

KU LEUVEN

DEPARTMENT OF ECONOMICS

Sugar rush or sugar crash? Experimental evidence on the impact of sugary drinks in the classroom

Fritz SCHILTZ and Kristof DE WITTE

FACULTY OF ECONOMICS AND BUSINESS



DISCUSSION PAPER SERIES DPS21.11

OCTOBER 2021

Sugar Rush or Sugar Crash?

Experimental Evidence on the Impact of Sugary Drinks in the Classroom.

Fritz Schiltz and Kristof De Witte

Abstract

Sugary drinks in schools have been demonized for their potential long-term contribution to rising obesity rates. Surprisingly, there is only little evidence on the immediate effects of sugary drinks in schools. This paper provides experimental evidence on the in-class effects of sugary drinks on behavior and student achievement. We randomly assigned 462 preschool children to receive sugary drinks or artificially sweetened drinks and collected data before and after consumption. Our findings suggest that the consumption of one sugary drink induces an initial ‘relaxing’ effect for boys, before making them more restless. Girls’ behavior is not significantly affected. We find a negative effect on student achievement for boys and a positive effect for girls. We show the robustness of the results across two field experiments.

Schiltz (corresponding author): Faculty of Economics and Business, KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium (e-mail: fritz.schiltz@kuleuven.be); De Witte: Faculty of Economics and Business, KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium; UNU-Merit, Maastricht University, Boschstraat 24, 6211 AX Maastricht, the Netherlands (e-mail: kristof.dewitte@kuleuven.be, k.dewitte@maastrichtuniversity.nl). The authors are grateful to Deni Mazrekaj, David Figlio, Michele Belot, Andrea Ichino, Jan Feld, Eric Hanushek, Dan Goldhaber, Anna Vignoles, Dinand Webbink, Nicky Rogge, Erwin Ooghe, Laurens Cherchye, Rachel Griffith, Veronica Minaya, Toni Mora, Hessel Oosterbeek, Erik Plug, Bert De Smedt and seminar participants in Leuven, Florence, and Reus for their helpful comments and suggestions, and to Arnout Baeyens, Tine Bollen, Maxime De Moor, Kaat Iterbeke, Gertjan Muyters, Ruben Van Assche, Giovanna D’Inverno, Vita Titl, and Kenneth Debecker for their research assistance. The authors received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 691676 (EdEN). The authors declare that they have no relevant or material financial interests that relate to the research described in this article. This study is registered in the AEA RCT Registry and the unique identifying number is AEARCTR-0002782.

1 Introduction

Parents and educators care a lot about the diet of their children. At the same time, the percentage of children in the US consuming sugary drinks¹ rose from 79 to 91 percent between 1989 and 2008 (Lasater et al., 2011). Data from the National Health and Nutrition Examination Survey (2007-2010) indicated that 62% of preschool children in the US (aged 2–5 years) consumed sugary drinks on a daily basis. This share increases with age, corresponding to 73% of children aged 6–11 years and 76% of adolescents aged 12–19 years (Bleich and Wolfson, 2015). These consumption patterns can also be seen in Europe, albeit to a much lesser extent. Data from the WHO’s European Health Information Gateway (2014) indicates that 15% of children aged 11 years consume soft drinks at least once a day. This share increases to 20% on average for children aged 15 years. In Belgium, where the data collection took place, more than 1 in 4 children aged 11 years, and more than 1 in 3 aged 15 years, is consuming soft drinks at least once a day.

Concerns are growing that this ‘Western diet’ high in salt, fat, and sugar - an increasingly global diet - immediately affects children’s physical and mental health (e.g. Gómez-Pinilla, 2008). It is a widely held belief that the consumption of sugary drinks causes children to experience a “sugar rush”². The associated potential effects on in-class performance have major policy implications. First, sugary drinks are still ubiquitously sold in schools (Papoutsi et al., 2013). For example, the School Nutrition Dietary Assessment of the USDA reported that one in eight elementary schools had beverage vending machines in 2010, and more than three quarters offered ‘competitive foods’ through a la carte offerings. Moreover, in US middle schools (67%) and high schools (85%), vending machines are omnipresent (Fox and Condon, 2012). School principals are often the staunchest opponents to the removal of vending machines as the profits

¹Caloric soft drinks, fruit drinks, and energy drinks consumed over a three consecutive days period.

²Oxford’s dictionary defines a sugar rush as “the sudden short period of extra energy that you get from eating or drinking a large amount of sugar, for example in sweet food or drinks”.

gained by selling these drinks are vital for many highly underfunded schools (Anderson and Butcher, 2006). From a policy perspective, given the lack of evidence on the effects of sugar on student achievement, there is currently little ground for a potential ban of sugary drinks on this basis. However, if sugar would affect student achievement, the availability of sugary drinks might result in increased test score inequality within and between schools.

Second, achievement differences accumulate over time, stressing the importance of schooling for children at a young age. Moreover, in education systems that use standardized tests to assign students to education tracks, a small change in test results (e.g. due to a sugary drink) might place children on the other side of the threshold, and consequently in another track. Given the large differences in average wages by education tracks (Mazrekaj et al., 2019), the consumption of sugary drinks at a young age might have long run impacts on children's lives.

Third, the consumption of sugary drinks (both in and outside schools) is typically higher among children from low-income households and among boys.³ Potential negative effects will then further widen (or partially explain) the achievement differences by socio-economic status and sex. At the same time, behavioral effects might affect all students in a class, also those who do not consume sugary drinks.

As a result of these alarming consumption patterns, public health advocates have proposed policy measures such as a ban on vending machines in schools, reducing children's exposure to advertisements, and sugar taxes (Scharf and DeBoer, 2016). For example, in Belgium a sugar tax was introduced on all soft drinks (including both artificially sweetened drinks and sugary drinks). Given the similarity in consumption patterns of children, the relevance of the results presented in this paper potentially extend beyond the Belgian, or European, context. Sound evidence on the impact of sugary drinks in a class-context can feed the discussion on dietary guidelines for schools, taxes on sugar and the potential cost of vending machines in schools.

³*Boys drink more sugary beverages than girls, federal survey shows*, Washington Post (2011).

This paper is the first to study the immediate effects of sugary drinks on student achievement and behavior of preschool children. Despite evidence supporting the toxicity of sugar, the potential for abuse and the negative impact on society (Lustig et al., 2012), causal studies linking the consumption of sugary drinks to educational and behavioral outcomes are limited. Previous studies in economics have considered sugar mainly as "empty calories" (Lustig et al., 2012, p.28) and focused on how nutrition as a whole matters for cognition. For example, earlier literature examined the role of breakfast (e.g. Frisvold, 2014; Leos-Urbel et al., 2013) and school meal programs (e.g. Schanzenbach, 2009; Millimet et al., 2010; McEwan, 2013) without targeting immediate effects of sugary drinks. Other studies evaluated the health and educational effects of more healthy schools meals (Anderson et al., 2017; Belot and James, 2011).⁴ Outside the economic literature, reviews on the link between children's diets and educational outcomes remain inconclusive (Sorhaindo and Feinstein, 2006; Bellisle et al., 2004; Ells et al., 2008; Smith et al., 2011; Philippou and Constantinou, 2014).

Although there are reports that sugary drinks (i.e. oral glucose ingestion) can improve some aspects of intellectual performance in the short run for adults (e.g. Smith et al., 2011), the question has been largely ignored for children. Existing studies with adult subjects cannot be generalized to young children. When compared to the size of their body, children have a relatively larger brain than adults. Also, for a given weight of brain tissue, a child's brain tissue utilizes more glucose than an adult's brain tissue (Kalhan and Kilic, 1999). This high glucose utilization rate is observed until the age of 9-10 years. Considering the importance of schooling at a young age (e.g. Heckman, 2008), and these biological differences, there is a need for the assessment of immediate effects on preschool children to guide policy decisions. One example of an existing meta-analysis on the relationship between cognition or behavior of children is

⁴A complementary literature focuses on behavioral economic assessments of food choices (Belot et al., 2016; Levitt et al., 2016; List and Samek, 2015), the effects of tax policies (Fletcher et al., 2010; Tiffin et al., 2015), and on the sugar effect with regards to time preferences and willpower (Kuhn et al., 2014).

the work by Wolraich et al. (1995) bringing together the evidence of 16 reports (covering 23 within-subject studies) on the topic. Overall, the authors find no evidence that sugar affects the cognition or behavior of children. However, the present paper differs from the Wolraich et al. (1995) study in 3 regards, motivating the need for further research on the link between sugar consumption and behavior or cognition of children: (1) Out of 23 studies, only 9 studies focus on ‘normal’ subjects, whereas the majority (14) of the studies focuses on children identified as ‘sugar reactors’, ‘ADHD’ or ‘Psychiatric inpatients’; (2) Out of these 9 studies, 5 studies focus on the group of children aged 5 and below; (3) Out of the 5 studies focusing on ‘normal’ children aged 5 and below, the sample size varies between 8 and 25 subjects. Hence, the generally small sample sizes and the diversity of the populations examined inhibits the authors of this review study from drawing powerful conclusions. Moreover, children (and parents) participating in the controlled trials are often tested in an experimental setting, and on a voluntary basis, which makes it difficult to generalize the findings to a classroom context.

