Postprandial suppression of appetite is more reproducible at a group than an individual-level: implications for assessing inter-individual variability

Javier T. Gonzalez[1], James Frampton[1] and Kevin Deighton[2].

[1]Department for Health, University of Bath, Bath, United Kingdom.
[2]Institute for Sport, Physical Activity & Leisure, Leeds Beckett University, Leeds, United Kingdom.

Running head: Variability in appetite measurement

Corresponding author
Javier T. Gonzalez
Department for Health, University of Bath, Bath, BA2 7AY, United Kingdom.
E-mail: J.T.Gonzalez@bath.ac.uk

Acknowledgements

All authors contributed to the conception and design of the study, collection, analysis and interpretation of data. J.T.G drafted the manuscript and all authors approved the final version of the manuscript. This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors. Data collection at Leeds Beckett University was supported internally by the Carnegie Research Fund. The authors declare no conflicts of interest.

Abstract

Individual differences in appetite are increasingly appreciated. However, the individual day-to-day reliability of appetite measurement is currently uncharacterised. This study aimed to assess the reliability of appetite following ingestion of mixed-macronutrient liquid meals at a group- and individual-level. Two experiments were conducted with identical protocols other than meal energy content. During each experiment, 10 non-obese males completed four experimental trials constituting high- and low-energy trials, each performed twice. Experiment one employed 579 kJ (138 kcal) and 1776 kJ (424 kcal) liquid meals. Experiment two employed 828 (198 kcal) and 4188 kJ (1001 kcal) liquid meals. Visual analogue scales were administered to assess appetite for 60 min post-ingestion. The typical error (standard error of measurement) of appetite area under the curve was 6.2 mm·60 min^{-1} (95%CI 4.3 to 11.3 mm·60 min^{-1}), 6.5 mm (95%CI 4.5 to 11.9 mm·60 min^{-1}), 7.1 mm·60 min^{-1} (95%CI 4.9 to 12.9 mm·60 min^{-1}) and 6.5 mm·60 min^{-1} (95%CI 4.5 to 11.8 mm·60 min^{-1}) with the 579, 828, 1776 and 4188 kJ meals, respectively. A systematic bias between first and second exposure was detected for all but the 4188 kJ meal. The change in appetite with high- vs. low-energy meals did not differ at a group level between first and second exposure (mean difference: -0.97 mm·60 min^{-1}; 95%CI -6.48 to 4.53 mm·60 min^{-1}), however, ~50% of individuals differed in their response with first vs second exposure by more than the typical error. Appetite responses are more reliable when liquid meals contain a higher- vs lower-energy content. Appetite suppression with high- vs low-energy meals is reproducible at the group- but not individual-level, suggesting that multiple exposures to an intervention are required to understand true individual differences in appetite.

Key words: Reliability; Hunger; Fullness; Personalized; Responder

Introduction

Understanding the regulation of appetite in humans can assist in the development of strategies to prevent and/or treat disorders of energy balance such as obesity. Subjective sensations of appetite are commonly captured using visual analogue scales (VAS), typically comprised of questions attempting to assess perceptions of hunger, fullness, satisfaction and prospective food consumption [1]. The methodology of administering these scales before, and at regular intervals after the consumption of meals/beverages, is supported as a standard and accepted tool to substantiate claims relating to the effects of foods on feeling states and motivations to eat [1]. In addition to assessing the effects of meal composition on appetite [2-6], these methods have also been applied more widely, to assess the effects of other interventions (such as acute [7-10] or chronic exercise [11], food restriction [7, 9] and environmental conditions [12]) on the subjective appetite response to a standard food or beverage.

Quantifying the day-to-day variability of a measure provides greater confidence on whether an intervention is the cause of an observed effect, as opposed to random (biological or behavioural) variability, measurement error or systematic bias [13, 14]. The day-to-day reliability of appetite perceptions in response to a meal - expressed as a coefficient of variation - has previously been reported to be in the range of 7 to 28%, in healthy, lean men [15-17]. The typical error (standard error of measurement) has been reported to be in the range of 8 to 13 $\mathrm{mm \cdot 120\ min^{-1}}$ [17].

Mixed-macronutrient liquid meals are commonly used in appetite research as "preloads" prior to ad libitum test meals, and covert manipulation of their energy content is used to assess the "sensitivity" of appetite regulation [18]. Moreover, liquid meals may produce more reliable appetite responses than semi-solid/solid meals [15-17]. Therefore, understanding the reliability of liquid meals with differing energy content is required in order to prescribe an appropriate preload energy content to detect subtle differences in appetite perceptions. However, it cannot necessarily be assumed that a measure shown to be reliable under one condition results in a reliable change in response to an intervention. For example, the measurement of appetite could be reproducible in response to a meal with a given energy content, but this does not provide insight into the reliability of the *suppression* of appetite with high- vs low-energy preloads.

With the growth of personalised approaches to nutrition and medicine [19], an increasing number of studies in the area of energy balance and appetite have attempted to understand the inter-individual differences in response to an intervention [20-25]. Whilst there is an increasing acknowledgement that measurement error needs to be considered in the interpretation of individual responses, there is still a common assumption that these individual responses are replicable. For example, an individual described as a "low responder" upon the first exposure to an intervention will remain a "low responder" upon repeated exposure to an intervention. It has therefore been suggested that to directly assess within-subject variability in response to an intervention, repeated exposure with an adequate washout is required [26]. Indeed, this appears to be relevant for appetite measurement, as the individual appetite response to a bout of exercise is not consistent enough to classify “compensators” and “non-compensators” [27]. The reliability of individual appetite responses to preloads (inducing appetite suppression by nutrition) has never been documented.

The present study aimed to investigate the day-to-day reliability of appetite perceptions in response to mixed-macronutrient liquid meals differing in energy content. In addition, by capitalising on repeated exposure to high and low-energy containing meals, it was also possible to assess both inter-individual variability and within-subject variability in appetite suppression with high-energy meals.

Methods

Study design

The data reported in this investigation are taken from two experiments previously described [28], which were both conducted according to the guidelines in the Declaration of Helsinki.

Both experiments involved a preload study design to investigate the influence of ad libitum meal composition on the compensatory energy intake response to different energy preloads. Both studies followed identical procedures, other than the energy content of the preloads. Here, the individual data have been rearranged to visit order to assess the day-to-day variability in appetite responses to mixed-macronutrient meals differing in energy content but matched for macronutrient composition and ingredients used. As previously described [28], experiment one was conducted at the University of Bath (UK) and utilised liquid meals containing a low (579 kJ; 138 kcal) and a moderate-energy content (1776 kJ; 424 kcal). Experiment two was conducted at Leeds Beckett University (UK) and utilised liquid meals containing a low- (828 kJ; 198 kcal) and a high-energy content (4188 kJ; 1001 kcal). The use of different energy contents enabled comparisons to be made regarding the reliability of subjective appetite measures in response to liquid meals of increasing energy content. Each experiment was approved by the respective Institutional Ethics Advisory Committee for the university at which experimental testing was performed, and informed written consent was obtained from all participants.

