***MS accepted for publication in Apptite***

Validation of an iPad visual analogue rating system for assessing appetite and satiety.

Louise Brunger1, Adam Smith1, Roberta Re2, Martin Wickham2, Andrew Philippides3, Phil Watten3  and Martin R Yeomans1

1. School of Psychology, University of Sussex, Falmer, Brighton, BN1 9QH, UK

2. Nutrition Research, Leatherhead Food Research, Randalls Road, Leatherhead, Surrey KT22 7RY, UK

3. Department of Informatics, University of Sussex, Falmer, Brighton, BN1 9QJ, UK

***Short title:***

Validation of appetite ratings

***Contact:***

Prof Martin R Yeomans

School of Psychology,

University of Sussex,

Brighton, BN1 9QH, UK

Email: martin@sussex.ac.uk

***Keywords:***

Appetite, satiety, electronic ratings

**Abstract**

The study aimed to validate appetite ratings made on a new electronic device, the Apple iPad Mini, against an existing but now obsolete electronic device (Hewlett Packard iPAQ). Healthy volunteers (9 men and 9 women) rated their appetite before and 0, 30, 60, 90 and 120 minutes after consuming both a low energy (LE: 77kcal) and high energy (HE: 274kcal) beverage at breakfast on two non-consecutive days in counter-balanced order. Rated hunger, desire to eat and how much participants could consume was significantly lower after HE than LE on both devices, although there was better overall differentiation between HE and LE for ratings on iPad. Rated satiation and fullness, and a composite measure combining all five ratings, was significantly higher after HE than LE on both devices. There was also evidence that differences between conditions were more significant when analysed at each time point than using an overall area under the curve (AUC) measure. Overall, these data confirm that appetite ratings made using iPad are at least as sensitive as those on iPAQ, and offer a new platform for researchers to collect appetite data.

**Key words**

Electronic rating system, appetite, hunger, satiety.

**Introduction**

Visual analogue scales (VAS) are frequently used to assess aspects of appetite both after food consumption as a measure of satiety and within meals, and remain one of the key tools used by appetite researchers. Although many studies continue to use paper-based VAS, there is increasing use of ratings made on electronic devices, originating desktop computers (e.g. [Blundell et al., 2010](#_ENREF_2); [Yeomans et al., 1997](#_ENREF_15)) and expanding to hand-held devices such as the Apple Newton (e.g. [Stratton et al., 1998](#_ENREF_12); [Stubbs et al., 2000](#_ENREF_13)) and iPAQ (e.g. [Hull et al., 2012](#_ENREF_5); [Maljaars et al., 2011](#_ENREF_6); [Peters et al., 2011](#_ENREF_10)), as well as specialised tools such as the Mandometer ([Bergh et al., 2002](#_ENREF_1); [Ford et al., 2010](#_ENREF_4); [Zandian et al., 2009](#_ENREF_17)). Ratings on such devices have several advantages over pen-and-paper ratings: (1) they negate risks of mis-measurement and bias in the measurement and transcription of paper ratings; (2) they generate immediate results, freeing up researcher time from laborious transcription of paper-based ratings; (3) they allow automatic randomisation of question order when relevant, so reducing risks of rating order effects; (4) they prevent the rater referring back to their previous responses to judge how experience has changed, and so provide data unbiased by self-determined contrasts ensuring ratings better reflect actual feelings at the point of rating.