Our study contributes to the literature by investigating whether and to what extent sugary drinks matter for student achievement and behavior of preschool children. The paper most related to our work is by Figlio and Winicki (2005), who provide quasi-experimental evidence on the immediate effect of nutrition on student achievement. They specifically focus on increases in nutritional intake (calories) resulting from schools’ responses to accountability pressures. Their results indicate positive short term effects of increased caloric content on student achievement of 10-11 year olds. Our study further extends their findings by presenting experimental evidence on both student achievement and behavior, and by focusing on preschool children (4-6 year-olds). Despite the generally acknowledged importance of early childhood development, and nutrition as a key determinant in this development, the limited literature on in-class implications of sugary drinks did not study preschool children.

In order to study the immediate effect of sugary drinks on preschool chil-

children’s behavior and student achievement, we set up two randomized, double-blind, placebo-controlled trials in a classroom setting. The two field experiments, with randomization at individual level, involved 25 classes and a total of 462 children in the community of Flanders, Belgium. One half of the class was assigned to the treatment group and received a sugary drink, the other half received an artificially sweetened drink. Test scores and behavioral assessments were collected both before and after consumption of the drinks and at different moments in time.

Figure 1: Conceptual representation of outcomes after treatment

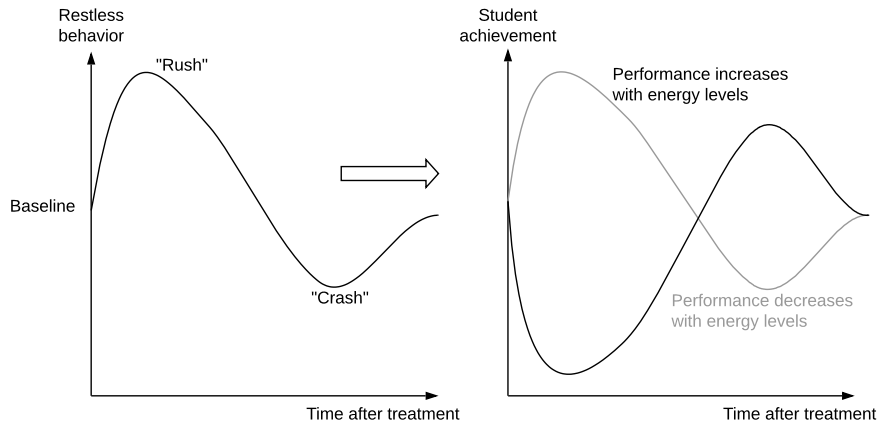


Figure 1 presents the conceptual flow of outcomes after treatment. The hypothesized impact of sugary drinks on restless behavior is positive initially, with a potential decline - the “sugar crash”- before returning back to the baseline level. The impact on student achievement can be either positive or negative depending on the way higher energy levels are channelled by the children when taking the test. This paper is the first to study the relationship between sugary drinks and student achievement and sheds more light on the potential mechanisms in Figure 1.

Our results indicate that boys significantly decrease their mathematics test

scores after drinking a sugary drink, compared to boys who received a placebo. We find the opposite effect for girls. Our findings are robust across both field experiments and to alternative empirical specifications. We consider several potential explanations for our results, and contemplate the relationship between stress, sugar consumption, and student achievement to be the most likely mechanism driving our heterogeneous findings. We also find that consuming one sugary drink results in more restless behavior among boys. By altering the timing of the second test for different classes, we find that positive effects for girls' mathematics achievement emerge after 45 minutes and disappear after 2 hours while mathematics achievement for boys receiving the sugary drink are significantly lower compared to boys in the placebo group, but only after 2 hours. In terms of behavior, our findings suggest that sugary drinks induce an initial 'relaxing' effect, before making boys more restless until 60 minutes after the treatment. Despite differences in timing and power, the robustness of the results is shown across the two field experiments.

2 Experimental design

We conducted two randomized, double-blind, placebo-controlled trials involving preschool children, aged 4 to 6 years, living in the community of Flanders, Belgium.⁵ Although 462 pupils were randomized, after attrition, the final sample included 399 children in 25 different classes. The derivation of this final sample from the contacted sample is presented in Figure 2. Figures A1 and A2 in the Appendix present this derivation separately for the first field experiment (denoted by 'Study I') and the second field experiment (denoted by 'Study II'). Empirical checks of selection by school principals and children (parents) are

⁵Note that the impact of sugary drinks will depend on differences in the final concentration of sucrose, glucose and fructose, which differs across regions. For example, drinks in Australia are sweetened with sucrose from sugar cane (50% glucose and 50% fructose), while in the US, drinks are typically sweetened with high fructose corn syrup. In Europe, and hence in our experiment, sugary drinks are sweetened with sucrose-rich sugar beet. The timing of the effect will also be affected by the respective concentration. Glucose (but not fructose) rapidly elevates plasma glucose and insulin levels, whereas sucrose must first be metabolised to free glucose and fructose (Varsamis et al., 2017). All findings reported here need to be interpreted in light of these regional variations.

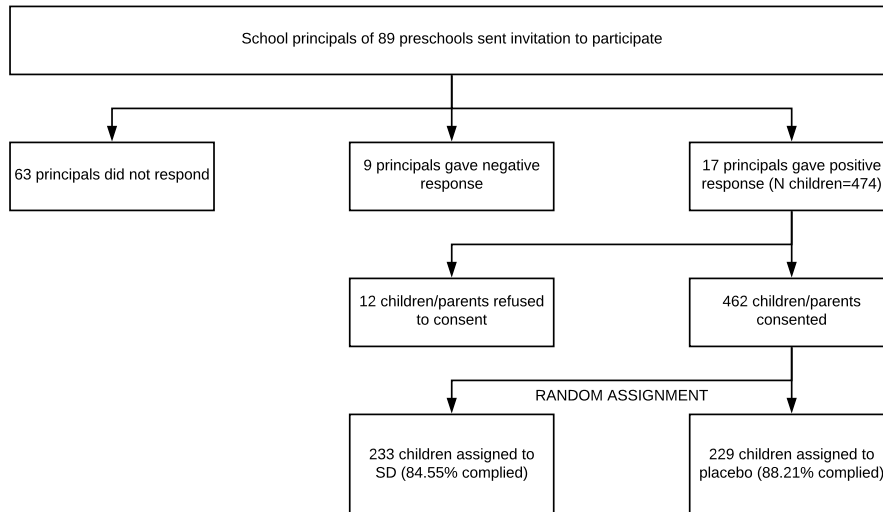
reported in Section 2.3.

The first study started on March 7, 2017 and ended on May 7, 2017. During this period, 54 preschools were contacted, out of which 6 preschools agreed to participate. The sample for this study consists of 187 children in 10 classes.

The second study started on September 10, 2017 and ended on November 28, 2017. During this period, 35 randomly selected preschools were contacted, out of which 11 preschools agreed to participate in the trial. The sample for this study consists of 275 children in 15 classes. In this second study, additional body weight data was collected and classes were assigned to different testing moments to accommodate possible time-varying effects of sugary drinks, as blood glucose levels rise and fall after consumption (Bellisle et al., 2004; Benton et al., 2007; Micha et al., 2011).

This study (covering both field experiments) is registered in the AEA RCT Registry and the unique identifying number is AEARCTR-0002782. Written consent was obtained from both parents and school principals.

Figure 2: Selection of participants Study I & II



Notes: The final sample consists of 399 preschool children in 25 classes in 17 schools. SD: sugary drink; Placebo: artificially sweetened drink. Non-compliers are children who did not drink their consumption, or only partially.

2.1 Procedure

Student achievement and behavior were evaluated before and after the treatment. The randomization took place at individual level and all the children in a participating class were included in the study design. Within each class, all schools desks were marked with a number before the children entered the class. Each school desk was randomly assigned either a sugary or an artificially sweetened drink by the study administrator. Thereafter, all children were seated in accordance with the number they received using a random number generator.

The children assigned to the sugar group were given a 150 mL lemon-flavored soft drink with added liquid sugar, corresponding to a total of 35 grams of sugar. A pilot study prior to the experiment revealed that children often did not drink a full can (250mL) within the short time-frame of our experiment. We therefore opted for smaller cups (150mL) with added liquid sugar to maximize compliance while the chosen dosage is equivalent to that found in one can of soda. The control group received a placebo that consisted of 150 mL drink containing aspartame and acesulfame-K as artificial sweeteners. This placebo drink can be seen as the ‘zero’ equivalent of the treatment drink.

The timing of the second test was administered at four time points: after 30 minutes, 45 minutes, 60 minutes, or 120 minutes. This was done to accommodate possible differential effects over time. These time points were chosen in accordance with the literature (Smith et al., 2011) and to capture a potential “sugar rush” and/or a “sugar-crash”. In the first randomized trial, all 10 classes were tested after 45 minutes. In the second randomized trial, each of the 15 classes were randomly assigned to one of the other three time points, i.e. 30, 60 or 120 minutes after consumption. Apart from timing of the second test, the design of the experiment was identical in both studies.