Participants and standardisation

All participants were non-smokers, weight stable for at least six months before participation and were not dieting or taking any medication. Participants had no known history of cardiovascular or metabolic disease, were classified as unrestrained eaters [29] and self-reported as recreationally active (engaging in structured exercise or sport ≥3 times/week). Participant characteristics have been previously reported [28] and are repeated for clarity. In experiment one the mean age, stature, body mass and body mass index were 22 ± 1 y, 1.80 ± 0.06 m, 81.1 ± 7.9 kg and 24.8 ± 1.6 kg/m^2, respectively. In experiment two, the mean age, stature, body mass and body mass index were 21 ± 4 y, 1.80 ± 0.05 m, 77.2 ± 6.4 kg and 24.2 ± 2.3 kg/m^2, respectively.

Diet and physical activity were standardised for 24 h prior to all trials by self-report and food diaries. Participants were asked to refrain from alcohol, caffeine and strenuous physical activity during this period. All trials commenced between 0800 and 0900 following an overnight fast (≥ 10 h).

Experimental protocol

At each testing location, 10 healthy men completed four experimental trials in a randomized (using online software: randomizer.org), double-blind, crossover design separated by ≥72 h. The four trials consisted of the low- or moderate/high-energy liquid meals each consumed on two occasions. Anthropometric measures, screening for eating behaviours [29] and self-reported habitual physical activity levels were obtained immediately before the first experimental trial.

Upon arriving at the laboratory for experimental trials, participants completed baseline visual analogue scales (VAS) to assess subjective appetite ratings before consuming the mixed-macronutrient liquid meal within a 5 min period. During the 60 min post-consumption, participants remained in the laboratory (seated and permitted to read or listen to music) whilst further VAS were administered every 15 min to assess appetite sensations. Whilst participants were not in isolation, any cues that could be seen to distort appetite perceptions were prohibited, e.g. discussions or radio/television programmes about food/appetite.

Liquid meals

Details of the mixed macronutrient liquid meals have been previously reported in detail. Briefly, each meal contained an identical macronutrient composition, but differed in energy content: 579 kJ (138 kcal) and 1776 kJ (424 kcal) in experiment one and 828 kJ (198 kcal) and 4188 kJ (1001 kcal) in experiment two. The macronutrient distribution was 58% carbohydrate, 26% fat, 16% protein comprised of single cream (Tesco, UK), maltodextrin (MyProtein, UK), whey protein isolate (MyProtein, UK), vanilla flavouring (MyProtein, UK) and tap water. The mass of each liquid meal was 550 g. All meals were consumed by participants in isolation. The meals were prepared by a third party external to the experimental trials in an attempt to ensure blinding was successful. We previously reported that participants were unaware of the energy content of the liquid meals [28].

Appetite assessment

Appetite sensations (hunger, fullness, satisfaction and prospective food consumption) were assessed at baseline and every 15 min following meal ingestion using 100 mm VAS with descriptors anchored at each end describing extremes (e.g. "I am not hungry at all" to "I have never been more hungry") [15]. Participants rated their appetite by placing a vertical line intersecting each horizontal line on paper and previous ratings were hidden to prevent the influence of a prior rating on subsequent reporting. The VAS were analysed by measuring the horizontal distance from the left-hand side of the scale to the vertical line indicated by the participant. Each VAS was analysed twice to maintain accuracy. A composite appetite score (herein referred to as "appetite" alone) was calculated for each time-point as previously described [30].

Statistical analyses

Data were analysed using Prism v7 (Graphpad Software, CA) and Excel v14.6.6 (Microsoft, WA) and are presented as means ± SD unless otherwise indicated. VAS ratings were converted into time-averaged area under the curve (AUC) values. Values are reported as 1) the absolute AUC, to compare the reliability across the different absolute energy content of meals; 2) as the satiety quotient (using μm rather than mm to equate to whole numbers [31]):

$$\text{Satiety Quotient} = \frac{\text{baseline appetite (μm)-postprandial appetite AUC (μm)}}{\text{energy content of meal (kJ)}}$$

and 3) as the difference between the moderate/high energy meals compared with the respective low-energy meals, to assess the reliability of appetite suppression. Reliability at the group level was assessed using a variety of statistical techniques [mean difference with 95% confidence intervals, typical error (standard error of measurement) and Bland-Altman plots] [13, 14, 32]. Coefficients of variation, expressed as a percentage (CV%) was also employed to compare across meals of differing absolute energy content. To assess the inter-individual variation in appetite suppression with high-energy vs low-energy meals, the SD of the true individual response to high- vs low-energy meals (SD_R) was used [33, 34]. This was calculated as:

$$SD_R = \sqrt{SD_I^2 - SD_C^2}$$

where SD_I is the SD of the difference between the high vs low-energy meals (intervention), and SD_C is the SD of the difference between the first and second exposure of the low energy meals (control). The SD_R was presented in both absolute units ($mm \oplus 60\ min^{-1}$) with 95% CI [35], and also in text as standardised, using the baseline SD [34]. Paired t-tests were used to identify differences in means. Differences were considered significant when $p < 0.05$.

Results

Absolute energy content of liquid meals

No differences in appetite perceptions were detected prior to ingestion of the drinks in either study (70 ± 12, 77 ± 10, 72 ± 11 and 74 ± 10 mm for 579 kJ visit one, 579 kJ visit two, 1776 kJ visit one and 1776 kJ visit two, and 72 ± 9, 77 ± 7, 72 ± 10 and 74 ± 9 mm for 828 kJ visit one, 828 kJ visit two, 4188 kJ visit one and 4188 visit two, respectively; p = 0.273 for between trial and p = 0.726 for between testing site comparisons). A systematic bias between the first and second exposure was detected for appetite AUC in response to meals with an energy content of 579 kJ (138 kcal) (p = 0.02), 828 kJ (198 kcal) (p = 0.03) and 1776 kJ (424 kcal) (p = 0.02), whereby higher appetite ratings were reported with the second exposure compared to the first exposure (**Table 1; Figures 1A, 1B and 1C**). In contrast, no systematic bias was apparent between the first and second exposure with the 4188 kJ (1001 kcal) meal (p = 0.2; **Table 1; Figure 1D**). When expressed in absolute units, typical errors were comparable between meals of different energy content (Table 1). However, when the satiety quotient was employed, the typical errors were higher with low-energy meals, compared to higher energy meals (Table 1).

Reliability of appetite suppression with moderate and large differences in energy content

With a moderate difference in meal energy content (1197 kJ; 286 kcal), the change in appetite AUC was –1.1 ± 10.9 and -0.2 ± 6.9 mm·60 min^{-1} with the first and second exposure, respectively. The mean difference in appetite AUC between the first and second exposure was 0.95 mm·60 min^{-1} (95% CI -9.10 to 11.00 mm·60 min^{-1}), indicating that there was not a systematic bias with the first, compared to the second exposure (**Figure 2A**). The typical error for appetite AUC with a moderate difference in meal energy content was 9.9 mm·60 min^{-1} (95% CI 6.8 to 18.1 mm·60 min^{-1}), which was similar to the typical error of hunger, fullness, satisfaction and prospective consumption AUCs (**Table 2**).