As newer and more efficient hand-held devices have emerged, the opportunity has arisen to implement VAS ratings on devices that will further simplify and standardise collection of appetite ratings, and allow more ubiquitous use of electronic rather than paper-based ratings. In this context, a key question is the length of the line on which the rating is made. Historically, paper-based VAS were 100mm long as a convenient length to generate ratings on a 100pt scale. Previous electronic versions have maintained the 100pt scale, but the actual physical length of such scales has varied depending on the size of the test screen: VAS presented on a monitor may be as long as 200mm, whereas the same apparent scale on a small handheld device such as the iPAQ is considerably shorter (64mm). This raised an important issue about discriminability: since differences as small as 5 pts on a VAS rating hunger may relate to significant differences in behaviour, can such small differences be observed on short versions of the VAS, and is discriminability improved by increasing scale length? To this end, this short report details the development and validation of VAS implemented on the Apple iPad Mini hand-held device. Just as previous studies have validated electronic versus paper-based ratings ([Stratton et al., 1998](#_ENREF_12); [Stubbs et al., 2000](#_ENREF_13)), the aim here is to validate an improved iPad based ratings system relative to an existing iPAQ-based system that has been used by various research groups in recent studies of satiety (e.g. [Hull et al., 2012](#_ENREF_5); [Maljaars et al., 2008](#_ENREF_7); [Sadoul et al., 2012](#_ENREF_11); [Zandian et al., 2009](#_ENREF_17)), at the same time contrasting a shorter (64mm: iPAQ) and longer (100mm: iPad) scale length. Past research has repeatedly shown that VAS in many forms can detect large changes in appetite, such as the reduction seen when a meal is consumed. Here we wanted to test the extent to which VAS with different physical scale lengths, were able to detect subtle effects of small, covert manipulations of energy intake as a more stringent test of sensitivity. Accordingly, we assessed changes in appetite and satiety following ingestion of both a low-energy (LE: 77kcal) and high energy (HE: 274kcal) drink consumed at breakfast, with the standard set of ratings recommended by Blundell et al. ([2010](#_ENREF_2)) made before ingestion and periodically over the two hours post-ingestion.

**Methods**

***Design***

A repeated measures (baseline, 0, 30, 60, 90 & 120 min), crossover design assessed rated appetite using visual analogue scales (VAS) on two different touch-screen devices, a Hewlett Packard iPAQ and Apple iPad Mini, before, and for the two hours after, the consumption of one of two preload drinks, a high energy drink (HE: 274kcal) and a sensory-matched low energy drink (LE: 77kcal).

***Participants***

Potential participants were recruited from Leatherhead Food Research, on the basis that they were participating in a study comparing two different touch-screen devices. Inclusion criteria were men and women aged 18-60 years of age whose body mass index (BMI) was within the range 18-28 kg/m2. Exclusion criteria included those who were pregnant, lactating, diagnosed with diabetes and any existing allergies or eversions to dairy. Participants gave written informed consent and the protocol approved by Reading Independent Ethics Committee. Two participants failed to attend all sessions and their data were excluded. The 18 participants (9 men and 9 women) who completed all sessions had a mean age of 28.5 years (SD: 5.50) and normal BMI of 23.4 kg/m2 (SD: 3.42).

***Test preload drinks***

The selected test drinks were based on those used in recent studies of satiety at University of Sussex ([McCrickerd et al., 2012](#_ENREF_8); [McCrickerd et al., 2014](#_ENREF_9); [Yeomans & Chambers, 2011](#_ENREF_14); [Yeomans et al., 2014](#_ENREF_16)) and were based on a combination of fromage frais and a commercial fruit juice drink (Mango, Peach and Papaya, Tropicana plc), along with small quantities of a thickener as that enhanced satiety in previous studies. The ingredients and energy content of the 320ml servings of both drinks can be seen in Table 1. Energy was added as maltodextrin, and concentrations of other ingredients and flavourings had been adjusted to minimise sensory differences. Drinks were made the day before serving and were stored in a refrigerator at 4oC. Each 320ml serving of the LE drink provides 77.4 kcal, 2.2g of protein, 12.2g of carbohydrate and less than 1g of fat whereas the HE drink provided 273.9 kcal, 1.2g of protein, 66.3g of carbohydrate and a negligible amount of fat.

***Electronic ratings***

At each test time point, participants made the same set of five ratings of their appetite on the two test devices, the iPAQ (Hewlett Packard) and iPad (iPad Mini, Apple). Ratings were made on vertical lines (visual analogue scales, VAS) and the five questions were those recommended for use in appetite studies ([Blundell et al., 2010](#_ENREF_2)). Three: “How hungry are you?”, “How full are you?” and “How satiated are you?”; were end-anchored with “Not at all” and “Extremely”. “How much do you think you could eat right now?” was end-anchored with “Nothing at all” and “Very much”, while “How strong is your desire to eat?” was end anchored with “Weak” and “Very strong”. The order of use of the devices (iPAQ or iPad first) was counterbalanced across participants and order of questions randomised by the device for each set of ratings. The main differences between iPAQ and iPad were: (a) the length of the line used to make the ratings, 64mm on iPAQ but 100mm on iPad; (b) the method of recording the data, with the line marker moved using a stylus on iPAQ but by hand using the touchscreen on iPad (see Figure 1); and (c) the distance between the end-anchors and line end on the devices, which was greater on iPad because of the larger screen size.