Between the first and second test, the teachers were instructed to continue

their classes as before in order to mimic a realistic setting. The data was collected either in the morning or afternoon. Although this timing of data collection could influence the findings of the study, including class fixed effects in our empirical specification should control this effect as all the children in a class were treated either in the morning or afternoon.⁶ Along the same lines, teacher-specific effects are also captured by including class fixed effects. The randomization procedure at the individual level should balance the intake of carbohydrates during breakfast or lunch between the treatment and control groups. Irrespective of the time of testing, the children were given breaks between the first and the second test. In order to limit the influence of possible nutritional interactions, all participating children were not allowed to drink or eat anything between the first and second test, and hence not during break time (i.e. at 10am or 2pm, depending on timing of data collection).⁷

2.2 Data collection

The collection of the data was done by five graduate students enrolled in the Masters of Economics (research program), trained in qualitative and quantitative methods. All graduate students received additional training on how to administer the cognitive and behavioral tests and were assigned classes in pairs. In each pair, one student was responsible for reading the test out loud and setting up the experiment before the children entered the class (study administrator), and the other was responsible for evaluating the behavior before and after the intervention took place (assessor). The student pairs and their roles (study administrator or assessor) were kept the same as much as possible to ensure consistency in the evaluations. Nonetheless, including class fixed effects accounts for possible bias through changing pairs.⁸

⁶Note that only 3 out of 25 classes participated in the study during the afternoon. Further research will be needed to assess differential effects of sugar intake in the morning or afternoon.

⁷Compliance with this was monitored. Moreover, when parents were asked to consent with their children participating in the experiment, parents were also asked not to provide snacks for their children during the break.

⁸Due to the within-class randomization, the class size of treatment and control groups is identical by construction. Participating classes varied between 9 and 28 children, with an

Due to the randomization procedure described in 2.1, the assessor was blinded to the treatment conditions when the children entered the class. This procedure prevents observer bias when completing the behavioral checklist. Only the name and number of each child was known to the assessor. Cognitive tests were corrected by the blinded assessor, and by both authors of this study, to ensure a valid assessment of student achievement.

Student achievement was examined using the mathematical test ‘Algebra for Preschoolers’.⁹ This standardized test was developed by the Central Institute for Test Development (CITO), in order to track the ability of children and to support teachers in tailoring education to the needs of all children. It is a valid and reliable measure of general numeracy skills of preschool children, and it has a high item discriminating power¹⁰, such that the test is able to measure (mathematical) achievement adequately. The test was originally designed for the Netherlands, which has the same language of instruction. Discussion and consultation with several preschool teachers led to the discovery and subsequent omission of six questions measuring contextual information irrelevant to the Flemish preschool children. The remaining 42 questions were equally divided into two parts, with 21 questions in each, before and after the intervention.¹¹ The test was read aloud in the classroom by the study administrator and the children were given 45 seconds to respond to each of the questions individually. The subsequent questions were covered by a piece of carton to help children focus on the question being read aloud, as instructed by the manual of the CITO test. For each part, the test duration was equal to 25 minutes.

We assessed the behavior of the children using a behavior checklist conforming to previous studies (Benton et al., 2007; Benton and Stevens, 2008; Blatchford, 2003). The blinded assessor evaluated the behavior during the cog-

average of 18.

⁹Throughout the remainder of this paper we use the term ‘mathematics achievement’ rather than student achievement to align closely with the unit of measurement.

¹⁰I. Koerhuis, J. Keuning, “Wetenschappelijke verantwoording van de toetsen Rekenen voor kleuters [Scientific justification of the test Algebra for preschoolers]” (Arnhem, 2011), (available at <http://toetswijzer.kennisnet.nl/html/tg/22.pdf>).

¹¹Figure A3 presents an example of such a question.

nitive test before and after the treatment. We used the mean of the following items as a proxy for restless behavior: ‘distracting others’, ‘looking around’, ‘fidgeting’, ‘arguing with others’, and ‘laughing loudly and standing out’. The score for each component is measured as a Likert scale ranging from 1 to 7, where higher scores indicate more restless behavior.

During the second study, (N=275, 15 classes), bodyweight data was collected. The children were weighed after the second test to avoid interference with the treatment effect.

2.3 Check of internal and external validity

Children in placebo or sugar conditions exhibited similar characteristics across pretreatment scores and basic demographic measures. From Table 1, it is clear that there are no significant differences between both groups in terms of children’s age, sex, mathematics achievement, behavior, and body weight. This result holds for both field experiments (Study I and Study II), as well as for the pooled sample. Moreover, repeating this exercise for boys and girls separately also fails to find significant differences between treatment and control groups.

Out of 462 consenting children (and parents), 63 (13.64%) children did not comply to their treatment. Non-compliers either did not drink because they did not like the taste (86.36%) or did not drink the entire cup in time (15.87%).¹² In Study I, there were 20 non-compliers (10.7%) and 43 in Study II (15.64%).

The children assigned to the sugary drink condition are equally likely to comply ($p = 0.252$) as those assigned to the control condition. While this difference in non-compliance between treatment (15.45%) and control (11.79%) is non-significant, it might still be numerically meaningful. Therefore, we construct upper and lower bounds that limit the range of the treatment effect, accounting for possible non-random sample selection, following Lee (2009). Necessary assumptions are random treatment assignment (satisfied by the RCT design and

¹²By choosing small cups (150mL), see 2.1, we minimized the share of non-compliers who did not drink their entire cup in time.

Table 1: Comparison before treatment.

Study I (n=187)			
	Sugar	Placebo	<i>P</i> -value
Age (years)	5.237	5.277	0.533
Boy (%)	0.548	0.436	0.126
Mathematics achievement	13.011	13.309	0.662
Behavior	2.669	2.760	0.483
Compliers	0.882	0.904	0.621
Study II (n=275)			
	Sugar	Placebo	<i>P</i> -value
Age (years)	4.871	4.881	0.807
Weight (kg)	20.309	20.259	0.888
Boy (%)	0.493	0.459	0.579
Mathematics achievement	10.779	10.370	0.452
Behavior	2.704	2.714	0.951
Compliers	0.843	0.881	0.335
Study I + II (n=462)			
	Sugar	Placebo	<i>P</i> -value
Age (years)	5.017	5.044	0.504
Boy (%)	0.515	0.450	0.161
Mathematics achievement	11.670	11.576	0.832
Behavior	2.690	2.733	0.702
Compliers	0.858	0.891	0.281

Notes: Comparison of means, and corresponding *p*-values of the treatment group (sugary drink) and control group (placebo) with robust standard errors. Mathematics achievement and behavioral scores are reported before standardization.

supported empirically in Table 1), and monotonicity (being assigned to treatment affects selection in the same way for any observation). The group that is least affected by selection is trimmed to balance both treatment and control groups. Upper and lower bounds for the treatment effect are then obtained by computing differences in group means. Results are presented in Table A1 and provide strong support for the existence of non-zero treatment effects.¹³

The low response rate of school principals could create an issue of external validity. Despite the high share of consenting parents (97.5%) in schools where principals were willing to participate, only 26 out of 89 school principals responded to the invitation for carrying out the present study (29.21%). To

¹³Note that we followed Lee (2009) by constructing bounds for the baseline specification where there is only a treatment indicator with no other covariates. In the remainder of this paper, we report treatment effects with a full set of covariates to improve precision (see 2.4). Stratifying the sample along this set of (discretized) covariates would narrow bounds further (Lee, 2009, p.1094), although it would require a large number of sample splits. From Table A3, it can be seen that treatment effect estimates are consistent when covariates are added to the baseline specification, especially for pooled estimates, which follows directly from the randomization procedure. As a result, treatment effect bounds will be in the same range for this full specification, while standard errors of the estimated coefficients will be smaller.

test the external validity of our findings, we compared the composition of different sets of schools in Table A2. This was done by matching school-level data on socio-economic indicators (% children with a mother having at most a high school degree, % children not speaking the official language (Dutch) at home, and % children receiving a school allowance) to unique school identifiers participating in the study. The results presented in Table A2 support the external validity of our findings. First, the sample of 89 contacted schools appears to be representative of the average school in Flanders, in terms of socio-economic composition (Panel A). Second, comparing actual participants to non-participants (nonresponse or decline) also does not suggest selective compliance (Panel B). Third, we compare the final participating sample of schools to the population of schools in Flanders (Panel C). Again, no significant differences are found in terms of socio-economic composition.

In sum, selective compliance of students and selective participation of schools appear to be unlikely sources of bias, strengthening both the internal and external validity of this study.

2.4 Empirical strategy

The randomized design at individual level enables straightforward analysis. We use the following regression framework:

$$y_{i,j} = \beta_1 T_{i,j} + \beta_2 age_{i,j} + \beta_3 boy_{i,j} + \beta_4 preT_{i,j} + \gamma_j + \epsilon_{i,j} \quad (1)$$

where $y_{i,j}$ represents the value of an outcome (mathematics achievement or behavior) for child i in class j ; $T_{i,j}$ indicates whether child i in class j received the sugary drink; $age_{i,j}$ is a categorical variable indicating the age (4, 5, or 6) of child i in class j ; $boy_{i,j}$ is an indicator variable equal to one if child i in class j is a boy; $preT_{i,j}$ is the pretreatment score of an outcome for child i in class j (mathematics achievement or behavior); γ_j captures the class fixed effect of class j , which includes the timing of the second test class j was assigned to; $\epsilon_{i,j}$ is an error term for child i in class j . Mathematics achievement and behavior are

measured in Z-scores, allowing the interpretation of treatment effect coefficients in terms of standard deviations (SD). Table A3 replicates our main results for different combinations of covariates.

3 Results

Table 2 shows that the consumption of one sugary drink significantly reduced mathematics achievement for boys, whereas test scores for girls significantly improved after consumption. The observed differential effects by sex are hidden in the average treatment effects. In terms of magnitude, drinking one sugary drink decreased mathematics achievement of boys by 0.26 standard deviations relative to boys drinking an artificially sweetened placebo. For girls, the results indicate that the consumption of one sugary drink improved mathematics achievement by almost 0.29 standard deviations, compared to girls who drank the placebo drink. The results also indicate that boys became more restless by 0.17 standard deviations after consuming a sugary drink, although this increase was only significant at the 10 percent level. A significant effect on the behavior of girls could not be found.

