Food Quality and Preference 127 (2025) 105435

Contents lists available at ScienceDirect

Food Quality and Preference

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

Trying tongs and spoiling spoons: Effort nudges influence food consumption and may motivate healthier food decisions

Tobias Otterbring [a,b,*], Erik Thomassen [a], Casper Solli Øritsland [a], Gastón Ares [c]

[a] Department of Management and Innovation, University of Agder, 4631 Kristiansand, Norway
[b] School of Electrical Engineering and Computer Science, Division of Media Technology and Interaction Design, KTH Royal Institute of Technology, 10044 Stockholm, Sweden
[c] Sensometrics & Consumer Science, Instituto Polo Tecnológico de Pando, Facultad de Química, Universidad de la República. By Pass de Rutas 8 y 101 s/n, CP 91000 Pando, Uruguay

ARTICLE INFO

Keywords:
Nudging
Choice architecture
Preregistration
Field study
Decision information
Decision structure
Translation
Effort
Satisfaction
Food consumption
Public health

ABSTRACT

Healthier eating is crucial to tackle the rapid rise of obesity and noncommunicable diseases worldwide. This research examined two nudging interventions intended to decrease food consumption: price display and serving utensils. Forecasting experiments showed that people predicted displaying the price of the food per kg (vs. hg) should decrease the amount of food purchased (Study 1 A), but that using tongs (vs. spoon) would be ineffective (Study 1B). In contrast to these results, a high-powered preregistered field study at a university canteen (Study 2) revealed that price display had no notable effect; however, tongs (vs. spoon) reliably decreased the average amount of food purchased per meal by 14 g or 3.1 %, also when compared to weeks when both types of serving utensils were available. An image-supported online experiment with enhanced rigor and control (Study 3) replicated the results regarding tongs (vs. spoon) for a particularly unhealthy food category (candy), while highlighting a psychological mechanism driving the effect. Using tongs required more effort, which decreased satisfaction tied to using said serving utensils, thereby reducing people's willingness to consume candy. Given the simplicity and cost effectiveness of swapping spoons with tongs, combined with the behavioral evidence underscoring its practical relevance, these findings might aid in steering consumers to healthier food decisions, ultimately benefiting public health.

1. Introduction

Overconsumption of food in general, and of unhealthy food in particular (e.g., energy-rich but nutrient-poor ultra-processed food products), is a major contributor to the obesity epidemic (Chaput, Klingenberg, Astrup, & Sjödin, 2011; Lane et al., 2021; Perkovic, Otterbring, Schärli, & Pachur, 2022) and has been liked to a wide array of adverse health outcomes, including cardiovascular diseases, cancer, and all-cause mortality (Pagliai et al., 2021; Stylianou, Fulgoni III, & Jolliet, 2021; Zech, van Dijk, & van Dillen, 2023). Such eating habits incur massive costs, not only to individual consumers but also to society at large.

Redesigning the environment has been identified as a feasible and cost-effective approach to reduce these costs (Callaway, Hardy, & Griffiths, 2023), often through nudges, defined as "any aspect of the choice architecture [the context in which people make decisions] that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives" (Thaler & Sunstein, 2021, p. 8). For an intervention to be considered a nudge, it must be easy and cheap to implement. Nudging involves implementing seemingly small and often unnoticeable interventions in the environment with the goal of facilitating decisions that may benefit individuals, communities, and societies. Despite their subtle nature, they have been used to improve judgment and decision-making across domains, including decisions in health, environmental work, finance, and food contexts (Hummel & Maedche, 2019; Prusaczyk, Earle, & Hodson, 2021; Vecchio & Cavallo, 2019; Stämpfli, Stöckli, & Brunner, 2017; Zumthurm & Stämpfli, 2024).

Although the performance of a nudge crucially depends on the context in which it is implemented, recent research has shown that the food domain is generally where nudging interventions seem to have had the greatest impact; however, the magnitude estimates of such

* Corresponding author at: Department of Management and Innovation, University of Agder, 4631 Kristiansand, Norway
E-mail address: tobias.otterbring@uia.no (T. Otterbring).

https://doi.org/10.1016/j.foodqual.2025.105435
Received 6 December 2024; Received in revised form 13 January 2025; Accepted 14 January 2025
Available online 18 January 2025

interventions differ widely from a moderate-to-large average effect size of $d = 0.65$ (Mertens, Herberz, Hahnel, & Brosch, 2022) to a small effect size of $d = 0.23$ when using experimental field evidence as an inclusion criterion (Cadario & Chandon, 2020). Some scholars have even claimed that, once publication bias has been accounted for, "no evidence for the effectiveness of nudges remains" (Maier et al., 2022, p. 1). Yet others have estimated that plausible effect sizes for various nudging interventions—after the implementation of certain bias-correction methods—fall somewhere on a continuum from virtually zero to $d = 0.08$ (Szaszi et al., 2022).