With a large difference in meal energy content (3360 kJ; 803 kcal), the change in appetite AUC was -8.3 ± 13.8 and -11.2 ± 14.8 mm·60 min^{-1}, with the first and second exposure, respectively. The mean difference between the first and second exposure was -2.90 mm·60 min^{-1} (95% CI -9.53 to 3.73 mm·60 min^{-1}), which suggests there was not a systematic bias between the first compared to the second exposure (**Figure 2B**). The typical error with a large difference in meal energy content was 6.6 mm·60 min^{-1} (95%CI 4.5 to 12.0 mm·60 min^{-1}), which was similar to the typical error of hunger, fullness, satisfaction and prospective consumption AUCs (**Table 2**).

Inter-individual variability

When data were combined from the two studies, the difference in the appetite AUC between moderate/high vs low energy liquid meals was -4.73 mm·60 min^{-1} (95% CI -10.66 to 1.21 mm·60 min^{-1}) with the first exposure (**Figure 3**). The SD_R for appetite AUC upon first exposure to high- vs low-energy meals was 9.4 mm·60 min^{-1} (95% CI 7.4 to 12.9 mm·60 min^{-1}; 1.1 in standardised units, 95% CI 0.8 to 1.5). When participants were exposed to the two meals for a second time, the difference in appetite AUC between moderate/high- and low-energy meals was -5.7 mm·60 min^{-1} (95% CI -11.6 to 0.2 mm·60 min^{-1}), which at a group level did not differ from the first exposure [mean difference = -0.97 mm·60 min^{-1} (95%CI -6.48 to 4.53 mm·60 min^{-1}); $p = 0.71$) and the SD_R was similar to the first exposure (9.2 mm·60 min^{-1}, 95% CI 7.3 to 12.7 mm·60 min^{-1}; standardized units: 1.0, 95% CI 0.8 to 1.4). However, when individual data are presented, there is a large variability in individual responses to the first and second exposure (**Figure 3**). For example, 10 of the 20 participants (50%) display a response to the second exposure that differs from the first exposure by more than the typical error.

Upon first exposure, the mean difference in ratings of hunger, fullness, satisfaction and prospective consumption with high vs low energy meals were -4.29 mm·60 min^{-1} (95% CI -10.81 to 2.24 mm·60 min^{-1}), 5.72 mm·60 min^{-1} (95% CI -0.22 to 11.67 mm·60 min^{-1}), 2.83 mm·60 min^{-1} (95% CI -5.6 to 11.26 mm·60 min^{-1}) and -6.09 mm·60 min^{-1} (95% CI -12.04 to -0.15 mm·60 min^{-1}), respectively. With the second exposure, the mean difference in ratings of hunger, fullness, satisfaction and prospective consumption were -7.03 mm·60 min^{-1} (95% CI -14.64 to 0.59 mm·60 min^{-1}), 5.74 mm·60 min^{-1} (95% CI -0.36 to 11.84 mm·60 min^{-1}), -4.10 mm·60 min^{-1} (95% CI -2.00 to 10.36 mm·60 min^{-1}) and -5.71 mm·60 min^{-1} (95% CI -11.26 to -0.16 mm·60 min^{-1}), respectively. The mean differences in hunger, fullness, satisfaction and prospective consumption did not differ with the first exposure compared with the second exposure (p = 0.4, >0.9, 0.7 and 0.9, respectively). However, there were large differences in the individual responses between first and second exposure for all ratings, with 9-11 of the 20 participants (45-55%) displaying a response to the second exposure that differs from the first exposure by more than the typical error (**Supplementary Figures 3A, 3B, 3C and 3D**).

Discussion

In the present study, we provide novel data demonstrating that the consumption of liquid meals with a higher energy content produces more reliable appetite responses compared with lower energy liquid meals. In addition, we demonstrate that the suppression of appetite by high- vs low-energy liquid meals is reproducible at the group level but not at an individual level. This suggests that repeated exposure to an intervention is required in order to assess true individual appetite responses.

Quantifying the day-to-day reliability of appetite perceptions in response to liquid meal ingestion can assist in the study design of future trials and interpretation of previous trials. The typical error of appetite AUC in response to ingestion of mixed-macronutrient semi-solid meals (1859 kJ; 444 kcal) by young healthy men has previously been reported to be in the region of 8.3 to 12.6 mm·120 min^{-1} [17]. In the present study, the typical errors ranged from 6.2 to 7.1 mm·60 min^{-1} between the liquid meals of differing energy content. It has previously been suggested that, compared with the ingestion of solid/semi-solid meals, the ingestion of liquid meals result in a more consistent metabolic and appetite response due to fewer sites where biological variation can act [17]. The energy content did not appear to influence the typical error in absolute terms, although there was a systematic bias detected for low and moderate-energy containing meals, whereby appetite ratings were higher upon second exposure to the meals containing 579 kJ (138 kcal), 828 kJ (198 kcal) and 1776 kJ (424 kcal) energy, which could result in an order effect in intervention studies. In contrast, there was no systematic bias detected between the first and second exposure to the meal containing 4188 kJ (1001 kcal) energy. Due to matching for total volume, the high-energy liquid meals would likely be more viscous than lower-energy meals in this study. However, it has previously been demonstrated that viscosity of liquid meals does not alter the subjective appetite responses to ingestion [36], and therefore the differences in viscosity between test-drinks are unlikely to have influence the findings in the present study. On the other hand, the higher palatability of the highest-energy liquid meal [28] could potentially explain the lack of systematic bias and greater reliability with higher vs lower-energy meals by eliciting stronger cognitive responses upon ingestion. Finally, the high-energy meal would likely perturb physiological signals to a greater extent than lower energy meals which would be more robustly detected by central appetite systems and manifest as more reliable appetite responses. This suggests that interventions aiming to assess the appetite response to a fixed preload should utilise a relatively high energy content. If a low energy meal or preload is desired, then a familiarisation trial may reduce or remove an order effect and researchers should ensure that the trial order is counterbalanced.

The reliability of appetite suppression with higher- relative to lower-energy containing meals, often used to assess appetite sensitivity, appeared to be dependent on the difference in energy content between the meals. For example, the typical errors for components of appetite (hunger, fullness, satisfaction and prospective consumption) ranged from 10.0 to 15.2 mm·60 min^{-1} with a modest difference in energy content (1197 kJ; 286 kcal; Table 2), compared to a range of 6.5 to 8.4 mm·60 min^{-1} with a large difference in energy content (3360 kJ; 803 kcal; Table 2). This was reflected in the typical error of the composite appetite AUC which was ~33% higher with the modest difference in energy content vs the large difference in energy content.

In order to assess the reliability of individual responses, data were combined from the two experiments. The different energy content in the meals provided by each experiment does not preclude this analysis, since this is still a within-subject comparison and the typical error to assess whether individual responses were reliable was specific to each experiment. Therefore the overall conclusion of this approach (appetite responses were reliable at the group but not individual level) is identical whether each experiment is considered in isolation (Supplementary Figures 1 and 2), or in combination (Figure 3).