Comparing the coefficients between boys and girls, we find that the impact of sugary drinks on mathematics achievement is significantly more positive for girls ($p < 0.01$) compared to boys. Although this differential achievement effect by sex was not hypothesized initially, the second study was able to replicate these findings.¹⁴ In contrast to achievement effects of sugary drinks, we could not find a significant difference in the coefficients for behavior between boys and girls (Study I+II: $p = 0.483$; Study I: $p = 0.593$; Study II: $p = 0.632$).

Comparing the coefficients between both studies, it can be seen in Table 2 that the coefficients for mathematics achievement and behavior are not sig-

¹⁴Note that there were no significant differences in baseline observables (age, mathematics achievement, behavior, and weight) between treatment and control groups, when comparing boys and girls separately. This finding holds for both Study I and Study II, and when pooling both studies. The importance of including control variables, and particularly pretreatment scores to mitigate pretreatment heterogeneity bias, is illustrated in Table A3. From this table, it can be seen that treatment effect estimates in Study I experience a change in size and significance when pretreatment scores are included.

Table 2: Effects of one sugary drink on mathematics achievement and behavior

	Mathematics achievement			Behavior		
	All	Boys	Girls	All	Boys	Girls
Study I (n=167)	0.127 (0.082)	-0.109 (0.111)	0.375 (0.121)	0.178 (0.076)	0.248 (0.105)	0.175 (0.105)
Study II (n=232)	-0.008 (0.078)	-0.347 (0.116)	0.215 (0.109)	0.036 (0.084)	0.109 (0.129)	0.032 (0.120)
Study I + II (n=399)	0.045 (0.058)	-0.260 (0.082)	0.286 (0.082)	0.095 (0.059)	0.170 (0.086)	0.092 (0.082)
Chi-square test (p-value)	1.56 (0.212)	2.56 (0.109)	1.14 (0.286)	1.69 (0.193)	0.83 (0.363)	0.95 (0.329)

Notes: This table reports treatment effects of drinking one sugary drink compared to drinking an artificially sweetened drink. Treatment effects reported for Study II aggregate results for classes assigned to assessments at different time points, corresponding to 30 minutes, 60 minutes, and 120 minutes after the consumption of drinks (sugar or placebo). Controls: age, sex, pretreatment test scores, and class fixed effects, including the timing of the second test. Robust standard errors in parentheses. The Chi-square test statistic is calculated to examine whether the treatment effect coefficients differ significantly across Study I and Study II.

nificantly different. Inevitably, pooled estimates are more precise. Table A4 probes robustness when standard errors are clustered at class level, and under different inference methods correcting for the small number of clusters: the Young (2016) degrees of freedom correction and the Cameron et al. (2008) wild-bootstrap method. All results are quantitatively and qualitatively comparable.

Figure 3 presents treatment effect estimates over time.¹⁵ The patterns observed in Figure 3 suggest that there is no impact on mathematics achievement after 30 minutes. After 45 minutes, positive effects for girls emerge, while effects for boys are negative and only become significant after 60 minutes. After 120 minutes, the positive effect for girls is no longer significant while there is still a significantly negative impact on boys' mathematics achievement. In terms of behavior, the boys assigned to the sugary drink condition became less restless (calmer) at 30 minutes after consumption, as compared to boys that were assigned artificially sweetened drinks. However, boys in the sugar condition who were evaluated 15 minutes later, became more restless vis à vis those in the placebo condition. Although the trend is parallel for boys and girls, the effect is only significant for boys. The discrepancy in effects for mathematics achieve-

¹⁵Due to the smaller number of classes at 30, 60, and 120 minutes between the first and second test, estimates are less precise. As a result, confidence intervals are narrower at 45 minutes since twice the number of classes were evaluated (10 classes) compared to other testing moments (5 classes).

ment between boys and girls and the absence of a significant increase in restless behavior among girls already hints at a potential mechanism: a differential way of channeling the sugars into mathematics achievement between boys and girls.

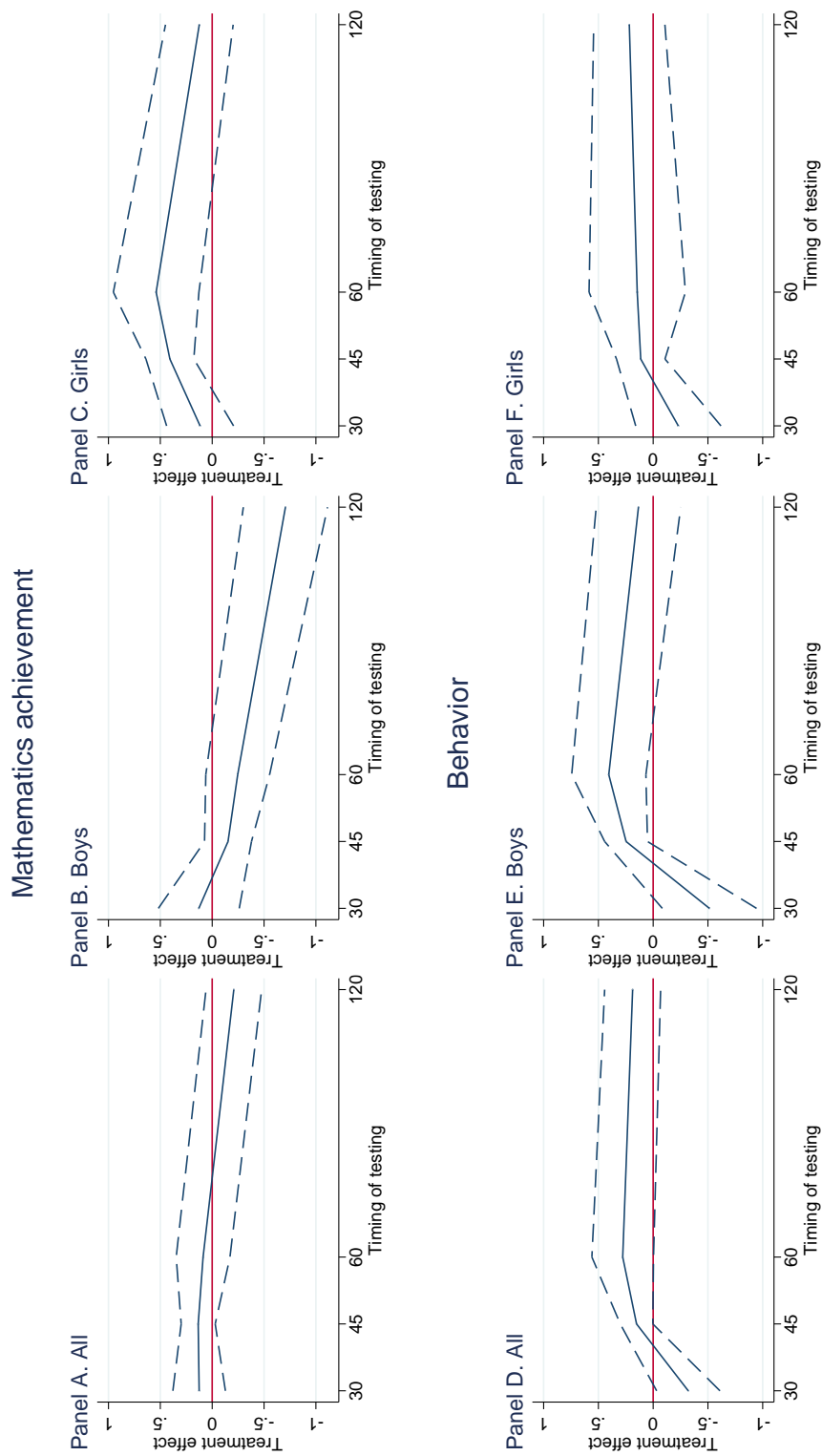
Figure 4 further disentangles our results by looking at the interaction between treatment effects and pretreatment outcomes. Overall, Figure 4 suggests that effects of sugary drinks are independent of children’s test scores and behavior prior to consumption, although the imprecision of the obtained estimates does not fully warrant this conclusion.

4 Mechanisms

The results of the experiment highlight a heterogeneous impact by sex and over time. In terms of behavior, we did not find a significant impact for girls, while boys responded by becoming more restless. Over time, our findings suggest that sugary drinks induce an initial ‘relaxing’ effect, before making boys more restless. In terms of mathematics achievement, the results indicate an opposite response between boys (negative) and girls (positive). Over time, this positive effect for girls emerged only after 45 minutes and disappeared after 2 hours while detrimental effects for boys emerged after 60 minutes and remained significant.

