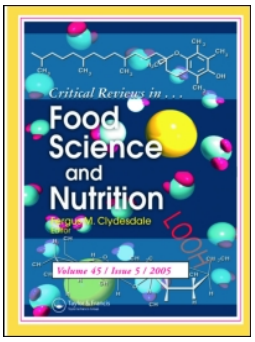

Critical Reviews in Food Science and Nutrition

ISSN: 1040-8398 (Print) 1549-7852 (Online) Journal homepage: http://www.tandfonline.com/loi/bfsn20

Systematic literature review shows that appetite rating does not predict energy intake

Guy M. Holt, Lauren J. Owen, Sophie Till, Yanying Cheng, Vicky A. Grant, Charlotte J. Harden & Bernard M. Corfe

To cite this article: Guy M. Holt, Lauren J. Owen, Sophie Till, Yanying Cheng, Vicky A. Grant, Charlotte J. Harden & Bernard M. Corfe (2017) Systematic literature review shows that appetite rating does not predict energy intake, Critical Reviews in Food Science and Nutrition, 57:16, 3577-3582, DOI: 10.1080/10408398.2016.1246414

To link to this article: http://dx.doi.org/10.1080/10408398.2016.1246414

CRITICAL REVIEWS IN FOOD SCIENCE AND NUTRITION
2017, VOL. 57, NO. 16, 3577–3582
https://doi.org/10.1080/10408398.2016.1246414

Systematic literature review shows that appetite rating does not predict energy intake

Guy M. Holt[a], Lauren J. Owen[b], Sophie Till[a], Yanying Cheng[a], Vicky A. Grant[a], Charlotte J. Harden[a], and Bernard M. Corfe[a,c]

[a]Molecular Gastroenterology Research Group, Academic Unit of Surgical Oncology, The Medical School, University of Sheffield, Sheffield, United Kingdom; [b]University of Central Lancashire, Psychology Department, Preston, Lancashire, United Kingdom; [c]Insigneo Institute for in Silico Medicine, University of Sheffield, Sheffield, United Kingdom

ABSTRACT

Ratings of appetite are commonly used to assess appetite modification following an intervention. Subjectively rated appetite is a widely employed proxy measure for energy intake (EI), measurement of which requires greater time and resources. However, the validity of appetite as a reliable predictor of EI has not yet been reviewed systematically. This literature search identified studies that quantified both appetite ratings and EI. Outcomes were predefined as: (1) agreement between self-reported appetite scores and EI; (2) no agreement between self-reported appetitescores and EI. The presence of direct statistical comparison between the endpoints, intervention type and study population were also recorded. 462 papers were included in this review. Appetite scores failed to correspond with EI in 51.3% of the total studies. Only 6% of all studies evaluated here reported a direct statistical comparison between appetite scores and EI. χ^2 analysis demonstrated that any relationship between EI and appetite was independent of study type stratification by age, gender or sample size. The very substantive corpus reviewed allows us to conclude that self-reported appetite ratings of appetite do not reliably predict EI. Caution should be exercised when drawing conclusions based from self-reported appetite scores in relation to prospective EI.

Abbreviations: EI: Energy intake; GLP: Glucagon-like peptide; MFB: Mealtime, food frequency and eating behavior; NMS: Nutrient, meal type or supplement; SLR: Systematic literature review; VAS: Visual analogue scale

KEYWORDS
Appetite; self-report appetite; appetite rating scales; visual analogue scales; energy intake; systematic review

Introduction

Appetite modification is a common strategy for modifying energy intake (EI) (Avena et al., 2013), for example, in assessment of interventions for management of obesity (Flint et al., 2000). Generic study design aimed at appetite modification involves subjects (or patients) rating appetite subjectively before and after an intervention (e.g., drug, supplement, and preload), followed by some form of test meal and evaluation of food intake or eating behavior. Controls may be different interventions in crossover or parallel format, or occasionally different subject groups. Appetite ratings are usually determined subjectively by the participant most often using visual analogue scales (VAS) and EI may be measured, overtly or covertly, using a suite of objective measures available.

