Appetite 107 (2016) 398–405

Contents lists available at ScienceDirect

Appetite

journal homepage: www.elsevier.com/locate/appet

Cross-over studies underestimate energy compensation: The example of sucrose-versus sucralose-containing drinks

Nouf S. Gadah, Jeffrey M. Brunstrom, Peter J. Rogers*

Nutrition and Behaviour Unit, School of Experimental Psychology, University of Bristol, Bristol, UK

ARTICLE INFO

Article history:
Received 29 April 2016
Received in revised form
2 August 2016
Accepted 22 August 2016
Available online 24 August 2016

Keywords:
Sugar
Low-calorie sweetener
Drink
Food
Satiety
Compensation

ABSTRACT

The vast majority of preload-test-meal studies that have investigated the effects on energy intake of disguised nutrient or other food/drink ingredient manipulations have used a cross-over design. We argue that this design may underestimate the effect of the manipulation due to carry-over effects. To test this we conducted comparable cross-over (n = 69) and parallel-groups (n = 48) studies testing the effects of sucrose versus low-calorie sweetener (sucralose) in a drink preload on test-meal energy intake. The parallel-groups study included a baseline day in which only the test meal was consumed. Energy intake in that meal was used to control for individual differences in energy intake in the analysis of the effects of sucrose versus sucralose on energy intake on the test day. Consistent with our prediction, the effect of consuming sucrose on subsequent energy intake was greater when measured in the parallel-groups study than in the cross-over study (respectively 64% versus 36% compensation for the 162 kcal difference in energy content of the sucrose and sucralose drinks). We also included a water comparison group in the parallel-groups study (n = 24) and found that test-meal energy intake did not differ significantly between the water and sucralose conditions. Together, these results confirm that consumption of sucrose in a drink reduces subsequent energy intake, but by less than the energy content of the drink, whilst drink sweetness does not increase food energy intake. Crucially, though, the studies demonstrate that study design affects estimated energy compensation.

1. Introduction

The preload-test-meal procedure has been used extensively to investigate the short-term effects of food and drink ingredients on energy intake. In this procedure participants consume a fixed amount of the preload, and the amount of food they then eat in the '*ad libitum*' test-meal is measured – the amount of food served in the test-meal exceeds what would usually be eaten for the particular occasion (e.g., a weekday lunch). The interval between preload and test-meal has varied between studies, ranging mostly between 20 and 90 min (Almiron-Roig et al., 2013; Rogers et al., 2016). In some studies sensory properties and/or information provided on the food or drink have been manipulated independently of nutrient content (e.g., McCrickerd, Chambers, & Yeomans, 2014). In other studies both nutrient content and visual and oro-sensory properties of the preload vary together (Almiron-Roig et al., 2013), but a common approach is to manipulate nutrient content while, as far as possible, matching the appearance, taste, flavour and texture of the different preloads. Arguably, this tests the effects of the nutrient or other ingredient in question free from the influences of differential expectations and anticipatory responses. A good example of the latter is a large number of studies which have investigated the satiety effects of sugars in drinks and semi-solid foods, using low-calorie sweeteners to control for sweetness in the reduced-sugar or sugar-free comparison preload.

We have recently reviewed these studies (Rogers et al., 2016), and found that across 68 studies (comparisons) test-meal energy intake was reduced on average by 94 kcal after sugar-containing versus low-calorie sweetener control drinks, representing overall 50% compensation for the difference in energy content of the respective drinks. Compensation was somewhat, although not significantly, greater in children (70%) than in adults (43%). Sixty-four of these studies used a cross-over (within-subjects) design. The only studies to use a parallel-groups (between-subjects) design were conducted by Reid and Hammersley (1995; 1999), and these were also unusual in allowing the preload to test-meal interval to

* Corresponding author. School of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol, BS8 1TU, UK.
E-mail address: peter.rogers@bristol.ac.uk (P.J. Rogers).

http://dx.doi.org/10.1016/j.appet.2016.08.113
0195-6663/

vary. Compensation based on the size of the first 'spontaneous meal' ranged across the four studies (comparisons) from −36% to 232% (Rogers et al., 2016). The preference for using a cross-over design would appear to be mainly pragmatic, in that it is recognised that test-meal energy intake across participants is highly variable, so having a design in which participants act as their own control increases power to detect effects of preload nutrient manipulations. Flint, Raben, Blundell, and Astrup (2000), make the same point for appetite ratings. A solution for appetite ratings is to measure appetite immediately before the consumption of the preload ('baseline' rating). The baseline scores can then be used to reduce the 'error' variance in the analysis of the post-preload scores. Obviously, however, this cannot be done for the test meal, as giving a baseline test meal immediately before the preload would substantially alter appetite for the post-preload meal.

Nonetheless, cross-over designs have potential drawbacks, the most worrisome being carry-over effects. In food intake studies it could be that, for example, more is eaten in the test-meal on the second than on the first test occasion due to familiarity with the food and eating environment (Gadah, 2014). In the simplest case of two comparisons that may not be much of a problem because in balanced cross-over designs the order of treatments can be accounted for in the statistical model. However, if there are more than two comparisons this then becomes problematic because relatively small sample sizes are common in studies of this kind. Much more significant though is the possibility of a carry-over effect related to learning following exposure to the preload. In the case where the different preloads (e.g., containing sugar or low-calorie sweetener) are matched in appearance, taste and flavour, it is possible that learned satiation (Brunstrom, 2005; Yeomans, 2012) occurs, such that the satiating effect of the preload consumed on the first test occasion becomes modified and that then influences the satiety that is experienced on the subsequent test occasion (cf. Gibson, Carnell, & Warle, 2006). This means that participants who consumed the sugar drink first and the control (low-calorie sweetener) drink second would experience greater satiety after the control drink than would participants who consumed the control drink first. Symmetrically, these latter participants (control drink first) would experience less satiety after the sugar drink (consumed second) than would the former participants (sugar drink first, control drink second). Put simply, cross-over designs risk underestimating the 'true' satiating effect of the manipulated ingredient (i.e., sugar in the present example).