When data were combined from both experiments, the suppression of appetite with higher- relative to lower-energy containing meals was reproducible at a group level when comparing the first exposure to the second exposure, as demonstrated by a small mean difference (<1 mm·60 min^{-1}) relative to the magnitude of appetite suppression (~5 mm·60 min^{-1}). There was large inter-individual variation in the suppression of appetite (Figure 3), with the characteristic spread of responses seen when normally-distributed data are plotted in rank order [26, 33]. However, when individual responses to the second exposure are presented, it is clear that individual responses of appetite suppression are not reproducible with ~50% of participants displaying a response to the second exposure that differs from the first exposure by more than the typical error for that meal (Figure 3). This is a consistent observation across hunger, fullness, satisfaction and prospective consumption (Supplementary Figures 3A, 3B, 3C and 3D) and also consistent with literature on the effect of exercise on individual appetite responses [27]. It is noteworthy that the individuals demonstrating the least reliable response with repeated exposure tend to cluster at the ends of the rank order as the “highest” and “lowest” responders, which supports the possibility that regression to the mean is contributing to this lack of consistent individual response [37]. These data suggest studying true individual variability in appetite regulation using mixed-macronutrient liquid meals is not possible without repeated exposure to the same intervention/stimulus. This study assessed the acute individual reliability of appetite responses, and therefore the sustained effect after repeated exposure remains currently unknown. In a chronic intervention, a pre-post comparison is (non-exclusively) influenced by 1) variability in measurement at baseline; 2) variability in measurement at follow-up and 3) variability in the “true” response to the intervention. The variability at each of these stages independently influences the ability to detect the true effect of an intervention.

This study is representative of baseline testing in a chronic intervention and therefore the lack of reliability at an individual-level would negatively impact on the ability to identify true responders and non-responders to a longer-term intervention and is, if anything, a conservative estimate on the variability with a longer-term intervention. Accordingly, labeling individuals as responders vs non-responders (or compensators versus non-compensators) and seeking investigations into the characteristics of these individuals warrants caution as the participants may not respond to an intervention in the same direction with repeated exposures.

In conclusion, liquid meals containing a high-energy content (4188 kJ; 1001 kcal) produce a more reliable appetite response compared to lower energy liquid meals (≤1776 kJ; ≤424 kcal). The appetite suppression induced by higher vs lower energy meals is reliable at the group level, but not at the individual level. Therefore, in order to understand individual appetite responses, repeated exposure to a given intervention is required.

Acknowledgements

All authors contributed to the conception and design of the study, collection, analysis and interpretation of data. J.T.G drafted the manuscript and all authors approved the final version of the manuscript. This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors. Data collection at Leeds Beckett University was supported internally by the Carnegie Research Fund. The authors declare no conflicts of interest.

References

1. Blundell J, de Graaf C, Hulshof T, Jebb S, Livingstone B, Lluch A, Mela D, Salah S, Schuring E, van der Knaap H, Westerterp M: **Appetite control: methodological aspects of the evaluation of foods.** *Obes Rev* 2010, **11:**251-270.
2. Gonzalez JT, Rumbold PLS, Stevenson EJ: **Appetite sensations at rest, during exercise and recovery: impact of a high-calcium meal.** *Applied Physiology Nutrition and Metabolism* 2013, **38:**1260-1267.
3. Gonzalez JT, Stevenson EJ: **Postprandial glycemia and appetite sensations in response to porridge made with rolled and pinhead oats.** *J Am Coll Nutr* 2012, **31:**111-116.
4. Gonzalez JT, Stevenson EJ: **Calcium co-ingestion augments postprandial glucose-dependent insulinotropic peptide1-42, glucagon-like peptide-1 and insulin concentrations in humans.** *Eur J Nutr* 2014, **53:**375-385.
5. Gonzalez JT, Green BP, Brown MA, Rumbold PL, Turner LA, Stevenson EJ: **Calcium ingestion suppresses appetite and produces acute overcompensation of energy intake independent of protein in healthy adults.** *J Nutr* 2015, **145:**476-482.
6. Astbury NM, Stevenson EJ, Morris P, Taylor MA, Macdonald IA: **Dose-response effect of a whey protein preload on within-day energy intake in lean subjects.** *Br J Nutr* 2010, **104:**1858-1867.
7. Gonzalez JT, Veasey RC, Rumbold PL, Stevenson EJ: **Breakfast and exercise contingently affect postprandial metabolism and energy balance in physically active males.** *Br J Nutr* 2013, **110:**721-732.
8. Alajmi N, Deighton K, King JA, Reischak-Oliveira A, Wasse LK, Jones J, Batterham RL, Stensel DJ: **Appetite and Energy Intake Responses to Acute Energy Deficits in Females versus Males.** *Med Sci Sports Exerc* 2016, **48:**412-420.
9. Deighton K, Batterham RL, Stensel DJ: **Appetite and gut peptide responses to exercise and calorie restriction. The effect of modest energy deficits.** *Appetite* 2014, **81:**52-59.
10. Deighton K, Barry R, Connon CE, Stensel DJ: **Appetite, gut hormone and energy intake responses to low volume sprint interval and traditional endurance exercise.** *Eur J Appl Physiol* 2013, **113:**1147-1156.
11. Martins C, Kulseng B, King NA, Holst JJ, Blundell JE: **The effects of exercise-induced weight loss on appetite-related peptides and motivation to eat.** *J Clin Endocrinol Metab* 2010, **95:**1609-1616.
12. Bailey DP, Smith LR, Chrismas BC, Taylor L, Stensel DJ, Deighton K, Douglas JA, Kerr CJ: **Appetite and gut hormone responses to moderate-intensity continuous exercise versus high-intensity interval exercise, in normoxic and hypoxic conditions.** *Appetite* 2015, **89:**237-245.
13. Atkinson G, Nevill AM: **Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine.** *Sports Med* 1998, **26:**217-238.
14. Hopkins WG: **Measures of reliability in sports medicine and science.** *Sports Med* 2000, **30:**1-15.

15. Flint A, Raben A, Blundell JE, 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* 2000, **24:**38-48.
16. Raben A, Tagliabue A, Astrup A: **The reproducibility of subjective appetite scores.** *Br J Nutr* 1995, **73:**517-530.
17. Gonzalez JT, Veasey RC, Rumbold PL, Stevenson EJ: **Consistency of metabolic responses and appetite sensations under postabsorptive and postprandial conditions.** *Appetite* 2012.
18. Beaulieu K, Hopkins M, Blundell J, Finlayson G: **Does Habitual Physical Activity Increase the Sensitivity of the Appetite Control System? A Systematic Review.** *Sports Med* 2016.
19. Betts JA, Gonzalez JT: **Personalised nutrition: what makes you so special?** *Nutrition Bulletin* Under Review.
20. Hopkins M, Blundell JE, King NA: **Individual variability in compensatory eating following acute exercise in overweight and obese women.** *Br J Sports Med* 2014, **48:**1472-1476.
21. King NA, Hopkins M, Caudwell P, Stubbs RJ, Blundell JE: **Individual variability following 12 weeks of supervised exercise: identification and characterization of compensation for exercise-induced weight loss.** *Int J Obes (Lond)* 2008, **32:**177-184.
22. Blundell JE, Stubbs RJ, Golding C, Croden F, Alam R, Whybrow S, Le Noury J, Lawton CL: **Resistance and susceptibility to weight gain: individual variability in response to a high-fat diet.** *Physiol Behav* 2005, **86:**614-622.
23. King NA, Horner K, Hills AP, Byrne NM, Wood RE, Bryant E, Caudwell P, Finlayson G, Gibbons C, Hopkins M, et al: **Exercise, appetite and weight management: understanding the compensatory responses in eating behaviour and how they contribute to variability in exercise-induced weight loss.** *Br J Sports Med* 2012, **46:**315-322.
24. Jebb SA, Siervo M, Fruhbeck G, Goldberg GR, Murgatroyd PR, Prentice AM: **Variability of appetite control mechanisms in response to 9 weeks of progressive overfeeding in humans.** *Int J Obes (Lond)* 2006, **30:**1160-1162.
25. Parr EB, Camera DM, Burke LM, Phillips SM, Coffey VG, Hawley JA: **Circulating MicroRNA Responses between 'High' and 'Low' Responders to a 16-Wk Diet and Exercise Weight Loss Intervention.** *PLoS One* 2016, **11:**e0152545.
26. Hecksteden A, Kraushaar J, Scharhag-Rosenberger F, Thiesen D, Senn S, Meyer T: **Individual responses to exercise training - a statistical perspective.** *Journal of Applied Physiology* 2015, **118:**1450-1459.
27. Unick JL, O'Leary KC, Dorfman L, Thomas JG, Strohacker K, Wing RR: **Consistency in compensatory eating responses following acute exercise in inactive, overweight and obese women.** *Br J Nutr* 2015, **113:**1170-1177.
28. Deighton K, Frampton J, Gonzalez JT: **Test-meal palatability is associated with overconsumption but better represents preceding changes in appetite in non-obese males.** *British Journal of Nutrition* Online ahead-of-print.