These vastly different estimates underscore a prevalent problem in the nudging literature: publication bias (Mertens et al., 2022; Szaszi, Goldstein, Soman, & Michie, 2024). In other words, questionable research practices and the file drawer problem—putting studies that show nonsignificant results in the file drawer and only submitting studies that show significant results to journals (Rosenthal, 1979)—might have distorted findings in this stream of research, failing to paint a complete picture of the true effects (Szaszi, Palinkas, Palfi, Szollosi, & Aczel, 2018). Consequently, there is an urgent need for high-quality research in this domain, not least given that public policy is increasingly based on nudging results from academic journals (Sanders, Snijders, & Hallsworth, 2018). To this end, the current research aims to rigorously examine the effectiveness of two distinct nudging interventions using large sample sizes in high-powered studies, preregistered hypotheses, and objective behavioral data from individual consumers in contexts characterized by habitual purchases (Collins et al., 2019; Migliavada, Ricci, & Torri, 2021; Raulio, Roos, & Prättälä, 2010). By running interventions in such settings, we put the validity of our theorizing to a particularly severe test (Mayo, 2018), considering that people rarely change their habits rapidly (Lally, Van Jaarsveld, Potts, & Wardle, 2010). In what follows, we briefly elaborate on our tested interventions.

1.1. Choice Architecture Classifications and Nudging Interventions

Münscher, Vetter, and Scheuerle (2016) proposed a framework to categorize intervention techniques into a series of choice architecture categories. This research focuses on displaying the price of food in different measurement units (hg vs. kg) and altering the serving utensils (tongs vs. spoon). These two nudges fall within the intervention categories of *decision information* and *decision structure*. These categories have been identified as the most effective in altering people's behavior, with average effect sizes when aggregated across decision-making domains of $d = 0.34$ and $d = 0.54$, respectively (Mertens et al., 2022).

Decision information includes increasing the availability, comprehensibility, and/or personal relevance of information by, for example, translating it through adapted attributes to facilitate processing of already available information (Münscher et al., 2016). By contrast, decision structure entails altering the utility of choice options through their arrangement in the environment by, for instance, changing option-related effort, which might be modified by adjusting the physical energy required to reach a certain choice option (Mertens et al., 2022).

Regarding the decision information nudge of varying measurement units, such as featuring the price of food per kg instead of hg, humans tend to focus more on the actual numbers rather than the measurement units through which these numbers are presented. This tendency relates to the numerosity heuristic, whereby humans typically evaluate higher (vs. lower) numbers as more expensive, while overlooking the unit in which the numbers appear (Pandelaere, Briers, & Lembregts, 2011; Pelham, Sumarta, & Myaskovsky, 1994). For example, people often overspend when traveling to a country where the currency is a fraction of one's home currency but underspend when it is a multiplier of said home currency (Lin & Fang, 2013; Raghubir & Srivastava, 2002). Fecher, Robbert, and Roth (2019) found that the same products were viewed as significantly more expensive when presented in price per kg (vs. hg). Similarly, Liu, Thakor, and Chen (2023) found that fruits and vegetables were viewed as more expensive when the price was displayed per kg (vs. hg). The numerosity heuristic also suggests that prices presented in large measurement units, with a higher numerical value, will result in the perception of higher prices (Bagchi & Davis, 2016).