Due to this uncertainty about the estimation of satiety effects, and therefore energy compensation, in preload-test-meal studies using cross-over designs, we conducted comparable parallel-groups and cross-over studies testing the effects of sucrose versus low-calorie sweetener (sucralose) in a drink preload on test-meal energy intake. We predicted that (1) the parallel-groups design study would demonstrate overall a larger effect of sucrose than the cross-over study, and that (2) in the cross-over study the effect of sucrose versus sucralose measured on the first test day (equivalent to the parallel-groups study) would be larger than that measured on the second test day. In the cross-over study the two test days were scheduled one-week apart. In the parallel-groups study we measured test-meal energy intake on a baseline day (no preload), scheduled the day before the test day. We also included a third preload condition in the parallel-groups study, namely water equal in volume to the sucrose and sucralose drinks. This enabled us to test (3) the prediction that sweetness (i.e., sucralose versus water) would not increase test-meal energy intake (Rogers et al., 2016). In both studies the preload to test-meal interval was 20 min. This was based on the observation of a maximum difference in hunger occurring between 15 and 30 min after sucrose versus control drink preloads (Anderson & Woodend, 2003), which is consistent with our subsequent finding of a peak in blood glucose concentration 20 min after consumption of a sucrose drink (the same drink used in the current studies) (Gadah, Kyle, Smith, Brunstrom, & Rogers, 2016).

2. Methods

2.1. Participants

Approximately equal numbers of healthy men and women who were 18–62 years old were recruited via volunteer databases, membership of which comprised members of the general public in Bristol and students and staff at the University of Bristol. Exclusion criteria were (1) currently dieting, (2) dieted > two times in the past year, (3) vegetarian or vegan, (4) having a food 'allergy' or 'sensitivity,' (5) did not like the test foods, (6) smoked >5 cigarettes/week or equivalent, and (7) doing >225 min/week vigorous physical activity and/or >445 min/week moderate physical activity. Sixty-nine participants (33 men and 36 women) completed the cross-over study and 72 (36 men and 36 women) competed the parallel-groups study. All participants gave signed consent prior to participating in the respective studies, and they were rewarded with £10 or course credits (psychology students) for taking part.

The study protocols were approved by the University of Bristol, Faculty of Science Human Research Ethics Committee.

2.2. Design

The designs of the two studies are summarised in Fig. 1. Within the constraint that there would be approximately equal or equal numbers of men and women in each treatment, participants in the cross-over were assigned randomly to receive either the sucralose drink or the sucrose drink preloads first, and in the parallel groups study participants were assigned randomly to receive one of the three preloads (water, or sucralose or sucrose drinks) on the test day. The cross-over study comprised two sub-studies, with 33 participants (15 men and 18 women) tested in the first sub-study and 36 participants (18 men and 18 women) tested in the second sub-study. There were some differences in the test meal between these two sub-studies (see below). The parallel-groups study was conducted after the first and before the second of the two sub-studies of the cross-over study.

Sample sizes for the sucrose versus sucralose comparisons (cross-over study $n = 69$, parallel-groups study $n = 48$) were determined from effect sizes observed in previous studies (Gadah et al., 2016; Rogers et al., 2016). We assumed 50% compensation in the cross-over study (Rogers et al., 2016) and, based on the arguments outlined in the Introduction, somewhat greater compensation of 60% in the parallel-groups study. We assumed the same standard deviation of energy intake for the two studies, as both studies control for between subject differences in intake in their designs. Based on effect sizes of $d = 0.4$ and 0.5 for the cross-over study and the parallel-groups study respectively, power for a 1-tail test at $\alpha = 0.05$ to reject the null hypothesis of no treatment effect was approximately 80% for both studies.

In both the cross-over and parallel-groups studies each participant attended on two occasions. In the cross-over study they consumed either the sucralose or sucrose drink on the first test day and the other drink on the second test day (one week apart). In the parallel-groups study participants first consumed the *ad libitum* test-meal without having consumed a preload (baseline day) and then, on the following day, they consumed their assigned preload (water, or sucralose or sucrose drinks) followed by the same test meal. Primary outcomes were (1) total energy consumed in the test meal on the test day (in the parallel-groups study adjusted for total

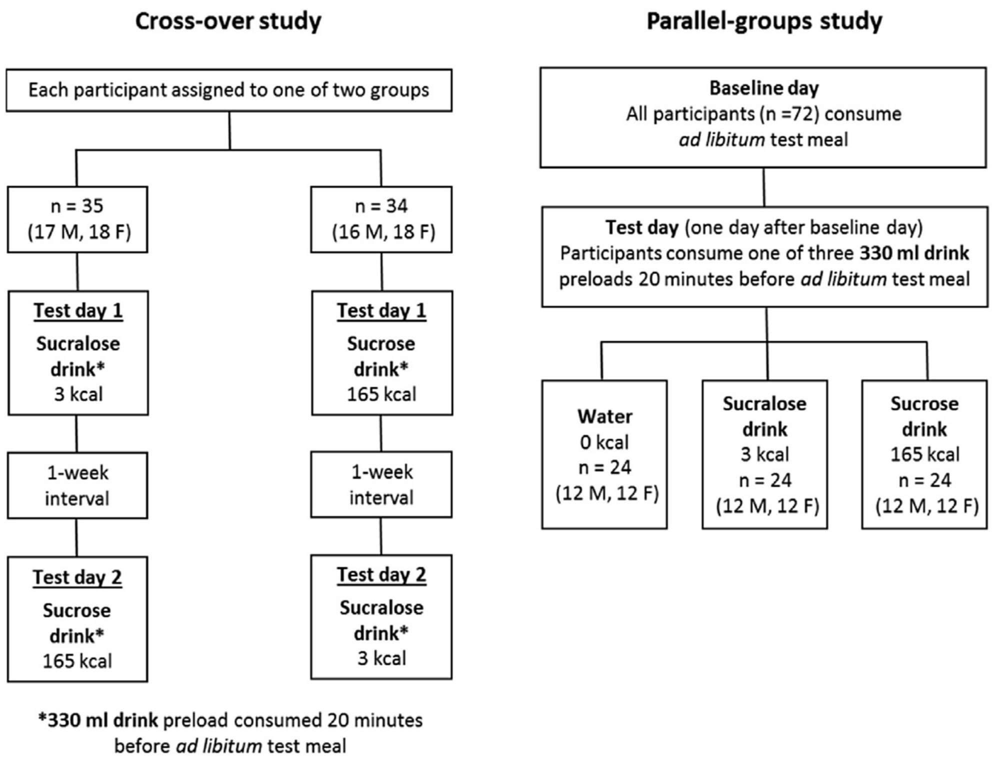

Fig. 1. Summary of the designs of the two studies.

