for any additional food desired. While this option is not ideal, with the participant unable to visualise their entire meal creation, it does allow for unlimited portion size selection. In the present study, no participant chose to complete a second test for any measure. In addition, the 252 measures obtained in total resulted in 564 different food item selections. Only 31 times (5.5%) were maximum portions selected (15 x salad). In addition, 26 of these 31 maximum portion selections occurred during instances where the participant selected 4 or more food items in the measure, when space on the plate for individual food item portions was limited. We are therefore confident that the VIMEC does not substantially restrict the upper limit of a subjective appetite measure.

While the use of photographic images of food is not a new concept within the area of appetite research, the VIMEC is, to our knowledge, the first subjective appetite tool that allows the user to create a whole meal. Similar tools have asked users to select a desired portion size of a range of individual food items [19] or a mixture of individual items and ready-made meals [20], showing potential as useful appetite measures. However, in neither of these studies did the technique demonstrate a relationship between desired portion size and *ad libitum* food intake. The progressive step evident with the VIMEC, allowing for the creation of a meal from an extensive menu of food items allows a stronger measure of food choice and preference that is limited with the aforementioned format of other tools. It is also possible that the more sophisticated nature of the VIMEC allows for a stronger prediction of feeding behaviour, as is supported by the findings of the current study.

CONCLUSION

In conclusion, the Visual Meal Creator would appear to be a strong predictor of between- and within- subject differences in energy intake. Test-retest reproducibility was good. Day-to-day reproducibility was quite large, but this may be due to the large degree of food choice allowable with the VIMEC. In comparison with the VAS technique, the VIMEC proved equally as proficient at detecting expected changes in subjective appetite, while exhibiting a similar degree of reproducibility. The VIMEC was shown to be a significantly stronger predictor of energy intake – a fundamental aspect of eating behaviour. Therefore, the VIMEC may prove a preferable tool for the measurement of subjective appetite, due to its strength as a predictor of eating behaviour.

REFERENCES

1. Blundell, J., De Graaf, C., Hulshof, T., Jebb, S., Livingstone, B., Lluch, A., et al., *Appetite control: methodological aspects of the evaluation of foods.* Obesity Reviews, 2010. **11**(3): p. 251-270.

2. Stubbs, R.J., Hughes, D.A., Johnstone, A.M., Rowley, E., Reid, C., Elia, M.,et al., *The use of visual analogue scales to assess motivation to eat in human subjects: a review of their reliability and validity with an evaluation of new hand-held computerized systems for temporal tracking of appetite ratings.* British Journal of Nutrition, 2000. **84**(4): p. 405-15.

3. Stratton, R.J., Stubbs, R.J., Hughes, D., King, N., Blundell, J.E., Elia, M. *Comparison of the traditional paper visual analogue scale questionnaire with an Apple Newton electronic appetite rating system (EARS) in free living subjects feeding ad libitum.* European Journal of Clinical Nutrition, 1998. **52**(10): p. 737-41.

4. Flint, A., Raben, A., Blundell, J.E., and Astrup, A. *Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies.* International journal of obesity and related metabolic disorders : Journal of the International Association for the Study of Obesity, 2000. **24**(1): p. 38-48.

5. Parker, B.A., Sturm, K., MacIntosh, C.G., Feinle, C., Horowitz., and Chapman, I.M. *Relation between food intake and visual analogue scale ratings of appetite and other sensations in healthy older and young subjects.* European Journal of Clinical Nutrition, 2004. **58**(2): p. 212-218.

6. Raben, A., Tagliabue, A., and Astrup, A. *The reproducibility of subjective appetite scores.* The British Journal of Nutrition, 1995. **73**(4): p. 517-30.

7. Arvaniti, K., Richard, D., and Tremblay, A. *Reproducibility of energy and macronutrient intake and related substrate oxidation rates in a buffet-type meal.* The British Journal of Nutrition, 2000. **83**(5): p. 489-95.

8. Gregersen, N.T., Flint, A., Bitz, C., Blundell, J.E., Raben, A., and Astrup, A. *Reproducibility and power of ad libitum energy intake assessed by repeated single meals.* The American Journal of Clinical Nutrition, 2008. **87**(5): p. 1277-1281.

9. Wansink, B., Painter, J.E., and North, J. *Bottomless Bowls: Why Visual Cues of Portion Size May Influence Intake[ast][ast]*. Obesity, 2005. **13**(1): p. 93-100.

10. Rolls, B.J., Morris, E.L., and Roe, L.S. *Portion size of food affects energy intake in normal-weight and overweight men and women.* American Journal of Clinical Nutrition, 2002. **76**(6): p. 1207-1213.

11. Hetherington, M.M., Foster, R., Newman, T., Anderson, A.S., and Norton, G. *Understanding variety: Tasting different foods delays satiation.* Physiology & Behavior, 2006. **87**(2): p. 263-271.

12. Yeomans, M.R., Lee, M.D., Gray, R.W., and French, S.J.*Effects of test-meal palatability on compensatory eating following disguised fat and carbohydrate preloads.* International Journal of Obesity, 2001. **25**(8): p. 1215-1224.

13. George, V.A. and Morganstein, A. *Effect of moderate intensity exercise on acute energy intake in normal and overweight females.* Appetite, 2003. **40**(1): p. 43-46.

14. Hill, A.J. and Blundell, J.E. *Nutrients and behaviour: Research strategies for the investigation of taste characteristics, food preferences, hunger sensations and eating patterns in man.* Journal of Psychiatric Research, 1982. **17**(2): p. 203-212.

15. Reid, C.H., Blundell, JE., and Stubbs, RJ, *What are psychometric assessments of appetite asking: a preliminary multivariate analysis.* International Journal of Obesity, 1998. **22**(Suppliment 3): p. 151.

16. Martins, C., Morgan, L.M., Bloom, S.R., and Robertson, M.D. *Effects of exercise on gut peptides, energy intake and appetite.* Journal of Endocrinology, 2007. **193(**2): p. 251-258.

17. Thompson, D.A., Wolfe, L.A., and Eikelboom, R. *Acute effects of exercise intensity on appetite in young men.* Medicine & Science in Sports & Exercise, 1988. **20**(3): p. 222-227.

18. Stubbs, R.J., Johnstone, A.M., O'Reilly, L.M., and Poppitt. *Methodological issues relating to the measurement of food, energy and nutrient intake in human laboratory-based studies.* Proceedings of the Nutrition Society, 1998. **57**(3): p. 357-72.

19. Sadoul, B.C., Schuring, E.A.H., Symersky, T., Mela, D.J., Masclee, A.A.M., and Peters, H.P.F. *Measuring satiety with pictures compared to visual analogue scales. An exploratory study.* Appetite, 2012. **58**(1): p. 414-417.

20. Farah, N.M.F., Brunstrom, J.M., and Gill, J.M.R. *Using a novel computer-based approach to assess the acute effects of exercise on appetite-related measures.* Appetite, 2012. **58**(1): p. 196-204.