Measurement unit effects like those linked to the numerosity heuristic have typically been examined in one-time hypothetical purchase situations under controlled but artificial conditions (e.g., Lembregts & Pandelaere, 2013; Monga & Bagchi, 2012; Wertenbroch, Soman, & Chattopadhyay, 2007), which differs from our focus on habitual purchase settings in "noisy" real-world contexts. Moreover, such studies are often contingent on factors ranging from the salience of the measurement units (Liu et al., 2023), whether customers have been reminded of alternative units in which information can be expressed (Pandelaere et al., 2011), and the extent to which customers are brand-oriented (Ohlwein, 2023) to customers' level of numeracy, albeit in different directions (Cadario, Parguel, & Benoit-Moreau, 2016; Kleber, Florack, & Chladek, 2016; Tangari, Burton, & Davis, 2014), and whether they are highly focused on restraint (Wilcox & Prokopec, 2019). Such boundary conditions indicate contextual sensitivity in the generality of numerosity effects. Nevertheless, as some studies have verified the effectiveness of numerosity manipulations even in the field (Ofir, Raghubir, Brosh, Monroe, & Heiman, 2008) and with representative samples (Fecher et al., 2019), at least under certain circumstances, we hypothesize that displaying the price of the food per kg (vs. hg) should reduce the amount of food purchased.

As for the decision structure nudge of changing serving utensils from spoon to tongs—and the more general notion of manipulating serving effort or convenience (Cadario & Chandon, 2020; Mertens et al., 2022; Münscher et al., 2016)—several studies indicate that such interventions can indeed influence people's food-related decisions (Garnett, Marteau, Sandbrook, Pilling, & Balmford, 2020; Luo, Li, Soman, & Zhao, 2023; Meiselman, Hedderley, Staddon, Pierson, & Symonds, 1994). With respect to nudges linked to serving utensils, Wansink et al. (2006) found that decreasing the size of an ice-cream scoop from 3 to 2 oz. (from 89 to 59 ml) reduced the amount of ice cream people served themselves by 14.5 %, although this effect was not statistically significant ($p = .10$). Rozin et al. (2011) found that swapping spoons with tongs at a salad bar resulted in a significant reduction in average food consumption by 16.5 %. Kanchanachitra et al. (2020) managed to significantly reduce people's consumption of fish sauce at a university canteen by placing the sauce in a bowl together with a special spoon with multiple holes. Other scholars have tried similar nudges (Brunner, 2013; Sugimoto, Tajiri, Nakashima, & Sakamoto, 2023; Yi, Kanetkar, & Brauer, 2022), with most but not all yielding significant effects. Therefore, we hypothesize that changing the serving utensils from spoon to tongs should also reduce the amount of food purchased and decrease people's willingness to consume the available food options; see Table 1 for a systematic overview of nudges linked to serving utensils or the effort associated with grabbing actual food items.

Rozin et al. (2011) suggested that the most plausible explanation for the reduction in food consumption when swapping spoon with tongs is Tolman's law of least effort, which postulates that choices tend to be made with a "minimum expenditure of physical energy" (Tolman, 1932, p. 448). As many food items likely require more effort when picked using tongs (vs. spoon) and given that effort is inherently aversive (for a meta-analysis, see David, Vassena, & Bijleveld, 2024), changing the serving utensils from spoon to tongs should reasonably decrease customers' satisfaction with using such serving utensils, ultimately reducing the amount of food purchased and their willingness to consume the available food options. The latter is likely considering the positive link between various indicators of customer satisfaction and people's purchase behavior (Otterbring, Arsenovic, Samuelsson, Malodia, & Dhir, 2024; Taylor & Baker, 1994). Accordingly, and consistent with spillover effects, defined as changes in responses on measures other than those primarily targeted (Galizzi & Whitmarsh, 2019; Raufeisen, Wulf, Köcher, Faupel, & Holzmüller, 2019), the increased effort of using tongs

Table 1
Literature overview of nudges linked to serving utensils or the effort associated with grabbing actual food items and their impact on consumers' food-related decisions.