energy consumed on the baseline day), and (2) hunger and fullness rated during the interval between the preload and test meal, adjusted for hunger or fullness rated 5 min before consuming the preload.

2.3. Preloads

The sweet drink preloads were based on supermarket brand, ‘dilute to taste’ blackcurrant squashes, namely Sainsbury's no added sugar, double concentrate blackcurrant squash (sweetened with sucralose) and Sainsbury's high juice blackcurrant squash (sweetened with sucrose) (Sainsbury's Supermarkets Ltd, London, UK). These were diluted as shown in Table 1, and small amounts of black food dye and thickener were added to the sucralose drink to match the colour and mouthfeel of the sucrose drink. Tap water was used for the water preload. These preloads were served in a glass, slightly chilled at between 15 °C and 17 °C.

2.4. Test meal

The test meal consisted of cheese sandwiches and ham sandwiches and other items. The sandwiches were made from crustless bread comprising 50% white and 50% wholemeal flour (Kingsmill 50/50 Crust Away Bread, Allied Bakeries, Maidenhead, UK), spread (Sainsbury's Butterlicious) and medium Cheddar cheese or honey roast ham. The sandwiches were cut into in small triangular pieces, each of which could be consumed comfortably in two bites, and served with a small amount of salad garnish (lettuce, without dressing). In the first sub-study of the cross-over study each participant was served 12 cheese (117 g) and 12 ham sandwich (132 g) triangles, along with 65 g Pringles original potato crisps, 108 g chocolate chip cookies, 250 g Onken fat-free strawberry yogurt and 200 g white grapes. The total energy content of this meal was 2383 kcal. The test meal in the second sub-study of cross-over study and in the parallel-groups study comprised 28 cheese (273 g) and 28 ham sandwich (308 g) triangles served along with a creamy yogurt dessert (80 g Sainsbury's double cream and 400 g Onken fat-free strawberry yogurt). The total energy content of this test meal was 2090 kcal. A glass of water (300 ml) was served with the meals.

2.5. Appetite and preload ratings

Participants rated their hunger and fullness on 100 mm line scales anchored on the left with the words ‘not at all’ (= score of 0) and on the right with the word ‘extremely’ (= score of 100) (Rogers & Hardman, 2015). Participants also rated their liking for (‘How much did you like the taste of the product?’) and familiarity with

Table 1
Composition of the preloads.

Preload	Juice, ml	Water, ml	[a]Thickening agent, g	Amount served, ml	Sugars content, g	[b]Energy value, kcal
Water	0	300	0	300	0	0
Sucralose drink	50 (50 g)	250	0.36	300	0.7	3
Sucrose drink	89 (105 g)	211	0	300	41.3	165

Also, 0.05 g black food dye was added to the sucralose preloads to match the darker colour of the sucrose preloads.
[a] Carrageenan, FMC Biopolymer, Brussels, Belgium, 1.24 kcal/g (0.1 g sugars and 0.42 g fibre).
[b] Sugars (4 kcal/g) and thickening agent. Sugars content was confirmed by Dionex ion chromatography analysis, conducted by British Sugar plc.

('How familiar was the product to you?') the sucralose and sucrose drink preloads, how filling they found them ('How filling did you find the product?'), as well as various oro-sensory attributes: thickness (not evaluated in the first sub-study of the cross-over study), sweetness and fruitiness. As for hunger, these ratings were made on 100 mm line scales anchored with the words 'not at all' and 'extremely.'

2.6. Procedures

On the first test day of the cross-over study, participants arrived at 11:55 h and gave their informed consent to participate. On the second test day they arrived at midday. On both days they completed hunger and fullness ratings at 12:05 h before being served the preload at 12:10 h. They were instructed to consume the drink preload within 5 min (i.e., by 12:15 h) and then to complete the various oro-sensory and other evaluations of the preload. Participants then rated their hunger and fullness again at 12:20 h, 12:25 h, and 12:35 h. The test meal was served immediately after they completed the ratings at 12:35 h. At the end of the second test day the weight and height of the participants was measured.

On the baseline day of the parallel-groups study, participants arrived at the laboratory at 12:15 h. After giving their informed consent to participate, they completed hunger and fullness ratings at 12:30 h and were served the test meal at 12:35 h. Their schedule for the second test day was the same as that for the second test day of the cross-over study.

For both studies participants were instructed to keep to their usual routine of physical activity, eating and drinking the evening before and on the morning of testing up to 9:00 h. They were told thereafter they should not consume any food or drink, except water, before the start of their test session. They were told that they could consume water up to 11:00 h. Participants were tested in groups of up to six, with each participant seated in a private booth within a larger room.

2.7. Data analysis

Data analysis was performed using IBM SPSS Statistics 23. We analysed the data on test-meal energy intake in the cross-over study using mixed factor ANOVA, with drink preload (sucralose and sucrose) as a within-subjects factor, and gender, drink 'order' (sucralose first test day, sucrose second test day and sucrose first test day, sucralose second test day) and sub-study (first sub-study and second sub-study) as between-subject factors. Note that drink order is actually drink x test day, so the main effect of day in this model is given by the drink × order interaction (i.e., drink x drink x test day). We included sub-study in the analysis as the test meal differed between the two sub-studies. We repeated this analysis for cumulative energy intake (i.e., drink plus test-meal energy intake). We were interested in the main effect of drink and its interaction with gender, order and sub-study.