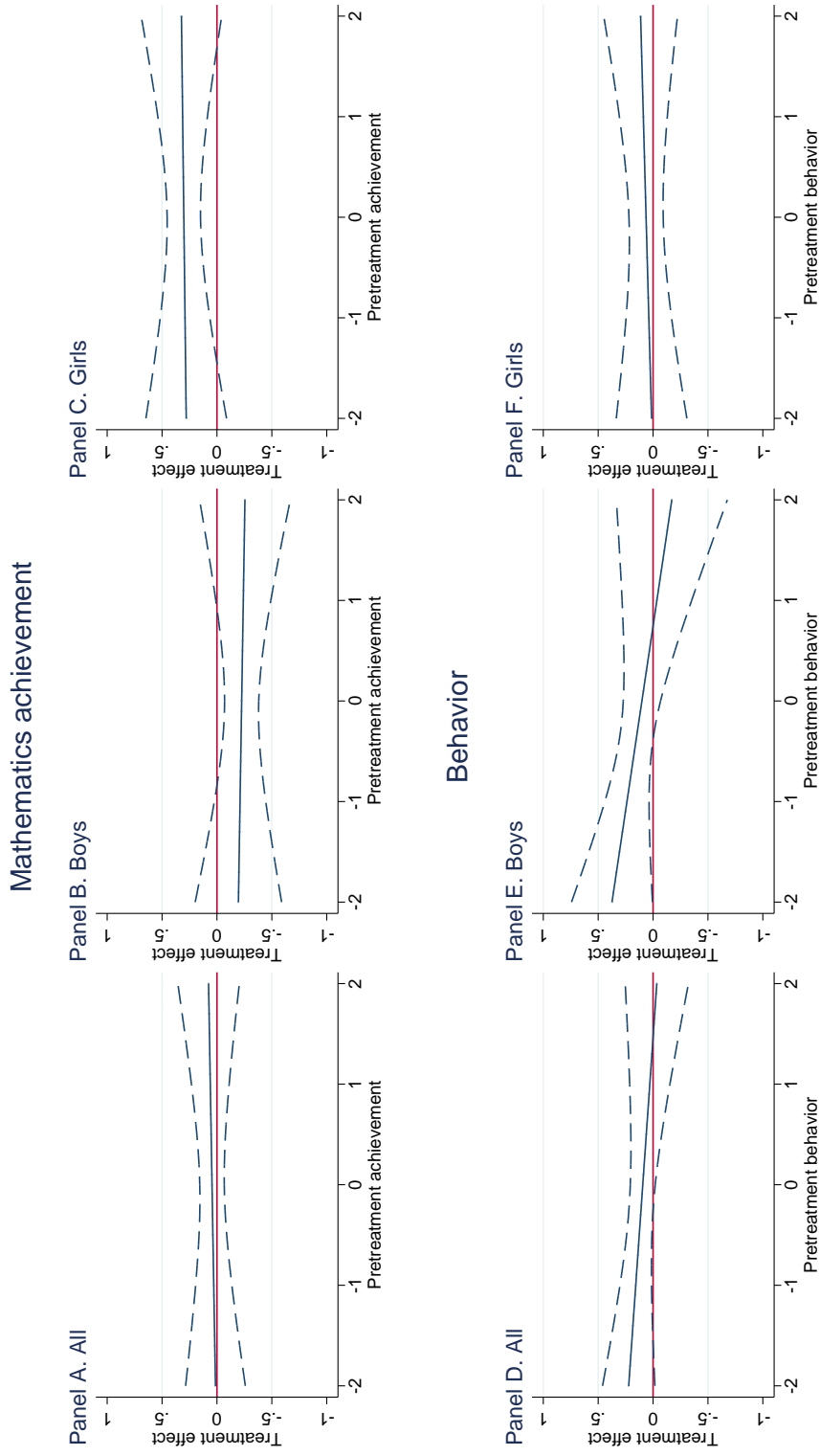
The relationship between nutrition and mathematics achievement, and the heterogeneity by sex, identified in the present study has been reported before. For example, one study suggested that sex influences the way breakfast affects cognitive processes (Mahoney et al., 2005). Girls seemed more affected than boys, as girls who voluntarily ate more of the nutritionally inadequate breakfast worked significantly faster on a number-checking test, but no differences were found on tests of multiplication or grammatical reasoning. Belot and James (2011) reported more pronounced effects of the ‘Feed Me Better’ campaign for girls, compared to boys. Wyon et al. (1997) found that lower energy intake

Figure 3: Treatment effect variation by the timing of the second test



Notes: This figure plots estimated coefficients and 95% confidence intervals of interactions between the treatment effect, and the timing of testing. Higher behavioral scores indicate more restless behavior of children in the classroom. All estimated regressions used here include the following set of controls: age, sex, pretreatment test scores, and class fixed effects. Confidence intervals were obtained using robust standard errors.

Figure 4: Treatment effect variation by pretreatment scores



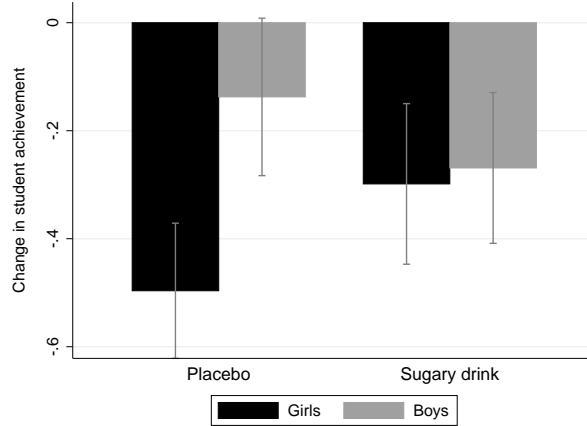
Notes: This figure plots estimated coefficients and 95% confidence intervals of interactions between the treatment effect, pretreatment test scores, and pretreatment behavior. Higher behavioral scores indicate more restless behavior of children in the classroom. All estimated regressions used here include the following set of controls: age, sex, pretreatment test scores, and class fixed effects, including the timing of the second test. Confidence intervals were obtained using robust standard errors.

during breakfast significantly and negatively affected creative thinking in 10-year-old boys but not in girls.

A possible explanation for the discrepancy in effects by sex, in the present and previous studies, is the difference in weight between boys and girls, as both received the same treatment. The average weight of girls in our sample (19.77kg) is significantly lower than that of boys in our sample (20.76kg), which is consistent with weight differences between healthy boys and girls at this age. To test the possibility that differences in effects between boys and girls are driven by a higher sugar dosage *per kg of bodyweight* for girls, we include an interaction effect between sex and treatment. We do this for the subsample of children where weight data was collected and check the robustness of the interaction effect when adding weight to the estimated equation. In both specifications, the interaction term is similar in size and remains significant when children’s weight is accounted for (see Table A3, panel C). Also, weight appears not significantly related to mathematics achievement or behavior, dismissing our initial hypothesis. The relationship between sugar consumption, stress, and mathematics achievement provides a more plausible explanation to reconcile heterogeneous findings by sex. Cognitive tests (e.g. IQ tests, or standardized tests as used in the present study) induce acute stress among participants (Vogel and Schwabe, 2016). In this type of competitive environments, previous studies pointed at differential effects of competition on mathematics achievement for boys and girls (Niederle and Vesterlund, 2007), suggesting a stronger stress response among girls. In adult women, sugar has been found to reduce stress-induced levels of cortisol when faced with mathematical problems (Tryon et al., 2015). A similar pattern could not be observed when subjects consumed aspartame. Although no evidence is available for preschool children, differential reductions in cortisol levels among boys and girls following the consumption of a sugary drink offers a plausible explanation for our findings.

Next, we compare the changes in mathematics achievement between the first and the second test in Figure 5. Regressing this change on the sex dummy indi-

Figure 5: Change in mathematics achievement between first and second test



Notes: This figure decomposes the change in mathematics achievement between the first and the second test by treatment status and sex. Mean changes are displayed along with 95% confidence intervals around the group means.

cates that girls in the placebo condition experience a significantly larger decrease in their scores compared to boys assigned to the placebo group (-0.496 SD vs -0.137 SD, $p < 0.01$). On the contrary, the change in scores between both tests does not significantly differ by sex for children assigned to the sugar condition (-0.298 SD vs -0.269 SD, $p = 0.776$). Hence, it appears that both boys and girls decrease their test scores after the consumption of a sugary drink. This decrease between the first and second test is contrary to expectations given that the children got acquainted with the test format during the first test. One explanation could be a decline in concentration when doing a similar test for the second time. Combining the effects in the treatment and placebo group yields a net effect for girls (see Table 2) that is clearly positive. Linking this observation back to the competition mechanism, sugary drinks potentially mitigate the detrimental effects of competition on girls' test scores. For boys, this mitigating effect might be more limited as they suffer less from competition, resulting in an overall negative effect of sugary drinks on mathematics achievement.

5 Discussion and conclusion

This study set up two field experiments to assess the in-class impact of one sugary drink on behavior and mathematics achievement of preschool children. Our findings indicate that the consumption of one sugary drink (a sugar dosage equivalent to one can of soda) decreases test scores for boys by 26 percent of a standard deviation. In contrast, mathematics achievement increased for girls by almost 29 percent of a standard deviation. This positive effect for girls emerges after 45 minutes and disappears after 2 hours. In contrast, the negative effect for boys emerges later, after 60 minutes, but remains significantly negative 2 hours after consuming a sugary drink. In terms of behavior, we did not find a significant impact for girls, while boys responded by becoming more restless after 45 minutes¹⁶ after an initial ‘relaxing’ effect, 30 minutes after the treatment. We argued that the relationship between stress, sugar consumption, and mathematics achievement might act as a possible mechanism behind differences in results for boys and girls.