A key issue in recording VAS data is security. Data on the iPad were stored locally in an inaccessible form incorporating the time stamp (date and time each rating was made) and this information was transmitted automatically into a secure database via a wireless connection. Data on the iPad is uneditable, ensuring that the raw data cannot be tampered with and so meeting the highest levels of data security. Data were then extracted from the secure database for analysis.

***Procedure***

Participants were instructed to eat as normal on the day before testing, but consume only water from 11pm the prior evening. On each test day, participants arrived at the unit fasted at 08.45h and completed baseline appetite ratings on both devices. The appropriate test drink was served as breakfast between 08.50-08.55h and participants repeated the ratings following complete consumption of the test drink. Participants then repeated these questions at four further time points: 30, 60, 90 and 120 min post-consumption, and were only allowed to consume water during this time.

***Data analysis***

In total, participants made six ratings of each of the five satiety-related questions on each device at each session. These ratings were contrasted using three analysis strategies. Firstly, individual ratings were contrasted using 3-way ANOVA, with device (iPAQ or iPad), drink (HE or LE) and time (baseline (-10 minutes) and 0, 30, 60, 90 and 120 minutes post-ingestion) as factors. Order of testing (HE or LE first) and order of device use (iPAQ or iPad) first were included in preliminary analyses but as these were not significant, these were not included in the final model to maximise power. Secondly, the five individual ratings were averaged to generate a composite satiety score, with scores for hunger, desire to eat and how much could you eat reversed. The same ANOVA was then conducted on this composite score. Finally, in line with current advice ([Blundell et al., 2010](#_ENREF_2)), changes in all measures from baseline were calculated and these change scores used to compute an overall area under the curve (AUC) measure for each rating across the 2-hours post-ingestion. These AUC measures were contrasted using ANOVA with drink energy and rating device as factors.

As the aim was to validate ratings on the newer iPad relative to established iPAQ, in addition for each participant we correlated their ratings for each question on both devices, and then looked at the range and average of these ratings to estimate consistency.

**Results**

***Effects of breakfast energy on appetite and satiety***

Hunger ratings (Figure 2a) at baseline did not differ significantly between treatment conditions and hunger decreased similarly after consuming both LE and HE versions of the drink rated on both devices. Hunger ratings averaged across all time periods and conditions were significantly higher on iPad than iPAQ (F(1,15) = 8.78, p<0.01, **η2**= 0.37: iPad 58.9, iPAQ 56.7), with the largest difference at baseline (iPad 77.0 ± 2.7, iPAQ 72.7 ± 3.4). The effect of device depended on drink energy (F(1,15) = 7.63, p = 0.015, **η2**= 0.34), with an average difference in hunger between HE and LE of 7.4 on iPad but only 4.1 on iPAQ. The recovery of hunger over the two hours post-ingestion also depended on drink energy content (energy x time, F(5,75) = 3.31, p = 0.009, **η2**= 0.18), as expected, with more rapid recovery of hunger in the LE than HE condition. As can be seen (Figure 2a), the difference between LE and HE versions tended to be more pronounced when rated on iPad than iPAQ, suggesting that the iPad was slightly more sensitive in discriminating recovery of hunger. Surprisingly, hunger ratings were not significantly different between female and male participants at any time point (data not shown). When changes in hunger rating were summarised as the AUC relative to baseline, hunger was suppressed more by HE than LE (F(1,15) = 5.24, p=0.037, **η2**=0.26). Although the AUC measure did not differ between device overall (F(1,15) = 1.74, p=0.22, **η2**= 0.10) the interaction between device and energy was significant (F(1,15) = 8.43, p=0.011, **η2**= 0.36) with the difference between HE and LE conditions larger when rated on iPad than iPAQ (Figure 2a inset).