29. de Lauzon B, Romon M, Deschamps V, Lafay L, Borys JM, Karlsson J, Ducimetiere P, Charles MA, Fleurbaix Laventie Ville Sante Study G: **The Three-Factor Eating Questionnaire-R18 is able to distinguish among different eating patterns in a general population.** *J Nutr* 2004, **134:**2372-2380.
30. Anderson GH, Catherine NL, Woodend DM, Wolever TM: **Inverse association between the effect of carbohydrates on blood glucose and subsequent short-term food intake in young men.** *American Journal of Clinical Nutrition* 2002, **76:**1023-1030.
31. Green SM, Delargy HJ, Joanes D, Blundell JE: **A satiety quotient: a formulation to assess the satiating effect of food.** *Appetite* 1997, **29:**291-304.
32. Bland JM, Altman DG: **Statistical methods for assessing agreement between two methods of clinical measurement.** *Lancet* 1986, **1:**307-310.
33. Atkinson G, Batterham AM: **True and false interindividual differences in the physiological response to an intervention.** *Exp Physiol* 2015, **100:**577-588.
34. Hopkins WG: **Individual responses made easy.** *Journal of Applied Physiology* 2015, **118:**1444-1446.
35. Zar JH: *Biostatistical Analysis.* Upper Saddle River, NJ.: Prentice Hall; 1999.
36. McCrickerd K, Chambers L, Yeomans MR: **Does modifying the thick texture and creamy flavour of a drink change portion size selection and intake?** *Appetite* 2014, **73:**114-120.
37. Bland JM, Altman DG: **Regression towards the mean.** *British Medical Journal* 1994, **308:**1499.

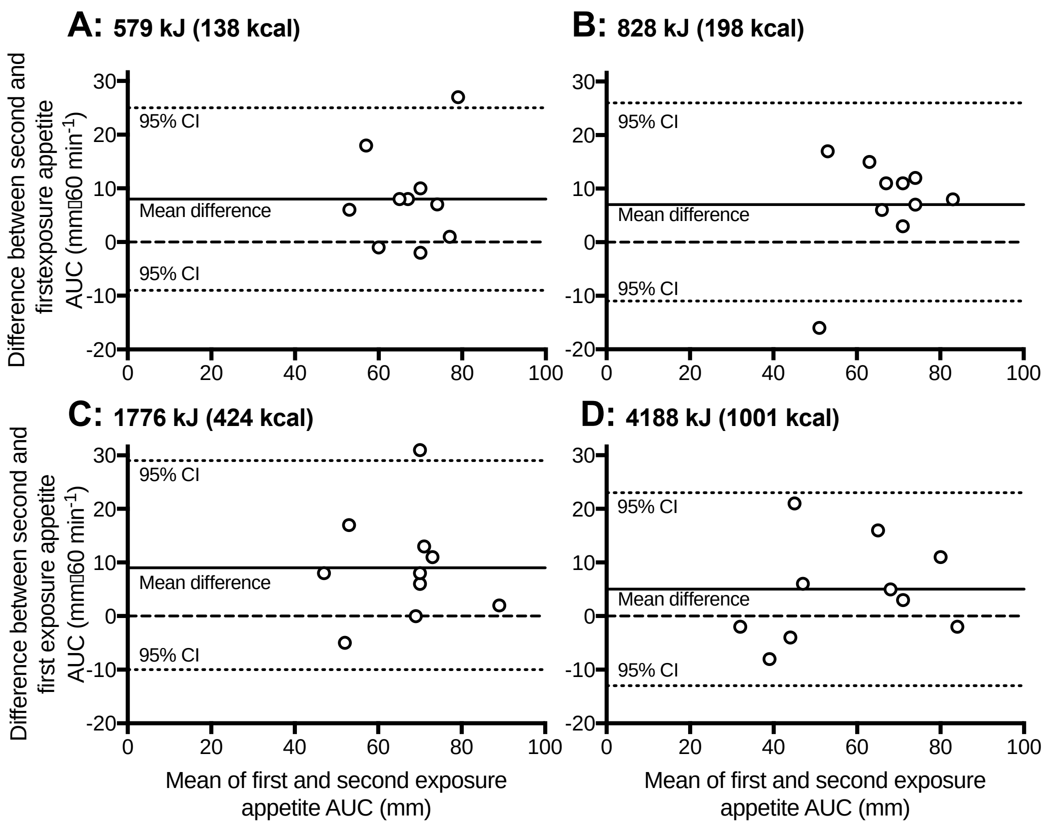


Figure 1. Bland and Altman plot of the composite appetite area under the curve (AUC) for 60 min following the ingestion of mixed-macronutrient liquid meals with an energy content of 579 kJ (A; 138 kcal), 828 KJ (B; 198 kcal), 1776 kJ (C; 424 kcal) and 4188 kJ (D; 1001 kcal).