VAS have been commonly used for over 30 years in numerous fields of research, sample populations, and age groups, spanning various clinical settings to community environments. VAS to assess appetite uses phrases representing polar extremes of a subjective feeling of appetite; these often include hunger, fullness, satisfaction, motivation to eat, and prospective food consumption (Blundell et al., 2010). Although it is suggested that VAS are a reliable and reproducible tool for assessing appetite (Flint et al., 2000), the relationship to EI has not previously been reviewed systematically. As product development, analysis of nutrients and the direction of strategies to modify appetite are a cornerstone of functional nutrition (in particular in obesity management) there is an unmet need to understand the true predictive value of appetite scores in relation to EI.

We aimed to address this question systematically by identifying and reviewing the corpus of studies that have measured both subjective ratings of appetite and EI in order to evaluate the predictive rigor of subjective appetite scores for EI.

Methods

Literature search strategy

This was a systematic literature review (SLR) using predefined and agreed search terms, inclusion criteria, and review strategy, as defined by the Cochrane Handbook for Systematic Reviews for Interventions (Higgins, 2011). The search terms (Table 1) and databases were selected through discussion between the

CONTACT Dr. Bernard M. Corfe b.m.corfe@sheffield.ac.uk Molecular Gastroenterology Research Group, Medical School Building, University of Sheffield, Beech Hill Rd., Sheffield, South Yorkshire S10 2RX, United Kingdom.
Supplemental data for this article can be accessed on the publisher's website.

Table 1. Search terms used in this study and their combinatorial permutations.

Appetite		Satiation		Energy consumption
OR		OR		OR
Hunger		Satiety		Energy intake
OR	AND		AND	OR
Hungry				Food consumption
OR				OR
Desire to eat				Food intake

research team and an independent research librarian in order to obtain a comprehensive generation of sources.

Following an initial combined search of Scopus, Ovid Medline, and Web of Science literature databases, the titles and abstracts were screened to generate a refined list of sources and these were deduplicated using Endnote Online™. Meta analyses, review papers, position statements, and animal studies were not included and studies that did not measure both appetite and EI, or showed nonrelevance to the topic, were excluded.

Scoring criteria and statistics

Initially, we aimed to evaluate the proportion of papers reporting a significant correlation between VAS appetite ratings and EI. However, only a small proportion (6.3%) of the corpus reported such analyses directly, so we devised a scoring matrix.

- "Link" studies included:

A. Studies reporting significant and codirectional effects on both appetite scores and EI (for example, increased hunger scores and increased EI, or decreased hunger scores and decreased EI).

B. Studies reporting no significant effect on either endpoint (for example, no alterations in hunger scores or EI). This is because neither variable changed in response to a particular event or intervention.

- "No link" studies included:

A. Studies reporting that appetite scores OR EI were significantly perturbed (for example, increased/decreased hunger scores but no changes in EI, or no changes in hunger scores but increased/decreased EI).

B. Studies reporting significant differences in appetite scores and EI, but in opposite directions (for example, significant increase in hunger scores and significant decrease in EI).

Further information extracted from each paper included the journal title, year of publication, author, health status of sample group, intervention or perturbation under study, type of VAS used, measure of EI, sample size, sex, and age of sample group. In addition, we ascertained if a link between appetite scores and EI was determined "directly" by the original authors through statistical analysis (for example, using regression or correlation) or "indirectly" inferred by the reader using reported significance and direction as described previously. Two researchers independently scored each paper in the corpus; where disagreement occurred a third scorer independently evaluated the paper.

Statistics

Inter-rater variability was quantified using Cohen's kappa analysis (Hayes and Hatch, 1999). Pearson's chi-square analysis was used to test the frequency of study characteristic distribution between the two outcomes (i.e., "link," "no link"). Cohen's kappa is a test of interobserver agreement for categorical data, it is considered more robust than simple percentage agreement calculation since this test accounts for the agreement occurring by chance. The value derived is proportional agreement. Landis and Koch (1977) suggested 0.61–0.80 is substantial agreement and any higher number represents almost perfect agreement.

Results

Overview of the literature search, screening, and studies included

The initial combined search of Scopus, Ovid Medline, and Web of Science during the period 1999–2015 returned 3,842 results. Following primary exclusion based on the abstract (generally due to lack of subjective appetite measurement) and deduplication, the revised list comprised 714 studies. Upon reading these papers in full, a further 252 were excluded due to the absence of EI measurement and/or nonrelevance to the topic. Four hundred sixty two studies were, therefore, included this review (Figure 1). For a tabulated list of all studies included in the review please refer to the Online Supplementary Material.