The pattern of change for rated fullness was essentially the reverse of hunger (Figure 2b), with an initial increase in fullness immediately post-ingestion, and then a more rapid decline in fullness in the LE than HE condition, evidenced by a significant time x energy interaction (F(5,75) = 4.68, p = 0.001, **η2**= 0.24). In contrast to rated hunger, there was no overall difference in fullness ratings between devices (F(1,15) = 0.41, p=0.53, **η2**= 0.03) and the device x energy interaction was not significant (F(1,15) = 0.07, p = 0.79, **η2**= 0.01). The change in fullness over time varied with sex, however, (F(5,75) = 2.39, p = 0.046, **η2**= 0.14), with men less full before breakfast and showing a smaller increase in fullness after consuming the drinks regardless of energy content. AUC measures for change in fullness after baseline (Fig 2b inset) did not detect the effects of the energy difference (F(1,15) = 2.34, p = 0.15, **η2**= 0.14), although the effect was in the predicted direction with a tendency for greater fullness after HE than LE. AUC measures did tend to differ between devices (F(1,15) = 4.14, p=0.06, **η2**= 0.22), with iPad greater than iPAQ, and AUC measures tended to be more variable on the iPAQ than iPad. While the contrast of HE versus LE approached significance for iPad (F(1,15) = 4.21, p = 0.058, **η2**= 0.22), this was not so for iPAQ (F(1,15) = 2.57, p =0.13, **η2**= 0.15). Participant sex also influenced the fullness AUC measure, with a significant sex x energy x device interaction (F(1,15) = 5.74, p = 0.03, **η2**= 0.28), with the iPad picking up the effects of added energy on fullness in women more clearly than did the iPAQ.

As would be expected, the pattern of change in ratings for “satiation” was similar to that for fullness (Figure 2c), and again there was clear differentiation between HE and LE conditions (energy x time: F(5,75) = 5.12, p < 0.001, **η2**= 0.26). There was a trend for lower overall ratings on iPad (38.2 ± 3.2) than iPAQ (39.6 ± 2.9), but no evidence that responses to the energy manipulation were detected better on one device than the other. As with fullness, there was evidence of sex-differences (sex x time: F(5,75) = 2.47, p = 0.04, **η2**= 0.14), with the drinks leading to lower satiation ratings in men than women regardless of energy content. AUC scores for satiation (Figure 2c inset) confirmed a significant effect of energy (F(1,15) = 8.96, p = 0.009, **η2**= 0.37), with lower satiation in LE than HE, but no significant differences between devices.

The pattern of ratings for desire to eat were similar to those for hunger (Figure 2d), with desire to eat decreasing immediately after consuming the drinks but recovering more rapidly after the LE than HE drink (energy x time: F(5,75) = 3.65, p = 0.005, **η2**= 0.20). Although there was no overall difference between devices, there was a significant device x energy interaction (F(1,15) = 6.13, p = 0.026, **η2**= 0.29), with a larger difference overall between HE and LE on iPad (8.5) than on iPAQ (5.5), similar to that seen with hunger. The AUC measure for desire to eat (Figure 2d inset) had lower desire after the HE than LE drink (F(1,15) = 15.61, p<0.001, **η2**= 0.51), and this was not affected significantly by the device used to make the ratings.

Combination of the five separate measures into a single composite satiety score further reduced variance and so increased the power to detect subtle effects of both the drink manipulation and the way participants used the two devices. There was a small but significant difference in overall composite satiety between the two devices (iPad 40.2 ± 3.3, iPAQ 41.5 ± 3.1: F(1,15) = 5.05, p = 0.04, **η2**= 0.25), and a significant interaction between device and drink energy (F(1,15) = 5.24, p = 0.037, **η2**= 0.26), with a larger difference between HE and LE overall when rated on iPad (7.1) than iPAQ (5.5). Composite satiety scores were higher overall in the HE than LE condition (HE: 44.0 ± 3.1, LE: 37.3 ± 3.8: F(1,15) = 4.96, p = 0.042, **η2**= 0.25), but this depended on time of rating (F(5,75) = 7.22, p < 0.001, **η2**= 0.33: see Figure 2f), with the largest differences at 60 and 90 minutes. When composite satiety was summarised as AUC from baseline, satiety was greater after HE (3098 ± 498) than LE (2034 ± 571: F(1,15) = 12.81, p = 0.003, **η2**= 0.46), but although AUC scores did not differ overall between devices (F(1,15) = 3.50, p = 0.081, **η2**= 0.19), the device x energy interaction was significant F(1,15) = 6.31, p = 0.024, **η2**= 0.30), with a larger difference in satiety between HE and LE when measured using the iPad than iPAQ.

***Correlational analysis***

The average correlations for each of the five ratings (Table 2) show a high level of correspondence in ratings between the two devices, with 55% of correlations in the range 0.9 – 1.0, and 78% 0.8 – 1.0. Ratings from one male participant did have lower correlations (range -0.24 to 0.73), but this arose because ratings from that participant varied much less over time and were less affected by the two drinks so restricting the variance. For the remaining participants, ratings on the two devices were similar across a wide data range. Where ratings differed this was mainly evident at the extreme ends of the scale, where participants seemed more willing to use the top end of the scale on the iPad than iPAQ.