For the parallel-groups study we were primarily interested in contrasting the sucralose drink with the sucrose drink, as in the cross-over study. We therefore analysed the data on test-meal energy intake using ANCOVA with drink (sucralose and sucrose) and gender as between subjects factors. Energy intake on the baseline day was included as a covariate to adjust for individual differences in food intake. Again, we did the same for cumulative energy intake. We then repeated the analyses of test-meal energy intake with (1) water, sucralose and sucrose, and then (2) water and sucralose included in the drink factor.

We report effect sizes (partial eta^2) for the effect of preload drink on energy intakes. As discussed by Levine and Hullett (2002), despite certain limitations, partial eta^2 can be more comparable than other measures of effect size when evaluating the size of an effect of the same manipulation across studies when, as in the present studies, additional manipulated and control variables are added to the design.

We calculated an energy intake compensation score (COMPX) (Cecil et al., 2005) for sucralose versus sucrose drinks. For the cross-over study this is simply sucralose minus sucrose test-meal energy intake divided by the difference (162 kcal) in the energy content of the sucralose and sucrose drinks multiplied by 100. For the parallel-groups study we calculated a compensation score that took account of baseline test meal energy intake. The equation we used was: $\text{COMPX} = \left(\frac{(x_{sucralose} - y_{sucralose}) - (x_{sucrose} - y_{sucrose})}{162}\right) \times 100$, where x = test day test-meal energy intake, y = baseline day test-meal energy intake, and 162 is the difference between the energy content of the sucrose and sucralose preloads. COMPX describes the extent to which adjustment in test-meal energy intake 'compensates' for the difference in energy content of the sucralose versus sucrose preload. If COMPX is <100% there is under-compensation for the greater energy content of the sucrose preload (higher cumulative energy intake). Note that throughout we report absolute mean and mean differences in energy intakes, rather than adjusted means, because they are the basis of the calculation of energy compensation.

For the analysis of hunger and fullness we first calculated difference from baseline scores, for example, hunger rated 5 min after consuming the drink minus hunger rated 5 min before (baseline) consuming the drink. We then analysed the difference scores using mixed model ANOVA, with drink, time (5, 10 and 20 min ratings), gender, drink order and sub-study as factors for the cross-over study, and time (5, 10 and 20 min ratings), drink (sucralose and sucrose) and gender, for the parallel-groups study.

For both studies we compared the ratings of liking, sweetness, etc. for the sucralose versus sucrose drink using *t*-tests.

3. Results

3.1. Cross-over study

Participants' mean ± SD age was 26.1 ± 10.4 years (men = 26.3 ± 11.4, women = 25.9 ± 9.5), their weight was 67.9 ± 13.1 kg (men = 76.4 ± 11.3, women = 60.1 ± 9.3), their BMI was 22.9 ± 3.2 kg/m^2 (men = 23.8 ± 3.1, women = 22.0 ± 3.1), and their DEBQ restraint score (minimum and maximum possible scores are 1 and 5) was 2.21 ± 0.83 (men = 1.78 ± 0.61, women = 2.61 ± 0.80).

Participants liked the two drinks equally (mean ± SE mm: sucralose = 60.2 ± 2.7, sucrose = 60.6 ± 2.9, $t(68) < 1$). They also rated them as equally thick (sucralose = 30.4 ± 4.1, sucrose = 32.1 ± 4.1, $t(35) < 1$), equally fruity tasting (sucralose = 64.1 ± 2.3, sucrose = 65.4 ± 2.8, $t(35) < 1$), equally filling (sucralose = 50.3 ± 2.7, sucrose = 47.3 ± 3.3, $t(35) < 1$) and equally familiar (sucralose = 66.4 ± 3.3, sucrose = 64.6 ± 4.1, $t(35) < 1$), but they rated the sucralose-sweetened drink as less sweet than the sucrose-sweetened drink (sucralose = 68.2 ± 2.0, sucrose = 75.4 ± 1.7, $t(68) = 3.34$, $p = 0.001$). Sucralose minus sucrose sweetness, however, was found not to be a significant covariate in analyses of the effects of drink on hunger, fullness or test-meal energy intake. Therefore the analyses were re-run without this covariate, and the results of those analyses are reported below.

There were main effects of drink, gender and sub-study on test-meal energy intake. Energy intake was lower after the sucrose- than after the sucralose-sweetened drink ($F(1,61) = 5.73$, $p = 0.020$, partial eta^2 = 0.086, Fig. 2 left-hand panel), lower in women than in men (mean ± SE: 706 ± 50 versus 1113 ± 58 kcal, $F(1,61) = 31.31$, $p < 0.001$), and lower in the second versus the first sub-study (811 ± 55 versus 999 ± 70 kcal, $F(1,61) = 7.90$, $p = 0.007$). There

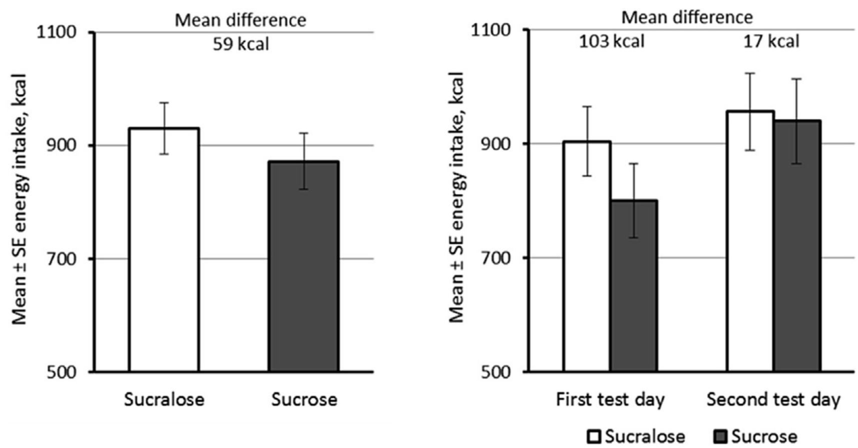

Fig. 2. In the cross-over (n = 69) study the sucrose-sweetened drink reduced test meal energy intake compared with the sucralose-sweetened drink. The overall difference was 59 kcal ($p = 0.020$, left-hand panel), however that difference was due mainly to the difference in the effect of the drinks measured on the first of the two test days (right hand panel).

was no interaction between the effect of drink and sub-study ($F < 1$) or the effect of drink and gender ($F < 1$) for test-meal energy intake.