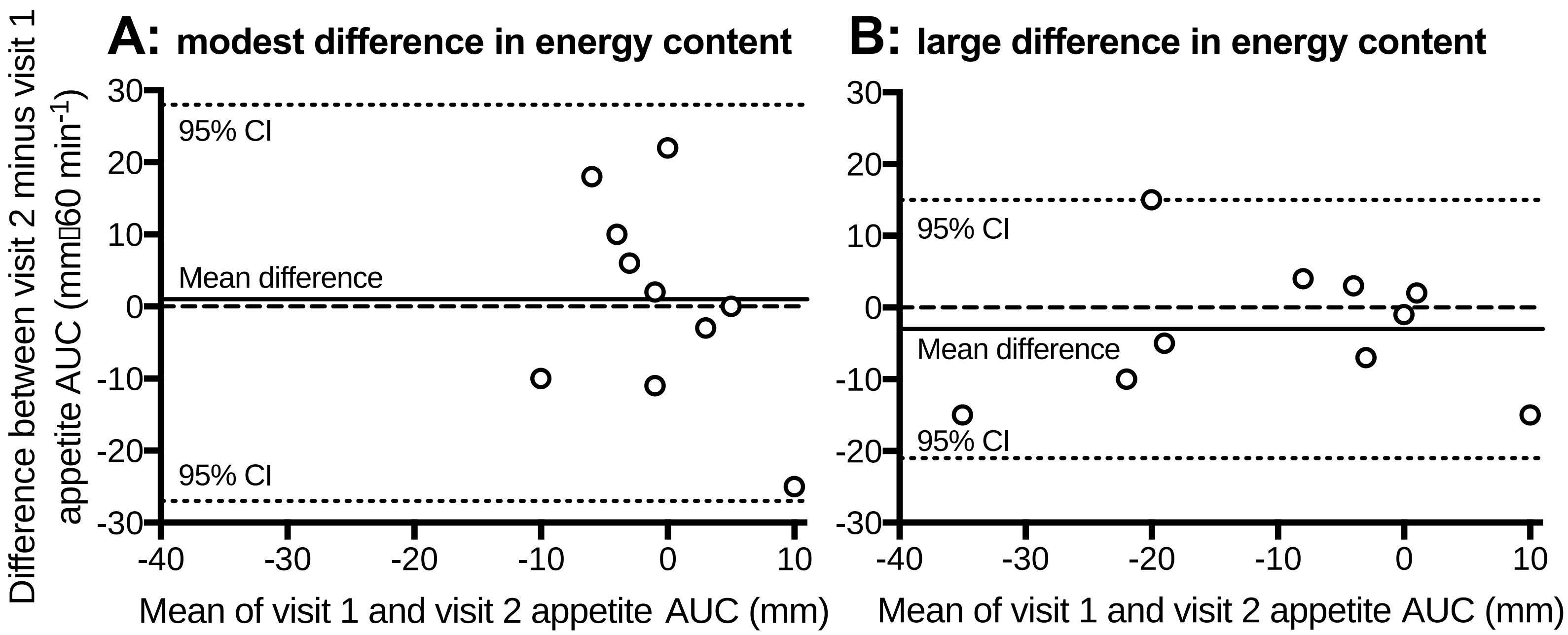


Figure 2. Bland and Altman plot of suppression of composite appetite area under the curve (AUC) for 60 min following the ingestion of mixed-macronutrient liquid meals differing in energy content by a modest (**A**; 1197 kJ; 286 kcal) or large (**B**; 3360 kJ; 803 kcal) degree.

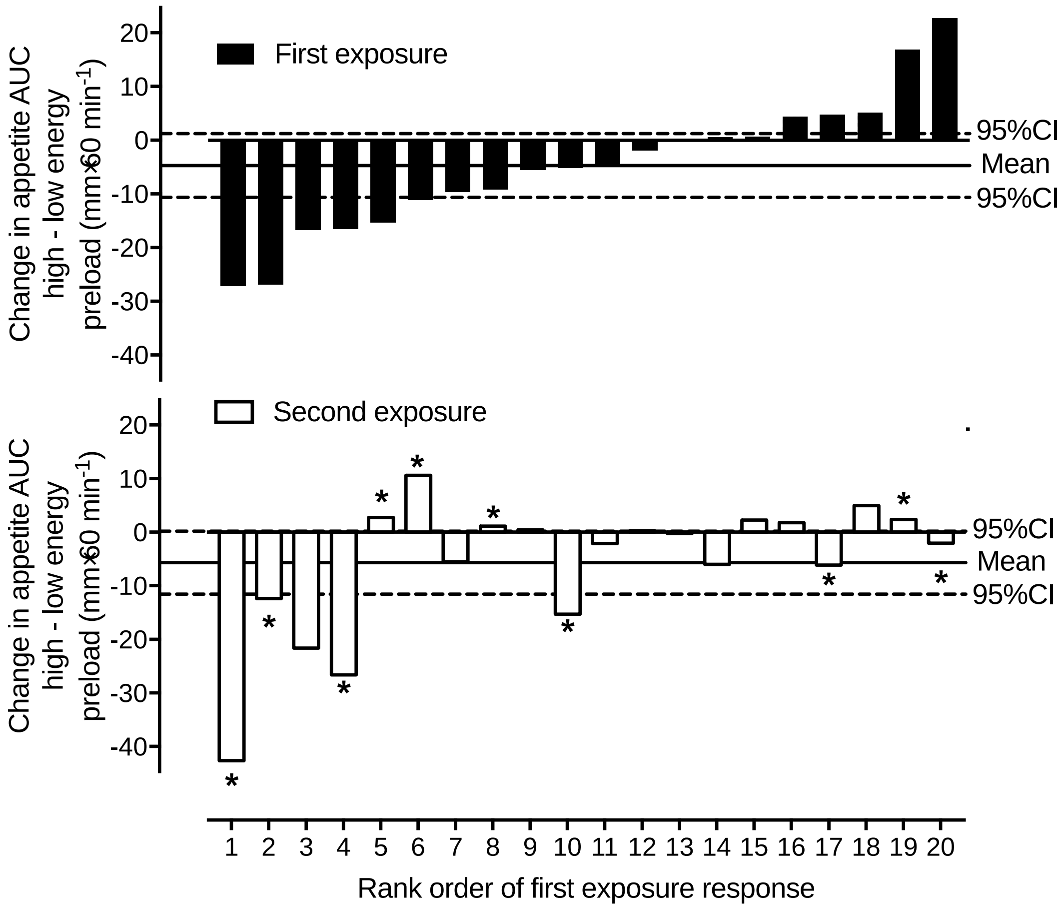


Figure 3. Individual responses in the change in composite appetite area under the curve (AUC) for 60 min following ingestion of mixed-macronutrient liquid meals with a higher- vs. a lower-energy content. *Response to second exposure differs from the first exposure by more than the typical error.

Table 1. Day-to-day variability of composite appetite area under the curve (AUC) and the satiety quotient in response to liquid meals of differing energy content.