There was good agreement between all scorers represented by Cohen's kappa values of 0.89 (authors ST/YC vs GMH, 1999–2014) and 0.96 (authors GMH vs LJO, 2015). Over the literature search time period, the number of included studies generally increased, with around 20 per annum in 2000–2006 rising to around 40 per annum in 2009–2014 (Figure 2). As the majority of papers (93.7%) did not report a direct statistical comparison between VAS ratings and EI, the relationship between these two measures was opaque in reports and required us to develop an independent scoring system (vide supra).

Overall relationship between appetite scores and EI

Out of the 462 studies included in this review, 225 (49%) demonstrated a "link" between self-reported appetite scores and measurement of EI. In 237 (51%) of the studies, there was "no link" between appetite scores and EI (definite disagreement between the outcomes).

Direct and indirect comparisons

The majority of papers (93.7%) did not directly (statistically) compare self-reported appetite ratings and EI, meaning that the relationship between the two measures was deduced by the researcher by identifying whether statistically significant changes occurred and were codirectional between the two outcome measures. Studies that were categorized as demonstrating a "link" between self-reported appetite scores and EI were approximately 30% more likely to have reported a direct statistical comparison of the two measurements ($p = .001$).

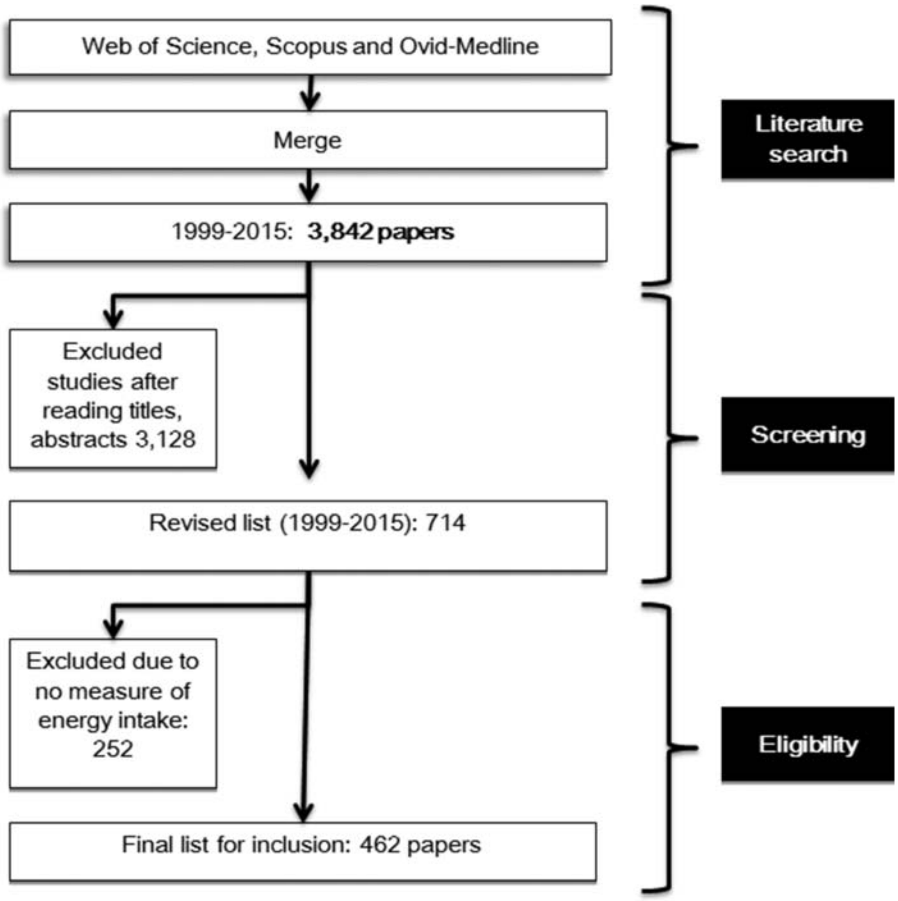

Figure 1. PRISMA flowchart. The figure summarizes the literature search workflow, starting with searches across three databases using defined terms, screening, merging, deduplicating, and production of the final corpus of papers.