There was also a main effect of test day on energy intake in the test meal. Energy intake was lower on the first than on the second test day (mean ± SE: 853 ± 44 versus 949 ± 50 kcal, $F(1,61) = 14.62$, $p < 0.001$). Additionally, the right hand panel of Fig. 2 shows that the difference in test-meal energy intake between the sucralose and sucrose drinks on the first test day was 103 kcal compared with 17 kcal on the second test day. However, neither of these (between subjects) sucralose versus sucrose differences in test-meal energy intake was statistically significant ($F(1,61) = 1.44$, $p > 0.1$, partial eta$^2 = 0.023$, and $F < 1$, partial eta$^2 = 0.003$, respectively), nor was the drink x test-day interaction statistically significant ($F < 1$).

The difference in test-meal energy intake after the two drinks (59 kcal) amounted to roughly a third of the difference in their energy content (COMPX = 36%). Consequently, cumulative (i.e., preload plus test-meal) energy intake was greater when the sucrose drink was consumed (mean ± SE: 1037 ± 50 kcal) compared with when the sucralose drink was consumed (933 ± 45 kcal) ($F(1,61) = 13.44$, $p < 0.001$, partial eta$^2 = 0.181$).

Baseline hunger did not differ between drink conditions (mean ± SD sucralose = 63.2 ± 20.1 mm and sucrose = 61.8 ± 20.0 mm $F < 1$) or test day (first test day = 61.1 ± 19.8 mm and second test day = 62.9 ± 20.1 mm, $F < 1$). Similarly, baseline fullness did not differ between drink conditions or between test day. There was no main effect of drink on hunger or fullness (largest $F(1,61) = 1.12$, $p > 0.1$ for hunger), and no drink by time effect for hunger or fullness (largest $F(2,136) = 1.22$, $p > 0.1$ for hunger) on post-preload ratings. Mean ± SD hunger and fullness ratings made 20 min after the preload (i.e., just before the commencement of the test-meal) were as follows: hunger, sucralose = 64.2 ± 17.9 mm and sucrose = 58.2 ± 20.1 mm; fullness, sucralose = 32.0 ± 19.3 mm and sucrose = 39.9 ± 19.1 mm.

3.2. Parallel-groups study

Participants' mean ± SD age was 22.5 ± 4.0 years (men = 21.6 ± 3.4, women = 23.4 ± 4.3), their weight was 65.4 ± 10.5 kg (men = 69.5 ± 9.1, women = 61.4 ± 10.4), their BMI was 22.1 ± 2.6 kg/m^2 (men = 22.1 ± 2.1, women = 22.2 ± 3.0), and their DEBQ restraint score (minimum and maximum possible scores are 1 and 5) was 2.05 ± 0.52 (men = 1.88 ± 0.57, women = 2.23 ± 0.40).

Participants liked the two sweet drinks equally (mean ± SE mm: sucralose = 61.8 ± 5.0, sucrose = 61.8 ± 4.9, $t(46) < 1$). They also rated them as equally sweet (sucralose = 77.4 ± 2.8, sucrose = 82.2 ± 2.6, $t(46) = 1.28$, $p > 0.1$), equally fruity tasting (sucralose = 57.9 ± 5.3, sucrose = 57.5 ± 4.8, $t(46) < 1$), equally filling (sucralose = 47.4 ± 4.6, sucrose = 52.8 ± 4.2, $t(46) < 1$) and equally thirst-quenching (sucralose = 57.7 ± 4.7, sucrose = 52.1 ± 4.5, $t(46) < 1$).

Baseline day energy intakes and test day test-meal energy intakes for the three drink groups are shown in Table 2. ANCOVA (baseline day energy intake as the covariate) comparing the effect of the two sweet drinks on test day test-meal energy intake revealed a main effect of drink ($F(1,43) = 4.26$, $p = 0.045$, partial eta$^2 = 0.090$), no effect of gender ($F < 1$, the effect of gender on energy intake is controlled for by the covariate) and no drink × gender interaction ($F(1,43) = 1.81$, $p > 0.1$). Test-meal minus baseline energy intakes for the two sweet drinks are shown in Fig. 3. These differences sum to 104 kcal (COMPX = 64%). Including preload energy content in these calculations showed that cumulative energy intake was higher after the sucrose-sweetened drink (58 kcal), but not significantly so ($F < 1$, partial eta$^2 = 0.015$).

Although not statistically significant ($F(1,44) = 2.42$, $p > 0.1$, gender included in the model), the sucrose group displayed considerably lower energy intake on the baseline day than did the sucralose group (Table 2). Inspection of the distributions of these energy intakes revealed two sucralose group participants whose baseline intakes (1689 and 2098 kcal) substantially exceeded the baseline intakes of any of the sucrose group participants (largest = 1120 kcal). Furthermore, the lowest baseline energy intake (369 kcal) was displayed by a sucrose group participant (lowest sucralose intake = 520 kcal). Repeating the analysis of test-meal energy intake with these three participants excluded (mean ± SD baseline day energy intakes: sucralose = 810 ± 244, sucrose = 782 ± 209) confirmed the main effect of drink ($F(1,40) = 6.68$, $p = 0.013$, partial eta$^2 = 0.14$), summing to 134 kcal (COMPX = 83%).

Table 2
Baseline day and test day test-meal energy intakes (kcal) for the three drink groups.

	Water	Sucralose	Sucrose
Baseline day	[a]914 ± 379	900 ± 389	765 ± 221
Test day	1002 ± 412	960 ± 430	721 ± 224
Test minus baseline day	88 ± 227	60 ± 176	−44 ± 203

[a] Data are means ± SDs.