Given earlier evidence that health education might not have the anticipated effect on behaviour (Kremer and Miguel, 2007), our findings have policy implications considering the wide availability of vending machines in schools and the high consumption of sugary drinks among preschool children in developed countries. Our results suggest that the detrimental and persisting effects on boys’ test scores and teachers’ classrooms are unlikely outweighed by a marginal increase in resources from selling sugary drinks at school. Moreover, from a societal point of view, our results should be viewed in light of the skewed consumption pattern of sugary drinks by sex and socio-economic status, and the harmful effects on children’s lives in terms of weight gain and obesity rates (Malik et al., 2013; Keller and Bucher Della Torre, 2015). Taking into account the higher intake of sugar among disadvantaged boys, our results could provide a partial explanation for the observed underperformance of boys in low socio-economic

¹⁶Despite the lack of precision, the obtained signs are consistent with previous findings for preschool boys (Prinz and Riddle, 1986)

status groups.¹⁷

Given the realistic dosage chosen in the treatment condition, and the strong external validity of the evaluated sample, the findings of this study can likely be generalized to school contexts in other countries. In some countries, where vending machines have already been banned from (primary) schools, such as Belgium, our findings have implications for school dietary guidelines as children could substitute their consumption of drinks bought from vending machines by drinks brought from outside the school. Therefore, our results can be seen as evidence to support a broader ban on the availability of sugary drinks in schools.

Several potential extensions can be identified for research following the present study. First, redefining the treatment and/or control conditions. Our study did not include an additional control condition without artificial sweeteners (e.g. sparkling water or no drink at all).¹⁸ Considering the possibility of the Hawthorne effect, a placebo is preferred rather than sparkling water because participants can easily distinguish between sparkling water and sugary drinks. Further research could explore the differential impact of artificially and sugar-sweetened drinks on the cognition or behavior of children.¹⁹ Also, differing treatment conditions could be a potential avenue for further research. In the current design, all children in the treatment group are receiving the same amount of sugar. Replicating the current study with varying doses could potentially help identify a sugar ‘tipping point’ that triggers the effects reported here.

Second, further research could consider additional heterogeneous effects. For example, changes over time in the behavioral-regulation sex gap (Raffaelli et al.,

¹⁷The OECD PISA survey indicated “Boys, especially disadvantaged boys, also need special support if they are to improve their reading performance. In all but 3 PISA-participating countries at least 20% of disadvantaged boys did not attain the minimum level of proficiency in reading. In 24 countries and economies, more than 70% of disadvantaged boys scored below that level.” (Schleicher, 2019, p. 18).

¹⁸Alternately, studying the effects of ‘natural sugars’ in fruit juice, instead of added sugars, provides another interesting research avenue due to different fructose-to-glucose ratio between fruit juice drinks and sugar-sweetened beverages (Walker et al., 2014).

¹⁹The blood glucose response to artificially and sugar-sweetened drinks has been studied in adult men (Bloomer et al., 2016). The results of this study indicated that blood glucose levels were responsive to the sugar-sweetened conditions whereas the glucose levels remained stable when artificially sweetened drinks were consumed. Along the same lines, sugar has been found to reduce stress-induced levels of cortisol in adult women, whereas a similar pattern could not be observed when subjects consumed aspartame (Tryon et al., 2015).

2005; DiPrete and Jennings, 2012) could alter the effects of sugary drinks as children progress through early childhood. Also, differential stress hormone levels for children with a high or a low socio-economic status (SES) (Lupien et al., 2000, 2001) could reveal interesting heterogeneous effects beyond those identified in this paper. This is especially important, as the consumption of sugary drinks is strongly skewed by race and SES. For example in the US, preschool children’s consumption (aged 2-5 years) is highest among African Americans and Hispanics, typically in the lowest quintile (Scharf and DeBoer, 2016; Bleich et al., 2018). Finally, one could look at hyperactive children or children with ADHD, who might be especially sensitive to sugar.

Another interesting avenue for further research is to consider the interaction between the effect of sugary drinks and children’s diets. This will require a detailed administration of food and drinks consumed by the children participating in the study. Although children in the present study were not allowed to drink or eat anything during the break, the nature and composition of the food during breakfast and lunch may influence the reaction to a glucose drink. For example, it has been suggested that previous breakfast consumption alters the effect of a glucose drink on cognitive outcomes (Benton and Parker, 1998), and affects the tolerance to glucose after lunch (Liljeberg et al., 1999). Although the children belonging to the treatment and control groups are expected, on average, to follow the same diet (which is supported by the balanced observables in Table 1), the interaction between their diet and the effect of sugary drinks cannot be identified in the present study. Nonetheless, the interventions were designed to measure the effect of sugary drinks during a regular school day, making our findings relevant in a realistic school context.

Replicating the present study using alternative measures of student achievement and behavior offers a final potential extension of the present study. For example, different effects could possibly have been obtained, depending on how achievement is defined (e.g. memory task, reading test). Considering behavioral effects, it could be that our behavioral scale is insufficiently able to capture vari-

ation in behavior induced by sugary drinks, offering a possible explanation for the generally imprecise behavioral estimates. Also, spill-overs of treated children might mitigate the observed behavioral effects. Peer effects are important in an education context (e.g. Schiltz et al., 2019; Sacerdote, 2014), and could impact the results observed here. In terms of mathematics achievement, peer effects could be limited by preventing children from cheating. However, as we tried to mimic a realistic classroom setting, there might still be an interaction between the behavior of children if more restless behavior of children in the treatment group affects the behavior of children in the control group. If this is the case, the present study could underestimate the behavioral effects of sugary drinks consumption. Further exploration of peer effects after consumption of sugary drinks will be needed to confirm this hypothesis.

References

- Anderson, M. L., J. Gallagher, and E. R. Ritchie (2017). School Lunch Quality and Academic Performance. *NBER Working Paper Series* (w23218).
- Anderson, P. M. and K. F. Butcher (2006). Reading writing, and refreshments: are school finances contributing to children’s obesity? *Journal of Human Resources* 41(3), 467–494.
- Bellisle, F., Hotel-Dieu, P. d. P. Notre-Dame, and F. Bellisle (2004). Effects of diet on behaviour and cognition in children France Bellisle. *British Journal of Nutrition* 92(S2), S227–S232.
- Belot, M. and J. James (2011). Healthy school meals and educational outcomes. *Journal of Health Economics* 30(3), 489–504.
- Belot, M., J. James, and P. Nolen (2016). Incentives and children’s dietary choices: A field experiment in primary schools. *Journal of Health Economics* 50, 213–229.
- Benton, D., A. Maconie, and C. Williams (2007). The influence of the glycaemic load of breakfast on the behaviour of children in school. *Physiology and Behavior* 92(4), 717–724.
- Benton, D. and P. Parker (1998). Breakfast, blood glucose, and cognition. *American Journal of Clinical Nutrition* 67(4), 772S–778S.
- Benton, D. and M. K. Stevens (2008). The influence of a glucose containing drink on the behavior of children in school. *Biological Psychology* 78(3), 242–245.
- Blatchford, P. (2003). A systematic observational study of teachers’ and pupils’ behaviour in large and small classes. *Learning and Instruction* 13(6), 569–595.
- Bleich, S. N., K. A. Vercammen, J. W. Koma, and Z. Li (2018). Trends in Beverage Consumption Among Children and Adults, 2003-2014. *Obesity* 26(2), 432–441.
- Bleich, S. N. and J. A. Wolfson (2015). Trends in SSBs and snack consumption among children by age, body weight, and race/ethnicity. *Obesity* 23(5), 1039–1046.
- Bloomer, R., S. Peel, R. Moran, and J. MacDonnchadh (2016). Blood glucose and insulin response to artificially- and sugar-sweetened sodas in healthy men. *Integrative Food, Nutrition and Metabolism* 3(1), 268–272.
- Cameron, C. A., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics* 90, 414–427.
- DiPrete, T. A. and J. L. Jennings (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research* 41(1), 1–15.

- Ells, L. J., F. C. Hillier, J. Shucksmith, H. Crawley, L. Harbige, J. Shield, A. Wiggins, and C. D. Summerbell (2008). A systematic review of the effect of dietary exposure that could be achieved through normal dietary intake on learning and performance of school-aged children of relevance to UK schools. *British Journal of Nutrition* 100(5), 927–936.
- Figlio, D. N. and J. Winicki (2005). Food for thought: The effects of school accountability plans on school nutrition. *Journal of Public Economics* 89(2-3), 381–394.
- Fletcher, J. M., D. E. Frisvold, and N. Tefft (2010). The effects of soft drink taxes on child and adolescent consumption and weight outcomes. *Journal of Public Economics* 94(11-12), 967–974.
- Fox, M. K. and E. Condon (2012). School Nutrition Dietary Assessment Study IV. Technical report, Food and Nutrition Service US Department of Agriculture.
- Frisvold, D. E. (2014). Nutrition and cognitive achievement: An evaluation of the School breakfast program. *Journal of Public Economics* 124, 91–104.
- Gómez-Pinilla, F. (2008). Brain foods: the effects of nutrients on brain function. *Nature Reviews Neuroscience* 9(7), 568–578.
- Heckman, J. J. (2008). School, Skills, and Synapses. *Economic Inquiry* 46(3), 289–324.
- Kalhan, S. and I. Kilic (1999). Carbohydrate as nutrient in the infant and child: range of acceptable intake. *European Journal of Clinical Nutrition* 53, S94–S100.
- Keller, A. and S. Bucher Della Torre (2015). Sugar-sweetened beverages and obesity among children and adolescents: a review of systematic literature reviews. *Childhood Obesity* 11(4), 338–346.
- Kremer, M. and E. Miguel (2007). The illusion of sustainability. *The Quarterly journal of economics* 122(3), 1007–1065.
- Kuhn, M. A., P. Kuhn, and M. C. Villeval (2014). Self Control and Intertemporal Choice : Evidence from Glucose and Depletion Interventions. *CESifo Working Paper Series* (No. 4609).
- Lasater, G., C. Piernas, and B. Popkin (2011). Beverage patterns and trends among school-aged children in the US, 1989-2008. *Nutrition Journal* 10(1), 103.
- Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies* 76(3), 1071–1102.
- Leos-Urbel, J., A. E. Schwartz, M. Weinstein, and S. Corcoran (2013). Not just for poor kids: The impact of universal free school breakfast on meal participation and student outcomes. *Economics of Education Review* 36, 88–107.