Article	Sample Size(s)[a] and Type(s)	Data Granularity	Main Intervention(s)	Preregistration	Double Difference Design[b]	Mediation[c]	Main Findings
Wansink et al. (2006)	$N = 85$ students and staff members	Individual	Bowls: large vs. small; Spoons: large vs. small (between-subjects)	No	No	No	● People ate more ice cream in large (vs. small) bowls. ● People ate directionally less ice cream with small (vs. large) spoons.
Rozin et al. (2011) (Study 4)	$N = 41$–44 data points, depending on analysis, reflecting the daily average consumption in grams per person at a self-service salad bar	Aggregated	Serving utensils: tongs vs. spoon (between-subjects)	No	No	No	● Using tongs (vs. spoon) reduced food consumption measured in grams at the salad bar.
Brunner (2013)	Study 1: $N = 60$ students Study 2: $N = 63$ students Study 3: $N = 99$ students Study 4: $N = 97$ students	Individual	Study 1: Candy: wrapped vs. unwrapped (between-subjects) Study 2: Grabbing candy: Sugar tongs vs. fingers (between-subjects) Study 3: Grabbing food: sugar tongs vs. fingers; Food: candy vs. fruit (between-subjects) Study 4: Grabbing fruit: sugar tongs vs. fingers; Cognitive load: high vs. low (between-subjects)	No	No	No	● People ate less candy when it was wrapped (vs. unwrapped). ● People ate less candy and dried fruit when they had to use sugar tongs (vs. their own fingers) to pick up these food items. ● People did not eat less or more dried fruit under high (vs. low) cognitive load.
Kanchanachitra et al. (2020)	$N = 124$ data points reflecting the daily amount of fish sauce added in grams divided by the total number of bowls sold that day across five canteens	Aggregated	1. Fish sauce in bottle (baseline) 2. Regular spoon + fish sauce in bowl (behavioral nudge) 3. Special spoon + fish sauce in bowl + information on spoon (behavioral and inconvenience nudge) 4. Fish sauce in bottle + health information (cognitive nudge) 5. Fish sauce in bottle + health information + affectively charged picture (cognitive and affective nudge) (Latin-square)	No	No	No	● Compared to the baseline, people served themselves directionally less fish sauce across all nudging conditions. ● The reduction was most notable in the combined behavioral and inconvenience nudging condition, followed by the behavioral nudge. ● The reduction was marginal in the combined cognitive and affective nudging condition and nonsignificant for the cognitive nudge.
Yi et al. (2022) (Study 5)	$N = 52$ data points reflecting the dollar amount of items sold daily at a self-service salad bar	Aggregated	Spoon size: large vs. small; Plate size order: medium first, large last vs. large first, medium last (between-subjects)	No	No	No	● Neither the plate size order nor the spoon size had any effect on the dollar amount of salad items sold.
Sugimoto et al. (2023)	$N = 36$ students	Individual	Spoon: Perforated with holes vs. regular (crossover between-within-subjects)	No	No	No	● Participants consumed less ramen soup and salt when they ate with the perforated (vs. regular) spoon.

(*continued on next page*)

Table 1 (*continued*)

Article	Sample Size(s)[a] and Type(s)	Data Granularity	Main Intervention(s)	Preregistration	Double Difference Design[b]	Mediation[c]	Main Findings
The current research	Study 1 A: $N = 200$ students and staff members Study 1B: $N = 200$ students and staff members Study 2: $N = 1965$ (preregistered) and $N = 6300$ (supplementary) meal transactions at a canteen Study 3: $N = 1074$ community members	Individual	Study 1 A: Price unit: per hg vs. kg (between-subjects) Study 1B: Serving utensils: tongs vs. spoon (between-subjects) Study 2: Price unit: per hg vs. kg; Serving utensils: tongs vs. spoon (between-subjects) + pre- and post-intervention baseline data Study 3: Serving utensils: tongs vs. spoon (between-subjects)	Yes (Study 2)	Yes (Study 2)	Yes (Study 3)	● People believed that consumers would purchase less food when the price of the food is featured per kg (vs. hg) ● People did not believe that different serving utensils would influence the amount of food purchased by consumers. ● Using tongs (vs. spoon) reduced the amount of food purchased, whereas featuring the price of the food per kg (vs. kg) had no notable effect. ● Grabbing candy with sugar tongs (vs. spoon) was more effortful and reduced individuals' willingness to consume candy. ● Decreased satisfaction of using tongs (vs. spoon) mediated this effect.

Note: The literature above originated from a systematic Scopus title/abstract/keyword search on peer-reviewed journal articles (as of October 25, 2024) using the following search string: (tongs OR spoon* OR "serving utensils") AND (effort OR nudg* OR "choice architecture") AND (consum* OR choice* OR purchas* OR buy*). This procedure resulted in a total of 32 articles, which were independently screened for suitability by the three first authors, while supplemented by their expertise in this topic domain, the latter of which resulted in the inclusion of one additional article (Wansink et al., 2006). Interventions that focused on the impact of serving utensils or the effort associated with grabbing actual food items rather than just making said items easier or more difficult to access or reach (e.g., by placing them closer or farther away through proximity nudges; Garnett et al., 2020) were selected for inclusion, resulting in a final sample of six empirical articles totaling nine studies plus the current research. Interrater reliability, as assessed by Fleiss' kappa, was indicative of excellent agreement between raters ($\kappa = 0.92$) according to common conventions (Cicchetti, 1994). Disagreements were resolved through discussion.