Assessment of factors potentially influencing outcome

Effect of age and sex

Twenty-one studies used a sample population of children and/or adolescents (4–17 years), whilst an elderly sample population (≥65 years) was used in nine studies. The 432 remaining studies used participants with large and variable age ranges (18–64 years) making it unfeasible to conduct a separate analysis of young adult and middle-aged subjects. Due to the small number of studies in the ≥65 years category, chi-squared analysis was not appropriate because the assumptions of this test would be violated. Assessment of the data reported in Table 2 shows that, across all age groups, the frequency of studies are equivocal in the "link" and "no link" categories. The chi-squared test showed no significant difference between the expected frequencies and the observed frequencies between participant sexes (Table 2). These data demonstrate no advantage of sex in predicting EI from self-reported appetite scores.

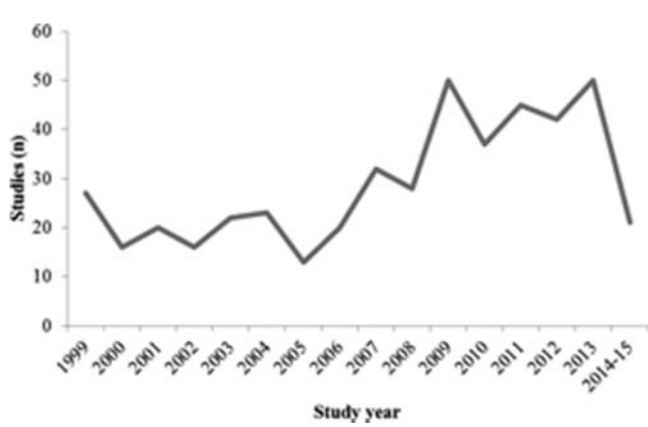

Figure 2. Number of studies within each search year. Analysis of number of eligible papers by year shows progressive increase in papers per annum in this area.

Effect of sample size

The chi-squared test showed no significant difference between the expected frequencies and the observed frequencies in any sample size category (Table 2). These data demonstrate no advantage in predicting EI from self-reported appetite scores by stratifying by sample size.

Effect of health status

Studies were divided according to the reported health status of the sample population. Three hundred forty two studies used "healthy individuals," 115 studies used individuals with clinically defined illness/disease ("nonhealthy"), and five studies did not report the health status of their chosen study population. The chi-squared test showed no significant difference between the expected frequencies and the observed frequencies in either category. These data demonstrate no advantage in predicting EI from self-reported appetite scores by stratifying for health

Table 2. Frequencies and expected frequencies of studies by population, intervention, and measurement tools used. Chi squared and p values shown. For expected cell frequencies less than 5, the chi-square approximation may not be reliable. A standard and conservative rule is that no more than 20% of the expected frequencies (shown in brackets) are less than 5 and all individual expected counts are 1 or greater. For age (years) category data reported in the majority of studies were not sufficiently detailed to better categorize studies based on age of population assessed; therefore, the data violate assumptions of the test and so χ^2 and p value are not reported here.

	Link		No-link				
Groups	Frequency; observed (expected)	%	Frequency; observed (expected)	%	Total	χ^2	p value
Age (years)							
4–17	10	47.6	11	52.4	21	—	—
18–64	210	48.6	222	51.4	432		
≥65	5	55.6	4	44.4	9		
Gender							
Male	78 (79.4)	47.8	85 (83.6)	52.2	163	0.209	0.900
Female	35 (36.0)	47.3	39 (37.9)	52.7	74		
Mixed	112 (109.6)	49.8	113 (115.4)	50.2	225		
Sample size							
≤10	18 (19.9)	43.9	23 (21.0)	56.10	41	1.954	0.744
11–30	133 (135.9)	77.7	146 (143.1)	22.3	279		
31–50	47 (45.3)	50.5	46 (47.7)	49.5	93		
51–100	22 (18.5)	57.9	16 (19.5)	42.1	38		
>100	5 (5.4)	45.5	6 (5.6)	54.5	11		
Health status							
Healthy	165 (166.5)	48.2	177 (175.4)	51.8	342	0.109	0.740
Nonhealthy	60 (58.4)	50.0	60 (61.6)	50.0	120		
VAS length							
<100 mm	5 (3.9)	62.5	3 (4.1)	37.5	8	9.764	0.439
100 mm	165 (164.1)	48.9	172 (172.9)	51.1	337		
≥101 mm	16 (20.9)	38.1	27 (21.1)	61.9	43		
Likert scale	15 (15.1)	48.4	16 (15.9)	51.6	31		
Not stated	24 (20.9)	55.8	19 (22.1)	44.2	43		
Intervention type							
Age/Gender	4 (3.9)	50.0	4 (4.1)	50.0	8	25.800	<.001
Exercise	4 (8.7)	22.2	14 (9.2)	77.8	18		
Food form	6 (17.5)	16.7	30 (18.5)	83.3	36		
MFB	27 (28.2)	46.6	31 (29.8)	53.4	58		
NMS	146 (136.0)	52.3	133 (143.0)	47.7	279		
Pharmacological	26 (19.5)	65.0	14 (20.5)	35.0	40		
Other	12 (11.2)	52.2	11 (11.8|)	47.8	23		