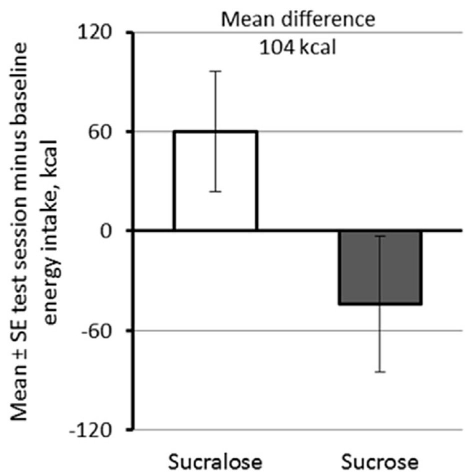

Fig. 3. In the parallel-groups study (n = 24 per group), compared with the baseline day, test-meal energy intake on the test day increased after the sucralose-sweetened drink and decreased after the sucrose-sweetened drink. The overall difference was 104 kcal ($p = 0.045$).

Including all three drink groups, water, sucralose and sucrose (no participants excluded) in the analysis of test-meal energy intake, revealed a main effect of drink ($F(1,65) = 3.61$, $p = 0.033$, partial eta$^2 = 0.100$), no effect of gender ($F < 1$, the effect of gender on energy intake is controlled for by the covariate) and no drink × gender interaction ($F(1,43) = 1.09$, $p > 0.1$). However, test-meal energy intake did not differ between water and sucralose conditions (difference between test day test-meal energy intake and baseline day energy intake for sucralose versus water = −28 kcal, $F < 1$, partial eta$^2 = 0.006$, Table 2).

In all of the above analyses of energy intakes relevant baseline day energy intake was a highly significant predictor of test day energy intake ($p < 0.001$).

In analyses comparing the three drinks, neither baseline hunger nor fullness on the test day differed between drink conditions. Results (mean ± SD) for hunger were: water = 70.8 ± 15.4, sucralose = 68.4 ± 18.6, sucrose = 68.8 ± 15.8 ($F < 1$). For fullness they were: water = 11.8 ± 12.7 sucralose = 16.7 ± 18.5, sucrose = 20.2 ± 17.9 ($F(1,69) = 1.71$, $p > 0.1$). There was no main effect of drink on post-preload hunger or fullness ($F < 0.1$), and no drink by time effect for hunger or fullness ($F < 0.1$). Mean ± SD hunger and fullness ratings made 20 min after the preload (i.e., just before the commencement of the test-meal) were as follows: hunger, water = 72.3 ± 15.5 mm and sucralose = 64.7 ± 25.1 mm and sucrose = 64.3 ± 21.6 mm; fullness, water = 21.1 ± 18.5 mm and sucralose = 30.9 ± 26.0 mm and sucrose = 33.4 ± 21.9 mm.

4. Discussion

The results of these studies support our hypothesis that using a cross-over design in preload-test-meal studies can lead to underestimation of effects of preload manipulations on energy intake. Consistent with our first prediction, the effect of consuming sucrose in a drink on subsequent energy intake was greater when measured in the parallel-groups study than in the cross-over study (64% versus 36% compensation). Furthermore, consistent with our second prediction, the effect of sucrose measured on the first test day in the cross-over study was larger than the effect measured on the second test day. Indeed, remarkably, the first day effect (absolute difference between sucrose and sucralose) in the cross-over study was almost identical in size to the effect of sucrose measured in the parallel-groups study (103 kcal versus 104 kcal, Fig. 2 right-hand panel and Fig. 3). This similarity is to be expected as these conditions are equivalent in that they are both 'uncontaminated' by prior consumption of a preload that tasted the same but differed in sucrose content. However, for the parallel-groups study the effect was statistically significant whereas the first-day effect for the cross-over study was not. This too was expected because, compared with the cross-over study, the influence of individual differences on test-day energy intake was reduced in the parallel-groups study by adjusting for baseline day energy intake. Note that, if anything, sample sizes for these comparisons (sucrose versus sucralose) favoured the cross-over study (n = 69) over the parallel-groups study (n = 48).

The different estimates of the effects of sucrose on test-meal energy intake in the present studies also accord very well with effects observed in similar previous research. In our meta-analysis of preload-test-meal studies comparing sugars and low-calorie sweeteners (Rogers et al., 2016) we found the average compensation for adults across all studies to be 43%, with a 95% confidence interval of 31%–55%. The vast majority of those studies used a cross-over design, and the effect observed in the present cross-over study (36% compensation) falls within that confidence interval. In contrast, the estimate of compensation from the present parallel-groups study (64%) is close to the value for compensation for sucrose in a drink (60%) we found in another very similar parallel-groups study, which additionally investigated effects of preload viscosity on energy compensation (Gadah et al., 2016). Neither of those values are included within the confidence interval for compensation we calculated for the earlier predominantly cross-over studies (Rogers et al., 2016).

We included plain water as a third preload condition in the parallel-groups study, and our third prediction, based on the results of our review of the short-term effects of low-calorie sweeteners (Rogers et al., 2016), was that test-meal energy intake would not differ between the water and sucralose conditions. This prediction was upheld. Test-meal energy intake was 28 kcal lower after sucralose compared with water but this difference was not statistically significant. This difference falls within the confidence interval estimated from our meta-analysis of 35 comparisons of low-calorie sweeteners with water: mean difference (low-calorie sweeteners minus water) was −2 kcal and the confidence interval was −30 kcal to 26 kcal (Rogers et al., 2016). Thirty-three of those comparisons were cross-over studies. Together, these results are important in demonstrating that exposure to sweetness in a drink does not increase short-term energy intake. They also show similar estimates for the effects of low-calorie sweeteners versus water derived from cross-over and parallel-groups studies. This is to be expected as, unlike the comparison between sugars and low-calorie sweeteners, water and a sweet drink are highly discriminable by their taste and (usually) flavour and appearance, so any preload-related carry-over effects in cross-over studies ought to be minimal.