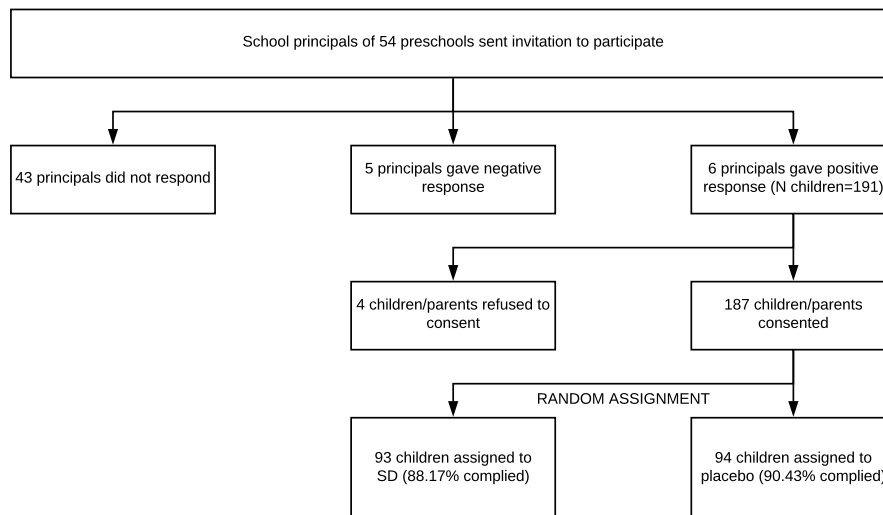
- Levitt, S. D., J. A. List, S. Neckermann, and S. Sadoff (2016). The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *American Economic Journal: Economic Policy* 8(4), 183–219.
- Liljeberg, H., A. Åkerberg, and I. Björck (1999). Effect of the glycemic index and content of indigestible carbohydrates of cereal-based breakfast meals on glucose tolerance at lunch in healthy subjects. *American Journal of Clinical Nutrition* 69(4), 647–655.
- List, J. A. and A. S. Samek (2015). The behavioralist as nutritionist: Leveraging behavioral economics to improve child food choice and consumption. *Journal of Health Economics* 39, 135–146.
- Lupien, S., S. King, M. Meaney, and B. McEwen (2000). Child stress hormone levels correlate with mother’s socioeconomic status and depressive state. *Biological Psychiatry* 48(10), 976–980.
- Lupien, S., S. King, M. Meaney, and B. McEwen (2001). Can poverty get under your skin? Basal cortisol levels and cognitive function in children from low and high socioeconomic status. *Development and psychopathology* 13(3), 653–676.
- Lustig, R. H., L. A. Schmidt, and C. D. Brindis (2012). The toxic truth about sugar. *Nature* 482(7383), 27–29.
- Mahoney, C., H. Taylor, R. Kanarek, and P. Samuel (2005). Effect of breakfast composition on cognitive processes in elementary school children. *Physiology and Behaviour* 85(5), 636–645.
- Malik, V. S., A. Pan, W. C. Willett, and F. B. Hu (2013, 08). Sugar-sweetened beverages and weight gain in children and adults: a systematic review and meta-analysis. *The American Journal of Clinical Nutrition* 98(4), 1084–1102.
- Mazrekaj, D., K. De Witte, and S. Vansteenkiste (2019). Labour market consequences of a high school diploma. *Applied Economics* 51(21), 2313–2325.
- McEwan, P. J. (2013). The impact of Chile’s school feeding program on education outcomes. *Economics of Education Review* 32(1), 122–139.
- Micha, R., P. J. P. Rogers, and M. Nelson (2011). Glycaemic index and glycaemic load of breakfast predict cognitive function and mood in school children: a randomised controlled trial. *British Journal of Nutrition* 106(10), 1552–1561.
- Millimet, D. L., R. Tchernis, and M. Hussain (2010). School Nutrition Programs and the Incidence of Childhood Obesity. *Journal of Human Resources* 45(3), 640–654.
- Niederle, M. and L. Vesterlund (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics* 122(3), 1067–1101.

- Papoutsis, G. S., A. C. Drichoutis, and R. M. Nayga (2013). The causes of childhood obesity: A survey. *Journal of Economic Surveys* 27(4), 743–767.
- Philippou, E. and M. Constantinou (2014). The Influence of Glycemic Index on Cognitive Functioning: A Systematic Review of the Evidence. *Advances in Nutrition: An International Review Journal* 5(2), 119–130.
- Prinz, R. J. and D. Riddle (1986). Associations between nutrition and behaviour in 5-year-old children. *Nutrition reviews* 44(s3), 151–158.
- Raffaelli, M., L. J. Crockett, and Y.-L. Shen (2005). Developmental Stability and Change in Self-Regulation From Childhood to Adolescence. *The Journal of Genetic Psychology* 166(1), 54–76.
- Sacerdote, B. (2014). Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward? *Annual Review of Economics* 6(1), 253–272.
- Schanzenbach, D. W. (2009). Do School Lunches Contribute to Childhood Obesity? *Journal of Human Resources* 44(3), 684–709.
- Scharf, R. and M. D. DeBoer (2016). Sugar-Sweetened Beverages and Children’s Health. *Annual Review of Public Health* 37, 273–296.
- Schiltz, F., D. Mazrekaj, D. Horn, and K. De Witte (2019). Does it matter when your smartest peers leave your class? evidence from hungary. *Labour Economics* 59, 79–91.
- Schleicher, A. (2019). Pisa 2018: Insights and interpretations. *OECD Publishing*.
- Smith, M. A., L. M. Riby, J. A. M. van Eekelen, and J. K. Foster (2011). Glucose enhancement of human memory: A comprehensive research review of the glucose memory facilitation effect. *Neuroscience and Biobehavioral Reviews* 35(3), 770–783.
- Sorhaindo, A. and L. Feinstein (2006). What is the relationship between child nutrition and school outcomes? Technical report, Centre for Research on the Wider Benefits of Learning, London.
- Tiffin, R., A. Kehlbacher, and M. Salois (2015). The effects of a soft drink tax in the UK. *Economic Journal* 24, 583–600.
- Tryon, M. S., K. L. Stanhope, E. S. Epel, A. E. Mason, R. Brown, V. Medici, P. J. Havel, and K. D. Laugero (2015). Excessive sugar consumption may be a difficult habit to break: A view from the brain and body. *Journal of Clinical Endocrinology and Metabolism* 100(6), 2239–2247.
- Varsamis, P., R. N. Larsen, D. W. Dunstan, N. Owen, B. A. Kingwell, G. L. Jennings, N. Owen, and B. A. Kingwell (2017). The sugar content of soft drinks in Australia, Europe and the United States. *Medical Journal of Australia* 206(10), 454–455.
- Vogel, S. and L. Schwabe (2016). Learning and memory under stress: impli-

- cations for the classroom. *Nature Partner Journals (npj) Science of Learning* 1(16011), 1–10.
- Walker, R. W., K. A. Dumke, and M. I. Goran (2014). Fructose content in popular beverages made with and without high-fructose corn syrup. *Nutrition* 30(7-8), 928–935.
- Wolraich, M. L., D. B. Wilson, and J. W. White (1995). The effect of sugar on behavior or cognition in children: a meta-analysis. *Jama* 274(20), 1617–1621.
- Wyon, D. P., L. Abrahamsson, M. Jartelius, and R. J. Fletcher (1997). An experimental study of the effects of energy intake at breakfast on the test performance of 10-year-old children in school. *International Journal of Food Sciences and Nutrition* 48(1), 5–12.
- Young, A. (2016). Improved, Nearly Exact, Statistical Inference with Robust and Clustered Covariance Matrices using Effective Degrees of Freedom Corrections.