[a] Denotes final sample sizes used in the analyses in main studies, excluding pretests and pilot studies.

[b] Denotes studies that both included treatments and pre- vs. post-intervention measures.

[c] Denotes hypothesized and empirically validated mediation.

(vs. spoon) should reduce people's satisfaction levels rather than solely decreasing their willingness to consume the available food options, with the presumed effect of serving utensils on people's willingness to consume food thus mediated by such satisfaction differences.

2. Overview of Studies

In this paper, we examine predicted and actual effectiveness of our tested interventions: displaying the price of food in different measurement units (hg vs. kg) and altering the serving utensils (tongs vs. spoon). Our two initial studies are designed as forecasting experiments, wherein people intuitively predict whether featuring the price of the food served per kg (vs. hg) would influence the amount of food purchased (Study 1 A; $N = 200$) and whether serving the food using tongs (vs. spoon) would play a role in shaping such consumption responses (Study 1B; $N = 200$). Next, we contrast these forecasts with actual behavioral responses through meal transactions in a large, preregistered field study at a university canteen (Study 2; $N = 1965$ in preregistered analyses and $N = 6300$ in robustness checks). Based on the field study findings, we finally conduct a follow-up online experiment ($N = 1074$) aimed at replicating our focal effect for a particularly unhealthy food category (candy), while testing for a psychological mechanism assumed to drive this effect. Unlike previous research in our topic domain, we use well-powered studies based on individual-level data, including preregistered hypotheses and a formal test of a plausible mediator. These aspects differ notably from previous studies, which have been characterized by small sample sizes, often combined with aggregated data, no formal tests of mechanisms, and no preregistered predictions (see Table 1).

3. Studies 1 A-B: Forecasting Consumer Behavior

These intuition-based experiments consistently included $N = 200$ participants ($n = 100$ per cell) in single-factor between-subjects designs, with price display (hg vs. kg) manipulated in Study 1 A (63.0 % female, 35.5 % male; 1.5 % other/preferred not to disclose; $M_{age} = 23.76$ years, $SD = 4.03$) and with serving utensils (tongs vs. spoon) manipulated in Study 1B (59.5 % female, 40.0 % male; 0.5 % other/preferred not to disclose; $M_{age} = 23.72$ years, $SD = 4.46$). The sample largely comprised university students and staff members at a Northern European university, with the sample sizes being sufficient to detect the typical effect sizes in psychological science and consumer research of approximately $d = 0.40$ (Bosco, Aguinis, Singh, Field, & Pierce, 2015; Eisend, Pol, Niewiadomska, Riley, & Wedgeworth, 2024; Funder & Ozer, 2019) with a power of 80 %, assuming two-tailed tests and the conventional alpha level of $\alpha = 0.05$ ($N = 198$ are required to meet the 80 % power criterion given the above stated assumptions).

Beyond providing basic demographics (gender and age), participants across studies received one forecasting question in a paper-and-pencil survey. We opted for such a forecasting approach given that it has been progressively promoted as a great asset for the development of novel theory across disciplines (DellaVigna, Pope, & Vivalt, 2019; Schaerer et al., 2023; The Forecasting Collaborative, 2023), including but not restricted to consumer research (Mai, Hoffmann, Lasarov, & Buhs, 2019; Nyhus, Frank, Król, & Otterbring, 2024; Philipp-Muller, Costello, & Reczek, 2023), and because people's predictions regarding consumer behavior are not always congruent with customers' actual codes of conduct (Goldstein, Cialdini, & Griskevicius, 2008; Kardes, 2006; Otterbring & Rolschau, 2021).

In Study 1 A, participants received the following question, after

which they were requested to make a prediction on a 7-point Likert scale (1 = less; 4 = no difference; 7 = more), with the text inside the brackets representing a different experimental condition: "Do you think people will purchase more or less taco and salad at the canteen if the price is displayed per kg instead of hg [per hg instead of kg]? In Study 1B, participants replied to the question: "Do you think people will purchase more or less taco and salad at the canteen if the serving utensils used are tongs instead of spoons [spoons instead of tongs]?" Of note, the canteen referred to in Studies 1 A-B and used for data collection purposes in Study 2 normally features the price of the food served per hg rather than kg. That said, it is common to use both these measurement units for a wide variety of food items in Scandinavia. As such, both our measurement units are prevalent in the country where the study took place, which should boost realism and ecological validity (Morales, Amir, & Lee, 2017; Otterbring, Folwarczny, & Gidlöf, 2023).