NMS = nutrient, meal type or supplement; MFB = mealtime, frequency or behavior.

status. The nonhealthy group included 99 studies which had recruited overweight and/or obese participants (47 = link; 53 = no link). Ten studies had recruited examined diabetic participants (9 = link; 1 = no link) and six studies had recruited participants with eating disorders (2 = link; 4 = no link). Of the five studies that did not report the health status of their participants, three demonstrated a link and two demonstrated no link. There were insufficient data available in these subgroups to conduct further chi-squared tests.

Effects of the type of self-reported appetite measurement tool and VAS length

The most commonly used VAS length was 100 mm, which was used in 337 (73%) of the studies examined in this review: 8 (2%) studies used VAS < 100 mm; 43 (9%) studies used VAS > 100 mm; and 43 (9%) papers did not state the length of VAS. 31 (7%) researchers reported using an alternative form of measurement for obtaining appetite scores (predominantly Likert scales).The most repeatedly used VAS was the one devised by Hill and Blundell (Blundell et al., 2010). Nine studies reported the use of an electronic form of VAS. The chi-squared test showed no significant difference between the expected frequencies and the observed frequencies in any category. These data demonstrate no advantage for any particular length of VAS scale for predicting EI from self-reported appetite scores.

Effect of intervention type

The different types of interventions observed in the corpus of papers were categorized as follows: nutrient, meal type, or supplement group (NMS, e.g., carbohydrate vs fat, glycaemic load) = 279, (60%) of studies; meal time, food frequency and eating behavior (MFB, e.g., TV viewing during meal times, portion size information) = 58 (12%); pharmacological (e.g., glucagon-like peptide, GLP, analogues, appetite hormone administration) = 40 (9%); food form (e.g., solid vs liquid) =

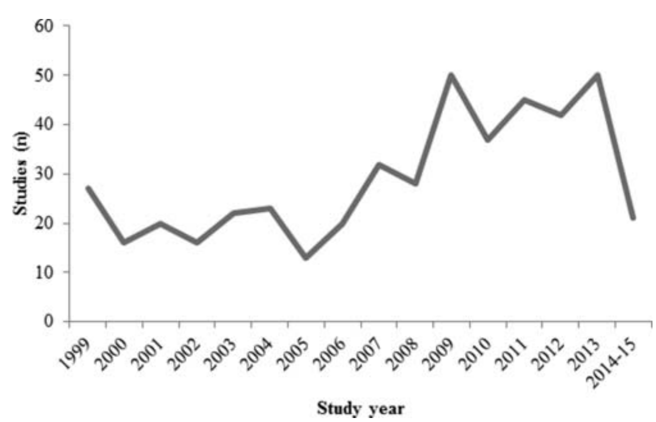

Figure 3. Frequency of studies reporting either a link or no link between self-report appetite and EI divided by study interventions ($n = 462$).

36 (8%); other = 23 (5%); exercise (e.g., intensity or duration) = 18 (4%); and age/sex 8 (2%) (Figure 3). The chi-squared test did show significant difference between the expected frequencies and the observed frequencies when stratified by intervention type ($p < .001$). The intervention type with the highest proportion of studies reporting a link between subjective appetite scores and EI was the "pharmacological" group: of the 40 studies identified in this category, 26 (65%) demonstrated a link and 14 (35%) showed no link. It is further of note that nine of the studies in this category explored the effect of GLP of which five showed a link between self-reported appetite scores and EI. The group of papers that were categorized as examining "food form" contained the highest number of studies that failed to observe a link between self-rated appetite scores and EI. Thirty out of the 36 studies (83.3%) reported no link.