**Discussion**

In line with other studies using drinks with similar formulation to those tested here ([Chambers et al., 2013](#_ENREF_3); [McCrickerd et al., 2012](#_ENREF_8); [McCrickerd et al., 2014](#_ENREF_9); [Yeomans & Chambers, 2011](#_ENREF_14)), rated hunger was lower, and fullness higher, over the two hours after consuming the HE drink relative to the LE version as predicted. Critically for the study aims, these differences were evident both using the existing device (iPAQ) and the new device (iPad). However, although the differences between devices was small and subtle, there was consistent evidence that ratings made on the iPad were better able to detect the energy differences in these drinks, particularly for those ratings relating to the strength of experienced appetite (rated hunger, desire to eat and amount they could consume). Examination of the data suggested this was because of a tendency for lower ratings at the higher ends of the scale on the iPAQ than on the iPad, so leading to some restriction in data range on the iPAQ. Users also found the iPad to be simpler to use: the need to use a stylus on a small line on the iPAQ made ratings fiddly to complete, and again this may have contributed to the slightly reduced sensitivity to detect the energy difference. Overall, these data suggest that ratings made on the iPad may be a useful methodological advance in studies requiring VAS ratings.

The data analysis also suggests that different approaches to analysis of appetite ratings yield subtly different results. Although there has been a suggestion that AUC measures should be preferred ([Blundell et al., 2010](#_ENREF_2)), it was notable that several of the measures taken here were not significant when analysed using AUC, but were when ratings were analysed over multiple time-points. This result questions whether the AUC method has the sensitivity needed to reliably detect small differences in the experience of appetite and satiety post-ingestion required to detect disguised nutritional manipulations.

Although the present study confirms that VAS presented on an iPad detected subtle effects of a small covert energy manipulation, it would be valuable to conduct further research with larger manipulations of appetite to further characterise and validate this system, and to explore the predictive validity of VAS ratings made in this way in predicting actual eating behaviour.

Overall, this simple study suggests that the collection of electronic VAS appetite and satiety ratings is both reliable and sensitive and may be a useful new tool for researchers needing to collect appetite and other VAS data.

**References cited**

Bergh, C., Brodin, U., Lindberg, G., & Sodersten, P. (2002). Randomized controlled trial of a treatment for anorexia and bulimia nervosa. *Proceedings of the National Academy of Sciences,* 99, 9486-9491.

Blundell, J., De Graaf, C., Hulshof, T., Jebb, S., Livingstone, B., Lluch, A., Mela, D., Salah, S., Schuring, E., Van Der Knapp, H., & Westerterp, M. (2010). Appetite control: Methodological aspects of the evaluation of foods. *Obesity Reviews,* 11, 251-270.

Chambers, L., Ells, H., & Yeomans, M. R. (2013). Can the satiating power of a high energy beverage be improved by manipulating sensory characteristics and label information? *Food Quality and Preference,* 28, 271-278.

Ford, A. L., Bergh, C., Sodersten, P., Sabin, M. A., Hollinghurst, S., Hunt, L. P., & Shield, J. P. H. (2010). Treatment of childhood obesity by retraining eating behaviour: Randomised controlled trial. *British Medical Journal,* 340.

Hull, S., Re, R., Tiihonen, K., Viscione, L., & Wickham, M. (2012). Consuming polydextrose in a mid-morning snack increases acute satiety measurements and reduces subsequent energy intake at lunch in healthy human subjects. *Appetite,* 59, 706-712.

Maljaars, P. W. J., Peters, H. P. F., Kodde, A., Geraedts, M., Troost, F. J., Haddeman, E., & Masclee, A. A. M. (2011). Length and site of the small intestine exposed to fat influences hunger and food intake. *British Journal of Nutrition,* 106, 1609-1615.

Maljaars, P. W. J., Symersky, T., Kee, B. C., Haddeman, E., Peters, H. P. F., & Masclee, A. A. M. (2008). Effect of ileal fat perfusion on satiety and hormone release in healthy volunteers. *International Journal of Obesity,* 32, 1633-1639.

McCrickerd, K., Chambers, L., Brunstrom, J. M., & Yeomans, M. R. (2012). Subtle changes in the flavour and texture of a drink enhance expectations of satiety. *Flavour,* 1, 20.