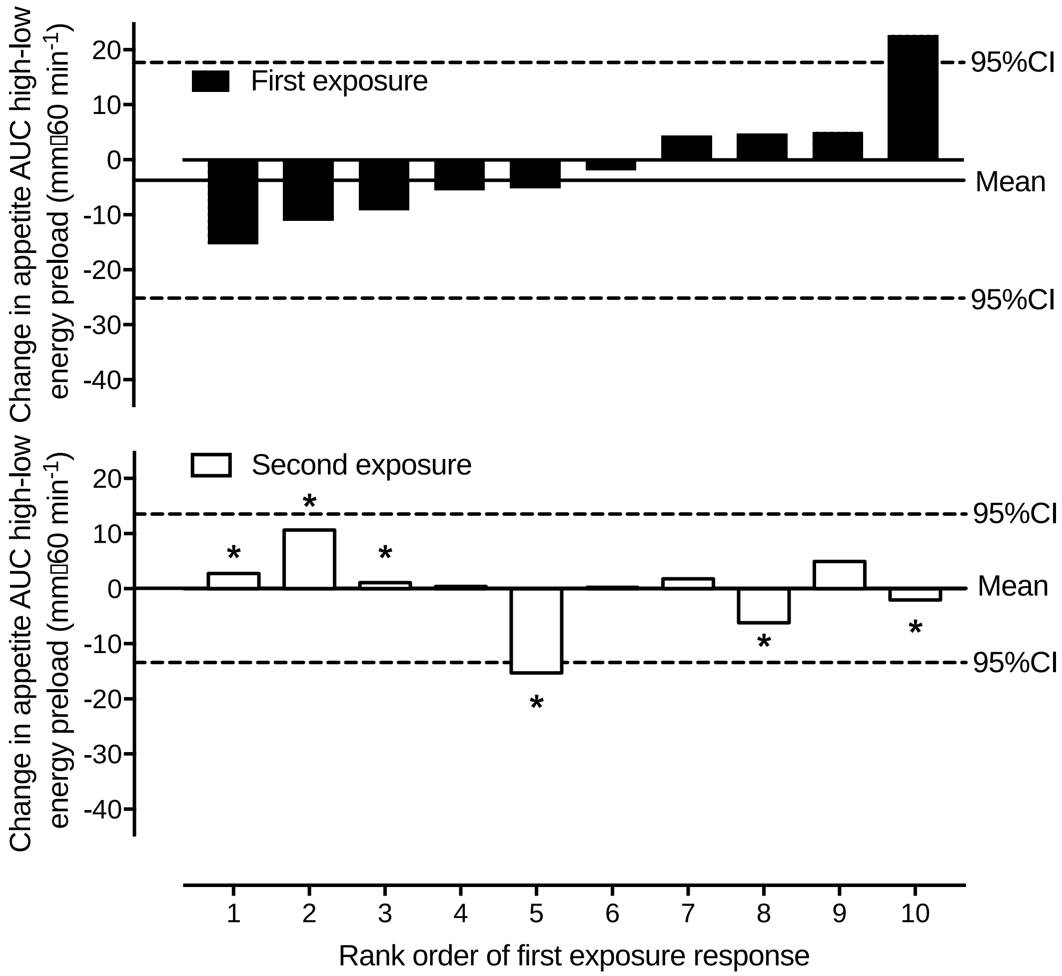
	Composite appetite AUC (mm·60 min^{-1})				Satiety Quotient [[Baseline composite appetite (μm) - Composite Appetite AUC (μm)]/energy intake (kJ)]			
	579 kJ (138 kcal)	**828 kJ (198 kcal)**	**1776 kJ (424 kcal)**	**4188 kJ (1001 kcal)**	**579 kJ (138 kcal)**	**828 kJ (198 kcal)**	**1776 kJ (424 kcal)**	**4188 kJ (1001 kcal)**
First exposure Mean (SD)	63 (9)	64 (9)	62 (13)	55 (18)	16 (23)	10 (22)	6 (5)	4 (4)
Second exposure Mean (SD)	71 (10)	71 (12)	71 (13)	60 (20)	10 (21)	2 (23)	1 (4)	4 (4)
Mean difference first vs second exposure (95% CI)	8.3 (2.0 to 14.5)	7.5 (0.9 to 14.1)	9.2 (2.0 to 16.4)	4.6 (-2.0 to 11.1)	-6.7 (-24.3 to 10.9)	-7.1 (-18.6 to 4.3)	-4.3 (-7.5 to -1.1)	-0.5 (-2.8 to 1.7)
Typical error (95% CI)	6.2 (4.3 to 11.3)	6.5 (4.5 to 11.9)	7.1 (4.9 to 12.9)	6.5 (4.5 to 11.8)	17.4 (12.0 to 31.7)	11.3 (7.8 to 20.7)	3.1 (2.2 to 5.7)	2.2 (1.5 to 4.1)
CV% (95% CI)	9.3 (6.3 to 17.6)	12.6 (8.5 to 24.3)	11.8 (8.0 to 22.6)	14 (9.6 to 27.7)	-	-	-	-

AUC, area under the curve; CV%, coefficient of variation expressed as a percentage; *n* = 10.

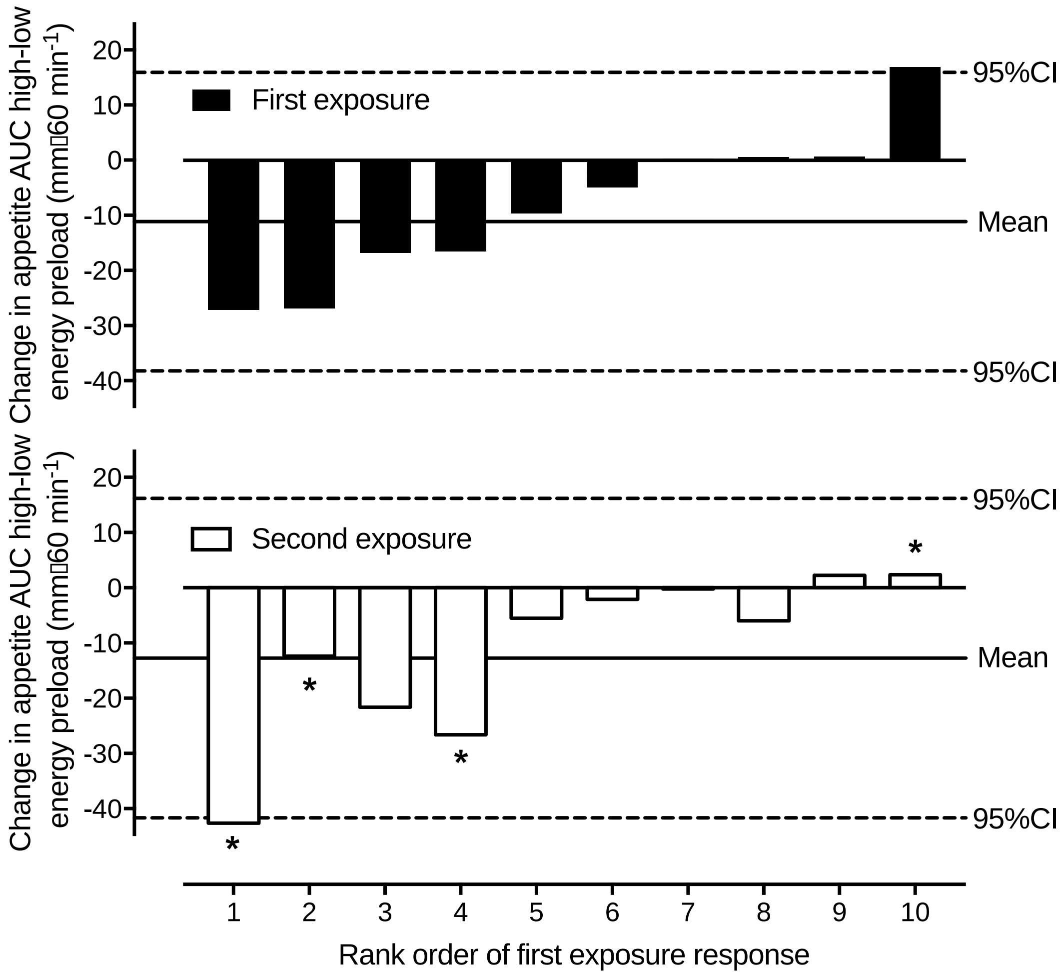
Table 2. Day-to-day variability of the change in subjective appetite ratings in response to high vs low-energy liquid meals.