Discussion

This was a systematic literature review addressing an unmet need to evaluate critically the validity of the putative link between subjectively rated appetite (implied by many of the reviewed papers as a proxy measure of prospective energy intake) and measures of energy intake. Studies were selected on the basis of reporting and quantifying both appetite and EI. A striking finding of this review was that the vast majority of studies (93.7%) did not undertake any statistical analysis to compare self-reported appetite scores and EI. This rendered formal approaches to meta-analysis impossible due to the degree of data disconnection, and required us to develop an assessment criterion to determine if the direction of significant change in response to an intervention was in agreement (link) or not (no link) between the two endpoints. Papers that had reported a direct statistical comparison between the two measures were more likely fall into the "link" category. This may suggest selective reporting in the "no link" studies: they either did not undertake statistical analyses or chose not to report their findings.

Over half of the studies included in our analysis did not demonstrate a link between appetite scores and EI. Of the 225 that demonstrated a link between appetite scores and EI, 54 (24%) scored a double negative (i.e., neither subjective appetite score nor EI changed from baseline; Supplementary Online Information). These studies were conservatively categorized as a "link," although it may be argued that a double-negative does not imply that both measures relied upon the same mechanism or processes. If double-negative studies are considered as an independent category, the number of papers evidencing a definitive link between appetite scores and subsequent EI declines to 37%, further strengthening our conclusion.

The heterogeneous nature of the studies included in this review provided scope for analysis according to factors such as age, sex, sample size, health status, self-reporting measurement tool (VAS length), and intervention type (Table 1). No study factor improved the utility of self-reported appetite scores on EI, other than "intervention type": of the 40 studies identified in the "pharmacological" group, 26 (65%) demonstrated a link, while 14 (35%) showed no link. A subset of nine studies belonging to the pharmacological group used appetite hormone administration as an intervention strategy, the most common of which was GLP. In response to subcutaneous (Flint et al., 2013; Horowitz et al., 2012) or intravenous administration (Gutzwiller et al., 1999a; Gutzwiller et al., 1999b; Naslund et al., 1999) of GLP, 55% of these studies found VAS appetite scores could predict EI reliably. It is well documented that GLP is an effective appetite modifier (Dailey and Moran, 2013) and its potent physiological mechanism of action is perhaps the reason why a more consistent relationship between appetite scores and EI was observed.

In regard to ad libitum measures, recording the number of calories consumed at a test meal is commonplace, and is considered to be an objective reflection of appetitive response (Blundell et al., 2010). The general protocol involves a test meal in which participants are instructed to consume as much as they like and/or until they reach satiation. The reproducibility of ad libitum test meals as an acute measure of EI is robust in normal weight, obese, and overweight sample groups of both children and adults (Arvaniti et al., 2000; Bellissimo et al., 2008; Lara et al., 2010). Booth (2009) however, questioned the nature of the subjective and objective divisions of appetite measurement, arguing that subjective ratings actually reflect an objective approach and that the rarified environmental control applied in so-called objective ad libitum test settings may inherently perturb intake. Our analysis supports the distinction made by Booth between emotive and behavioral aspects of appetite and feeding behavior and demonstrates unequivocally that one is not a proxy for the other. A sequella of this separation is that the research community needs to reflect on the meaning and implication of "appetite" as an emotive and self-reported measure, and re-evaluate its purpose. Rated appetite may function similarly to eating-behavior questionnaires (Lowe and Butryn, 2007) as a proxy of individuals' attitude to food coupled to state at time of assessment. Other factors governing actual food intake include: sensorial environment (including exposure to food cues and palatable foods) social factors; entrained behavior related to food timing; consequences of energy expenditure and homeostasis. These factors are all underwritten by the physiological governance of EI which includes gut hormone release, stretch sensing, and micronutrient sensing in the gut. We suggest there is an unmet need to (1) undertake further systematic studies that weight the impact of these variables in the governance of EI and (2) identify biomarkers of behavior. Until modeling or biomarkers are validated and become widely available, we argue that the best estimate of EI remains the measurement of EI itself.

A second sequella of our analysis of critical relevance to policy and regulation is that health claims on the energy intake-modifying potential of formulations based on their impact on rated appetite should not be supported. Reformulation and product development of satiating foods and nutraceuticals is a key strategy in obesity control (Van Kleef et al., 2012). This report demonstrates that appetitive sensation is not a robust endpoint, or even screening tool, for the development of functional foods.

Conclusions

Subjective appetite scores are ineffective in predicting energy intake in studies attempting to modify and/or assess appetite. Emotional and behavioral aspects of appetite are distinct and not proxies for each other, more work is needed to explore their independence and independent value. The development of novel methods to predict EI is needed.