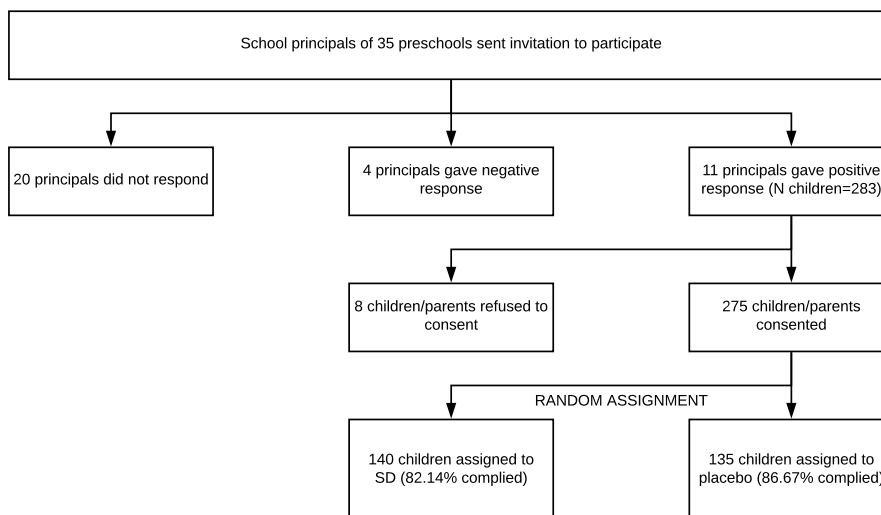
Appendix

Figure A1: Selection of participants Study I



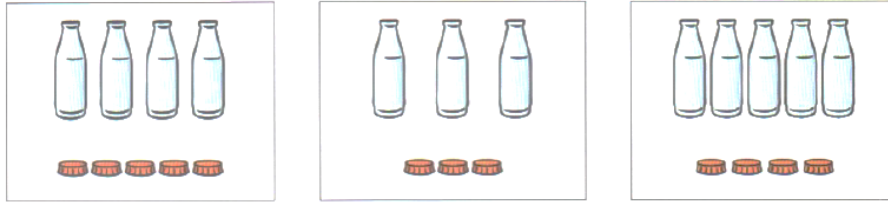
Notes: The final sample consists of 187 preschool children in 10 classes in 6 schools. SD: sugary drink; Placebo: artificially sweetened drink. Non-compliers are children who did not drink their consumption, or only partially.

Figure A2: Selection of participants Study II



Notes: The final sample consists of 232 preschool children in 15 classes in 11 schools. SD: sugary drink; Placebo: artificially sweetened drink. Non-compliers are children who did not drink their consumption, or only partially.

Figure A3: Example question (№8)



Notes: The study administrator reads the following aloud: “Each bottle requires a cap. Which picture indicates a missing cap? Draw a line underneath this picture”.

Table A1: Treatment effect bounds (Lee, 2009) for baseline specification without controls

	Mathematics achievement			Behavior		
	All	Boys	Girls	All	Boys	Girls
Study I (n=167)	0.108 [0.047, 0.129]	-0.339 [-0.386, -0.304]	0.556 [0.494, 0.627]	0.114 [0.067, 0.150]	0.323 [0.285, 0.323]	0.005 [-0.024, 0.039]
Study II (n=232)	0.018 [-0.056, 0.113]	-0.296 [-0.323, -0.296]	0.304 [0.200, 0.376]	0.023 [-0.050, 0.134]	0.093 [0.093, 0.149]	-0.034 [-0.151, 0.105]
Study I + II (n=399)	0.053 [-0.013, 0.119]	-0.299 [-0.332, -0.278]	0.383 [0.300, 0.501]	0.059 [0.000, 0.140]	0.192 [0.144, 0.271]	-0.033 [-0.192, 0.053]

Notes: This table reports treatment effects and associated upper- and lower bounds of drinking one sugary drink compared to drinking an artificially sweetened drink. Treatment effects reported for Study II aggregate results for classes assigned to assessments at different time points, corresponding to 30 minutes, 60 minutes, and 120 minutes after the consumption of drinks (sugar or placebo). Controls: none. Treatment effect bounds (Lee, 2009) are reported in brackets.

Table A2: Comparison of schools

Variable		Contacted sample (participants and non-participants)	Population	<i>P</i> -value
Panel A	% low educated mothers	18.324	21.079	0.143
	% on allowance	21.693	22.341	0.713
	% language not Dutch	17.437	20.203	0.270
		Participants	Non-participants	<i>P</i> -value
Panel B	% low educated mothers	16.524	18.747	0.586
	% on allowance	16.676	22.877	0.155
	% language not Dutch	13.148	18.449	0.215
		Participants	Population	<i>P</i> -value
Panel C	% low educated mothers	16.524	21.079	0.285
	% on allowance	16.676	22.341	0.154
	% language not Dutch	13.148	20.203	0.216

Notes: Comparison of means, and corresponding *p*-values of school composition characteristics when comparing different samples. Panel A compares the contacted sample to the population of schools in Flanders. Panel B compares participating schools to non-participating schools, within the contacted sample. Panel C compares the final sample of participants to the population of schools.

Table A3: Effects of one sugary drink on mathematics achievement and behavior: Alternative specifications

		Mathematics achievement											
		All				Boys				Girls			
Panel A	Study I	0.108	0.154	0.142	0.127	-0.339	-0.093	-0.093	-0.109	0.556	0.379	0.383	0.375
	(n=167)	(0.139)	(0.085)	(0.086)	(0.082)	(0.198)	(0.111)	(0.113)	(0.111)	(0.183)	(0.127)	(0.128)	(0.121)
	Study II	0.018	-0.009	-0.014	-0.008	-0.296	-0.265	-0.274	-0.347	0.304	0.223	0.220	0.215
	(n=232)	(0.101)	(0.080)	(0.081)	(0.078)	(0.149)	(0.117)	(0.120)	(0.116)	(0.134)	(0.105)	(0.107)	(0.109)
Study I + II	0.053	0.052	0.047	0.045	-0.299	-0.194	-0.211	-0.260	0.383	0.268	0.291	0.286	
(n=399)	(0.089)	(0.061)	(0.060)	(0.058)	(0.126)	(0.086)	(0.086)	(0.082)	(0.121)	(0.085)	(0.082)	(0.082)	
Pretreatment scores		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Class FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
		Behavior											
		All				Boys				Girls			
Panel B	Study I	0.114	0.152	0.173	0.178	0.323	0.232	0.235	0.248	0.005	0.129	0.132	0.175
	(n=167)	(0.101)	(0.078)	(0.077)	(0.076)	(0.115)	(0.098)	(0.099)	(0.105)	(0.151)	(0.118)	(0.118)	(0.105)
	Study II	0.023	0.041	0.035	0.036	0.093	0.062	0.058	0.109	-0.034	0.003	-0.015	0.032
	(n=232)	(0.110)	(0.087)	(0.088)	(0.084)	(0.166)	(0.123)	(0.126)	(0.129)	(0.147)	(0.122)	(0.118)	(0.119)
Study I + II	0.059	0.084	0.093	0.095	0.192	0.133	0.129	0.169	-0.033	0.033	0.046	0.092	
(n=399)	(0.077)	(0.062)	(0.061)	(0.059)	(0.109)	(0.082)	(0.084)	(0.086)	(0.110)	(0.091)	(0.085)	(0.082)	
Pretreatment scores		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Class FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	

Notes: This table reports treatment effects of drinking one sugary drink compared to drinking an artificially sweetened drink. Treatment effects reported for Study II aggregate results for classes assigned to assessments at different time points, corresponding to 30 minutes, 60 minutes, and 120 minutes after the consumption of drinks (sugar or placebo). Controls: age, sex, and timing of testing. Robust standard errors in parentheses. The Chi-square test statistic is calculated to examine whether the treatment effect coefficients differ significantly across Study I and Study II.

Table A3: Effects of one sugary drink on mathematics achievement and behavior: Alternative specifications (ctd)

Panel C	Without weight as control		With weight as control	
	Mathematics achievement	Behavior	Mathematics achievement	Behavior
(n=232)				
Treatment	0.254 (0.107)	0.007 (0.121)	0.260 (0.108)	0.014 (0.121)
Boy	0.433 (0.11)	0.008 (0.130)	0.428 (0.113)	0.008 (0.131)
Treatment x Boy	-0.545 (0.158)	0.061 (0.181)	-0.549 (0.159)	0.054 (0.181)
Weight			-0.005 (0.013)	0.006 (0.017)
Panel D	30-45-60 minutes (n=318)		120 minutes (n=81)	
	Mathematics achievement	Behavior	Mathematics achievement	Behavior
Treatment (all)	0.116 (0.063)	0.069 (0.065)	-0.248 (0.149)	0.177 (0.406)
Boys	-0.153 (0.083)	0.159 (0.093)	-0.734 (0.274)	0.209 (0.231)
Girls	0.324 (0.094)	0.036 (0.092)	0.124 (0.190)	0.271 (0.177)

Notes: This table reports treatment effects of drinking one sugary drink compared to drinking an artificially sweetened drink for the subsample of children where weight data was collected. Treatment effects reported for Study II aggregate results for classes assigned to assessments at different time points, corresponding to 30 minutes, 60 minutes, and 120 minutes after the consumption of drinks (sugary drink or placebo). Controls: age, sex, (weight), pretreatment test scores, and class fixed effects, including the timing of the second test. Robust standard errors in parentheses.

Table A4: Robustness of treatment effect (p -values of estimates)

Dependent variable	Mathematics achievement			Behavior		
	All	Boys	Girls	All	Boys	Girls
Unclustered SEs	0.438	0.002	0.001	0.106	0.051	0.261
Clustered SEs	0.402	0.012	0.001	0.194	0.084	0.226
Cameron et al. (2008)	0.305	0.004	0.000	0.156	0.082	0.206
Young (2016)	0.387	0.009	0.000	0.181	0.067	0.199

Notes: Treatment effect p -values are reported under alternative clustering scenarios. Row 1 reports the p -values based on unclustered standard errors, as reported in the main text. Row 2 reports the p -values based on standard errors clustered at the class level. Row 3 reports the p -value based on standard errors clustered at the class level computed using the wild bootstrap method of Cameron et al. (2008). Row 4 reports the p -value based on standard errors clustered at the class level with the degrees of freedom adjustment as in Young (2016). All treatment effects are obtained using the full set of controls (pretreatment scores, age, sex) and class fixed effects, including the timing of the second test, as detailed in Section 2.4.