	Modest difference in meal energy content (1197 kJ; 286 kcal)				Large difference in meal energy content (3360 kJ; 803 kcal)			
	ΔHunger AUC (mm·60 min^{-1})	**ΔFullness AUC (mm·60 min^{-1})**	**ΔSatisfaction AUC (mm·60 min^{-1})**	**ΔProspective consumption AUC (mm·60 min^{-1})**	**ΔHunger AUC (mm·60 min^{-1})**	**ΔFullness AUC (mm·60 min^{-1})**	**ΔSatisfaction AUC (mm·60 min^{-1})**	**ΔProspective consumption AUC (mm·60 min^{-1})**
First exposure Mean (SD)	-0.1 (11.3)	2.5 (12.0)	-3.2 (19.2)	-4.9 (8.8)	-8.4 (15.6)	8.9 (13.2)	8.9 (15.5)	-7.3 (16.1)
Second exposure Mean (SD)	-1.2 (12.1)	1.5 (6.7)	-1.5 (5.6)	0.5 (9.7)	-12.9 (18.3)	10.0 (16.5)	9.9 (16.3)	-12.0 (10.8)
Mean difference first vs second exposure (95% CI)	-1.0 (-15.2 to 13.2)	-1.1 (-11.2 to 9.1)	1.7 (-13.8 to 17.1)	5.5 (-5.41 to 16.3)	-4.4 (-12.1 to 3.2)	1.1 (-5.5 to 7.7)	1.0 (-6.5 to 8.6)	-4.7 (-13.2 to 3.8)
Typical error (95% CI)	14.0 (9.7 to 25.6)	10.0 (6.9 to 18.2)	15.2 (10.5 to 27.8)	10.7 (7.4 to 19.6)	7.6 (5.2 to 13.8)	6.5 (4.5 to 11.9)	7.4 (5.1 to 13.6)	8.4 (5.8 to 13.4)
First exposure SD_R (95%CI)	4.3 (3.2 to 4.3)	5.2 (3.8 to 8.5)	2.6 (1.9 to 4.4)	4.7 (3.4 to 7.7)	10.8 (7.9 to 17.7)	9.6 (7.0 to 15.8)	10.9 (7.9 to 17.9)	13.0 (9.5 to 21.5)
Second exposure SD_R (95%CI)	6.2 (4.5 to 10.1)	11.3 (8.2 to 18.6)	18.0 (13.1 to 29.6)	6.1 (4.5 to 10.1)	14.4 (10.5 to 23.7)	13.9 (10.1 to 22.8)	11.9 (8.7 to 19.6)	5.3 (3.8 to 8.7)

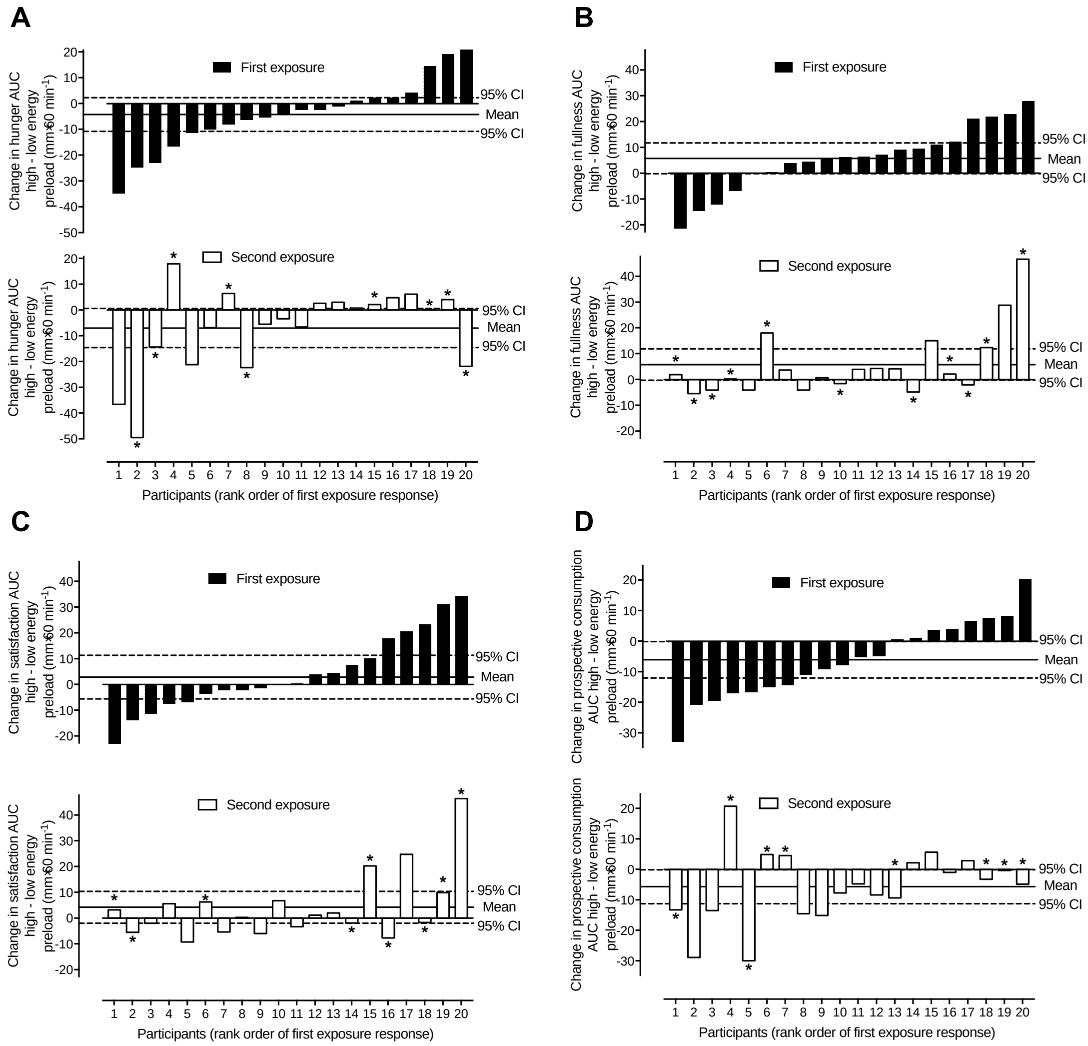
AUC, area under the curve. n = 10.



Supplementary Figure 1. Individual responses in the change in composite appetite area under the curve (AUC) following ingestion of mixed-macronutrient liquid meals with a moderate- vs. low-energy content. Data from experiment 1. *Response to second exposure differs from the first exposure by more than the typical error.



Supplementary Figure 2. Individual responses in the change in composite appetite area under the curve (AUC) following ingestion of mixed-macronutrient liquid meals with a high- vs. low-energy content. Data from experiment 2. *Response to second exposure differs from the first exposure by more than the typical error.



Supplementary Figure 3. Individual responses in the change in hunger (**A**), fullness (**B**), satisfaction (**C**) and prospective consumption (**D**) area under the curve (AUC) for 60 min following ingestion of mixed-macronutrient liquid meals with a higher- vs. a lower-energy content. *Response to second exposure differs from the first exposure by more than the typical error.