Another finding from the present cross-over study was that test-meal energy intake was higher (by 96 kcal) on the second test day than a week earlier on the first test day. It is unclear what the explanation for this is, but it could be an effect of increased familiarity with the test-meal, the study environment and/or the study protocol. Participants were equally hungry before consuming the preload on test days one and two, suggesting that their experience on their first test day did not cause them to reduce their food intake at breakfast in order to be able to eat more of the test-meal food on test day two. However, perhaps their higher intake on the second test day was planned in anticipation of eating less later in the day. Whatever the explanation for the test-day effect on energy intake, it might have contributed to the smaller effect of sucrose versus sucralose on test day two (Fig. 2 right-hand panel). Perhaps higher test-meal energy intake reduces sensitivity to pre-meal differences in satiety. Against this, however, is the finding that the sucrose versus sucralose effect on test-meal energy intake did not differ between the two sub-studies (drink x sub-study interaction, $F < 1$), despite the fact that overall test-meal energy intake did differ substantially (by 188 kcal) between these sub-studies. The higher test-meal energy intake in the first sub-study is presumably explained by the greater variety of food and somewhat larger portion size provided in that meal (Hetherington, 2007; Norton, Anderson, & Hetherington, 2006; Zlatevska, Dubelaar, & Holden, 2014). This leaves a carry-over effect linked to the preloads as the most likely explanation for the smaller sucrose versus sucralose effect on test-day two.

We predicted a carry-over effect based on the argument that energy intake on the second test day is affected by learning about the satiating consequences of the preload (plus test meal) consumed on the first test day. Gibson et al. (2006) make the same suggestion, and provide supporting evidence from a preload-test-meal cross-over study in preschool children. Eighty-five children consumed 200 ml sensorily-matched low-energy (5 kcal) and high-energy (174 kcal, added maltodextrin) fruit-flavoured drink preloads. The test meal was a buffet lunch served 30 min after the preload. Lunch intake was lower by 64 kcal after the low-energy preload when it was consumed on the second test day (i.e., after the high-energy preload had been experienced on the first test day) than when it was consumed on the first test day. However, energy intake after the high-energy preload did not differ between test days.

Although this result, like ours, is consistent with an influence of learned satiation, neither study directly evaluates learning. Demonstration of learned satiation requires evidence that a flavour previously paired with high or low energy intake (satiation) affects intake or fullness independently of the energy content of the product (drink or food). In practice, that means, for example, contrasting the satiating effects of two products of the same intermediate energy content, differing only in flavour – one flavour having been consumed previously in a high-energy version of the product and the other having been consumed in a low-energy version (e.g., Birch & Deysher, 1985; Booth, Mather, & Fuller, 1983). A problem is that the evidence from such studies suggests that learned satiation is not a particularly robust phenomenon, at least in humans (Brunstrom, 2005; Yeomans, 2012). So, while learned satiation may have contributed to our findings and those of Gibson et al. (2006), other explanations need to be considered as well. A possibility is that consumption of the preload and test meal on the first test day sets up a personal norm (cf. Lewis et al., 2015) for intake of what is, in terms of appearance, taste and flavour, the same drink and food presented on the second test day, and it is that norm that guides intake on the second day, working against the nutrient-related difference in physiological state arising from consumption the higher- or lower-energy preload. If this was the case, then in our study it was combined with a separate tendency to eat more overall on the second test day, which might have been related to plans for further eating later in the day (see above). If nothing else this testifies to the complexity of potential carry-over effects introduced by using cross-over designs.

In the parallel-groups study, test-meal energy intake on the baseline day was somewhat, though not significantly, lower in the sucrose condition than in either the water or sucralose conditions. Participants were randomised to conditions, with the constraint of having equal numbers of men and women in each condition, so the somewhat lower baseline energy intake in the sucrose condition participants is presumably due to chance. This could have been avoided had we allocated participants across conditions based on matching baseline energy intakes. Nevertheless, we are confident that the effect of sucrose in reducing test-day energy intake measured in the study is robust. If anything, regression to the mean (whereby under- or over-estimates of energy intake at baseline move towards the true value on the test day) would predict that a lower energy intake at baseline in the sucrose versus other conditions would lead to an underestimation of compensation. This is supported by our *post hoc* analysis which found that removing participants with, respectively, particularly low and particularly high baseline energy intakes from the sucrose and sucralose conditions actually increased the estimate of the compensation effect to 83%.

We aimed to match the sucralose and sucrose drinks in appearance and oro-sensory qualities, and the results of the participant evaluations of the drinks showed that we succeeded in this, except for a small difference in sweetness in the cross-over study. However, as sweetness was not found to predict test-meal energy intake, it would seem that the slightly higher sweetness of the sucrose preload cannot account for its effect on energy intake compared with sucralose.

Neither the cross-over nor parallel-groups study revealed effects of sucrose versus sucralose on hunger or fullness. This was unexpected, as previous studies (e.g., Almiron-Roig & Drewnowski, 2003; Anderson & Woodend, 2003; Rogers, Carlyle, Hill & Blundell, 2008), including our recent similar parallel-groups study (Gadah et al., 2016), have demonstrated decreased hunger and increased fullness after consumption of sugar in a drink. The finding of a difference in hunger between sucrose and low-energy sweetener after around 10–20 min but little difference more immediately post-consumption is consistent with a post-ingestive action of sucrose (in the context of oro-sensory stimulation) rather than an oro-sensory effect alone (Gadah et al., 2016). Why we did not observe this pattern in the present studies is unexplained; however, as expected, we did find effects of sucrose on energy intake, which is the important outcome in respect of implications for energy balance.