Acknowledgments

We thank Prof HJ Powers for a critical reading of the manuscript prior to submission.

Funding

We thank the University of Sheffield for part-funding ST's contribution.

References

Arvaniti, K., Richard, D. and Tremblay, A. (2000). Reproducibility of energy and macronutrient intake and related substrate oxidation rates in a buffet-type meal. *Br. J. Nutr.* **83**(05):489–495.

Avena, N. M., Murray, S. and Gold, M. S. (2013). The next generation of obesity treatments: Beyond suppressing appetite. *Front. Psychol.* **4**:1–3.

Blundell, J., de Graaf, C., Hulshof, T., Jebb, S., Livingstone, B., Lluch, A. and Westerterp, M. (2010). Appetite control: Methodological aspects of the evaluation of foods. *Obe. Rev.* **11**(3):251–270.

Booth, D. A. (2009). Lines, dashed lines and "scale" ex-tricks. Objective measurements of appetite versus subjective tests of intake. *Appetite* **53** (3):434–437.

Dailey, M. J. and Moran, T. H. (2013). Glucagon-like peptide 1 and appetite. *Tre. Endocr. Meta.* **24**(2):85–91.

Flint, A., Kapitza, C. and Zdravkovic, M. (2013). The once-daily human GLP-1 analogue liraglutide impacts appetite and energy intake in patients with type 2 diabetes after short-term treatment. *Diab. Obe. Meta.* **15**(10):958–962.

Flint, A., Raben, A., Blundell, J. E. and Astrup, A. (2000). Reproducibility, power and validity of visual analogue scares in assessment of appetite sensations in single test meal studies. *Interna. J. Obe.* 24(1):38–48.

Gutzwiller, J. P., Drewe, J., Goke, B., Schmidt, H., Rohrer, B., Lareida, J. and Beglinger, C. (1999a). Glucagon-like peptide-1 promotes satiety and reduces food intake in patients with diabetes mellitus type 2. *Am. J. Phy.-Reg. Integ. Compa. Phy.* **276**(5):R1541–R1544.

Gutzwiller, J. P., Goke, B., Drewe, J., Hildebrand, P., Ketterer, S., Handschin, D. and Beglinger, C. (1999b). Glucagon-like peptide-1: A potent regulator of food intake in humans. *Gut* 44(1):81–86.

Hayes, J. R. and Hatch, J. A. (1999). Issues in measuring reliability—Correlation versus percentage of agreement. *Written Communication* **16** (3):354–367.

Higgins, J. P. T. and Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions*(Version 5.1.0 [updated March 2011] ed.): The Cochrane Collaboration.

Horowitz, M., Flint, A., Jones, K. L., Hindsberger, C., Rasmussen, M. F., Kapitza, C. and Chapman, I. M. (2012). Effect of the once-daily human GLP-1 analogue liraglutide on appetite, energy intake, energy expenditure and gastric emptying in type 2 diabetes. *Diab. Res. Clin. Pra.* **97** (2):258–266

Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* **33**:159–174.

Lara, J., Taylor, M. A. and Macdonald, I. A. (2010). Is ad libitum energy intake in overweight subjects reproducible in laboratory studies using the preload paradigm? *Eur. J. Clin. Nutr.* **64**(9):1028–1031.

Lowe, M. R. and Butryn, M. L. (2007). Hedonic hunger: A new dimension of appetite? *Physiol. Behav.* **91**(4):432–439.

Bellissimo, N., Thomas, S. G., Pencharz, P. B., Goode, R. C. and Anderson, G. H. (2008). Reproducibility of short-term food intake and subjective appetite scores after a glucose preload, ventilation threshold, and body composition in boys. *Appl. Physiol. Nutr. Metab.* **33**:326–337.

Naslund, E., Barkeling, B., King, N., Gutniak, M., Blundell, J. E., Holst, J. J. and Hellstrom, P. M. (1999). Energy intake and appetite are suppressed by glucagon-like peptide-1 (GLP-1) in obese men. *Int. J. Obes. Relat. Metab. Disord.* **23**(3):304–311.

Van Kleef, E., Van Trijp, J., Van Den Borne, J. and Zondervan, C. (2012). Successful development of satiety enhancing food products: Towards a multidisciplinary agenda of research challenges. *Crit. Rev. Food Sci. Nutr.* **52**(7):611–628.