McCrickerd, K., Chambers, L., & Yeomans, M. R. (2014). Does modifying the thick texture and creamy flavour of a drink change portion size selection and intake? *Appetite,* 73, 114-120.

Peters, H. P. F., Koppert, R. J., Boers, H. M., Strom, A., Melnikov, S. M., Haddeman, E., Schuring, E. A. H., Mela, D. J., & Wiseman, S. A. (2011). Dose-dependent suppression of hunger by a specific alginate in a low-viscosity drink formulation. *Obesity,* 19, 1171-1176.

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

Stratton, R. J., Stubbs, R. J., Hughes, D., King, N., Blundell, J. E., & Elia, M. (1998). 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,* 52, 737-741.

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

Yeomans, M. R., & Chambers, L. C. (2011). Satiety-relevant sensory qualities enhance the satiating effects of mixed carbohydrate-protein preloads. *American Journal of Clinical Nutrition,* 94, 1410-1417.

Yeomans, M. R., Gray, R. W., Mitchell, C. J., & True, S. (1997). Independent effects of palatability and within-meal pauses on intake and subjective appetite in human volunteers. *Appetite,* 29, 61-76.

Yeomans, M. R., McCrickerd, K., Brunstrom, J. M., & Chambers, L. (2014). Effects of repeated consumption on sensory-enhanced satiety. *British Journal of Nutrition,* 111, 1137-1144.

Zandian, M., Ioakimidis, I., Bergh, C., Brodin, U., & Sodersten, P. (2009). Decelerated and linear eaters: Effect of eating rate on food intake and satiety. *Physiology and Behavior,* 96, 270-275.

Table 1. Composition of the two test drinks.

|  |  |  |  |
| --- | --- | --- | --- |
| Ingredient | Supplier | Low energy (LE) | High energy (HE) |
| Grams | Kcals | Grams | Kcals |
| Mango peach and papaya juice | Tropicana plc, UK | 100 | 46 | 100 | 46 |
| 0% fat Fromage Frais | Sainsbury’s plc, UK | 55 | 27.5 | 30 | 15 |
| Low calorie peach and barley squash | Robinsons plc, UK | 35 | 3.9 | 35 | 3.9 |
| Maltodextrin C-dry md01910 | Cargill, UK | 0 | 0 | 55 | 209 |
| Yellow food colour | Silverspoon | 8 drops | 0 | 0 | 0 |
| Red food colour | Silverspoon | 2 drops | 0 | 2 drops | 0 |
| Milk caramel flavour | Synrise, DE | 0.5 | 0 | 0.5 | 0 |
| Vanilla extract | Nielsen-Massey, NL | 1 | 0 | 1 | 0 |
| Tara gum | Kaly’s Gastronomie, France | 1.2 | 0 | 1 | 0 |
| Aspartame | Ajinomoto, EU | 0.03 | 0 | 0 | 0 |
| Water |  | 130 | 0 | 100 | 0 |
|  | Total | 320 | 77.4 | 320 | 273.9 |

Table 2. Average correlation coefficients (Mean ± SEM) between ratings on the iPad and iPAQ for the different measures of appetite and satiety.

|  |  |
| --- | --- |
| Rating | Correlation (n=17) |
| How hungry do you feel? | 0.83 ± 0.04 (range 0.27 – 0.94) |
| How full do you feel? | 0.76 ± 0.08 (range -0.23 – 0.98) |
| How satiated are you? | 0.84 ± 0.05 (range 0.17 – 0.98) |
| How strong is your desire to eat? | 0.85 ± 0.05 (range 0.16 – 0.99) |
| How much could you eat? | 0.87 ± 0.04 (0.73 – 0.99) |

**Figure Legends**

Figure 1. The two touchscreen devices, iPAQ on the left and iPad right, used in the study displaying the same rating (How hungry do you feel?). Ratings on the iPAQ were made using a stylus to move the marker: on the iPad the participant used the target circle to allow them to scroll the marker to the correct position.

Figure 2. The line graphs show ratings of (A) hunger, (B) fullness, (C) satiation, (D) desire to eat and (E) how much could be eaten, and (F) a composite measure across all five ratings, before and for the two hours after consuming a drink with low (77kcal: broken lines) or high (274kcal: solid lines) energy content rated both on an iPAQ () and iPad (). Data are mean rating ± SEM at each time point. Inset for each panel are the same data summarised as mean (±SEM) area under the curve (AUC) measures calculated in each condition on both devices relative to baseline.