Taken together, these various results demonstrate unequivocally that sucrose consumed in a drink reduces subsequent energy intake. This effect, however, is not sufficient to prevent increased cumulative energy intake, even when measured using a parallel-groups design (non-significant increase of 58 kcal), and optimising the timing of the test-meal interval to coincide with the peak effect of the preload on blood glucose concentration (Gadah et al., 2016). On the other hand, we hid the manipulation of the sucrose content of the drink from participants, as has been done in almost all of the other similar studies (Rogers et al., 2016). In every-day life, awareness of the sugar content of a drink may add to the compensation effect. This could be tested in future preload-test-meal studies by manipulating information about the sugar content of the preload, but to be relevant such studies would need to be conducted with credible labels and within a setting that encourages participants to attend to the information provided.

Finally, in light of our findings we recommend using parallel-groups (with a baseline test day) rather than cross-over study designs when testing the effects of disguised preload manipulations on energy intake. The practical significance lies equally in potentially overestimating the effect of nutrient dilution in reducing energy intake, as in the present example of reducing sucrose content, and underestimating the effect of, for example, a novel ingredient designed to increase satiety, with the result that it might be mistakenly abandoned as ineffective.

Acknowledgements and disclosures

We thank Lesley Kyle and Dr Jessica Smith for their assistance with data collection for this study. This research was funded by Sugar Nutrition Research UK (grant reference number 47190) and the King Abdullah Scholarship Program (NSG) (grant reference number S7283). Part of the work leading to the completion of this paper received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 607310 (JMB and PJR). Results from first sub-study of the cross-over study were presented at the Winter Meeting of the Nutrition Society, held in London, UK, December 2011, and a summary of the additional results was presented at the 38th Annual Meeting of the British Feeding and Drinking Group, held in Portsmouth, UK, April 2014. Peter Rogers has provided consultancy services to Coca-Cola Great Britain and received speaker's fees from the International Sweeteners Association.

References

Almiron-Roig, E., & Drewnowski, A. (2003). Hunger, thirst, and energy intakes following consumption of caloric beverages. *Physiology and Behavior, 79*, 767–773.

Almiron-Roig, E., Palla, L., Guest, K., Ricchiuti, C., Vint, N., Jebb, S. A., et al. (2013). Factors that determine energy compensation: A systematic review of preload studies. *Nutrition Reviews, 71*, 458–473.

Anderson, G. H., & Woodend, D. (2003). Consumption of sugars and the regulation of short-term satiety and food intake. *American Journal of Clinical Nutrition, 78*, 843S–849S.

Birch, L. L., & Deysher, M. (1985). Conditioned and unconditioned caloric compensation: Evidence for self-regulation of food intake in young children. *Learning and Motivation, 16*, 341–355.

Booth, D. A., Mather, P., & Fuller, J. (1983). Starch content of ordinary foods associatively conditions human appetite and satiation, indexed by intake and eating pleasantness of starch-paired flavours. *Appetite, 3*, 163–184.

Brunstrom, J. E. (2005). Dietary learning in humans: Directions for future research. *Physiology and Behavior, 85*, 57–65.

Cecil, J. E., Palmer, C. N., Wrieden, W., Murrie, I., Bolton-Smith, C., Watt, P., et al. (2005). Energy intakes of children after preloads: Adjustment, not compensation. *American Journal of Clinical Nutrition, 82*, 302–308.

Flint, A., Raben, A., Blundell, J. E., & Astrup, A. (2000). Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. *International Journal of Obesity, 24*, 38–48.

Gadah, N. S. (2014). *Liquid calories: An investigation of the satiety effects of sugar-sweetened beverages*. PhD thesis. UK: University of Bristol.

Gadah, N. S., Kyle, L. A., Smith, J. A., Brunstrom, J. M., & Rogers, P. J. (2016). No difference in compensation for sugar in a drink versus sugar in semi-solid and solid foods. *Physiology and Behavior, 156*, 35–42.

Gibson, E. L., Carnell, S., & Warle, J. (2006). Learnt satiation in effects of disguised preloads on lunch intake and appetite in schoolchildren. *Appetite, 47*, 281.

Hetherington, M. M. (2007). Cues to overeat: Psychological factors influencing overconsumption. *Proceedings of the Nutrition Society, 66*, 113–123.

Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squatted, and misreporting of effect size in communication research. *Human Communication Research, 28*, 612–625.

Lewis, H. B., Forwood, S. E., Ahern, A. L., Verlaers, K., Robinson, E., Higgs, S., et al. (2015). Personal and social norms for food portion sizes in lean and obese adults. *International Journal of Obesity, 39*, 1319–1324.

McCrickerd, K., Chambers, L., & Yeomans, M. R. (2014). Fluid or fuel? The context of consuming a beverage is important for satiety. *PLoS One, 9*, e1000406.

Norton, G. N. M., Anderson, A. S., & Hetherington, M. M. (2006). Volume and variety: Relative effects on food intake. *Physiology and Behavior, 87*, 714–722.

Reid, M., & Hammersley, R. (1995). Effects of carbohydrate intake on subsequent food intake and mood state. *Physiology and Behavior, 58*, 421–427.

Reid, M., & Hammersley, R. (1999). The effects of sucrose and maize oil on subsequent food intake and mood state. *British Journal of Nutrition, 82*, 447–485.

Rogers, P. J., & Hardman, C. A. (2015). Food reward: What it is and how to measure it. *Appetite, 90*, 1–15.

Rogers, P. J., Carlyle, J.-A., Hill, A. J., & Blundell, J. E. (2008). Uncoupling sweet taste and calories: comparison of the effects of glucose and three intense sweeteners on hunger and food intake. *Physiology and Behavior, 43*, 547–552.

Rogers, P. J., Hogenkamp, P. S., de Graaf, C., Higgs, S., Lluch, A., Ness, A. R., et al. (2016). Does low-energy sweetener consumption affect energy intake and body weight? A systematic review, including meta-analyses, of the evidence from human and animal studies. *International Journal of Obesity, 40*, 381–394.

Yeomans, M. R. (2012). Flavour-nutrient learning in humans: An elusive phenomenon? *Physiology and Behavior, 106*, 345–355.

Zlatevska, N., Dubelaar, C., & Holden, S. S. (2014). Sizing up the effect of portion size on consumption: A meta-analytic review. *Journal of Marketing, 78*, 140–154.