3.1. Results and discussion

In the first forecasting experiment, an independent samples *t*-test found a significant difference between the hg ($M = 5.15$, $SD = 1.59$) and kg ($M = 3.17$, $SD = 1.53$) conditions, $t(198) = 9.00$, $p < .001$, $d = 1.23$. However, a similar test on the data from the second forecasting experiment yielded no significant difference between the spoon ($M = 4.36$, $SD = 1.24$) and tongs ($M = 4.35$, $SD = 1.37$) conditions, $t(198) = 0.05$, $p = .957$, $d = 0.01$. Thus, participants predicted that most customers would purchase substantially less food when the price of the food is displayed per kg rather than hg (Study 1 A). Moreover, they did not think that the serving utensils would play a major role in shaping the amount of food purchased (Study 1B).

4. Study 2: Consumer Behavior in the Wild

This study was preregistered (https://aspredicted.org/C9W_CM4) and included $N = 1965$ meal transactions with information of the amount of food purchased in grams in the preregistered analyses and $N = 6300$ meal transactions with information of the amount of food purchased in grams in the robustness checks that included the 8 baseline weeks of data collection prior to (4 weeks) and after (4 weeks) the tested interventions. The study met the national regulations for research involving human subjects and the transactions came from the canteen's cash register system, with fully anonymized transactional information without any identifying or personal details (e.g., credit card numbers or banking references).

Given our preregistered reliance on planned contrast analyses to test our formal hypotheses coupled with our one-tailed predictions, we designed the study such that our preregistered sample size would have a statistical power considerably greater than 80 % to detect small effects equivalent to $d = 0.20$, assuming the regular alpha level of $\alpha = 0.05$ ($N = 620$ meal transactions are needed to detect such subtle effects with 80 % power, given the alpha level stated above and one-tailed tests).

Based on historical sales data from the canteen, we specified an estimated total sample size of 2000–3000 meal transactions during the intervention period as part of the preregistration, thus corresponding to approximately 500–750 transactions per condition. This estimate was largely met given that we almost reached 2000 meal transactions ($N = 1965$) during the intervention period, which spanned 4 consecutive Fridays.

In the transaction data, some receipts contained more than one food order. When receipts contained more than one food order, each food order was treated as an individual order, and not as an aggregated one, consistent with what we specified as part of our preregistration. This is because we assumed that each weighted food option corresponds to one plate or bowl of food that was to be consumed by one individual, and that transactions containing two or more items of weighted food correspond to food that was purchased to be consumed by more than one individual.

The study followed a 2 × 2 between-subjects design, with price display (kg vs. hg) and serving utensils (tongs vs. spoon) as the between-subjects factors. Each condition was implemented during a Friday at the canteen, when the food served is always a taco buffet. The decision to keep the food constant across conditions was made to reduce variance from different foods being served. Cell sizes were roughly equal across conditions: $n = 480$ (hg + spoon), $n = 520$ (hg + tongs), $n = 524$ (kg + spoon), $n = 441$ (kg + tongs). To allow the canteen to operate as usual, it was not feasible to randomize treatments on an individual level. Instead, condition assignment was randomized on a weekly basis using a random number generator. The intervention period was implemented during the first Fridays of March 2024 (March 1, 8, 15, and 22). In addition, we used baseline data at the canteen for the four Fridays before the intervention period ($n = 1919$) in February 2024 (February 2, 9, 16, and 23) as well as the four Fridays after the intervention period ($n = 2416$) in April 2024 (April 5, 12, 19, and 26; the canteen was closed the Friday of March 29 due to public Easter holiday) when the food served was identical to the intervention period.

The canteen offers a variety of foods and beverages, including a self-service buffet and a salad bar. It is a hotspot among students, faculty members, and other individuals (e.g., administrative and HR staff) at the university, serving as a place to eat, socialize, and work. The sample represents all customers who purchased a meal at the canteen during the test period.

The price of the food (169 NOK per kg or approximately £11.89 per kg) was shown both on wall-mounted digital displays and printed signs (Figs. 2A-B) at the food counter of the self-serving buffet and the salad bar. In total, the price was shown on four individual digital displays and two printed signs across the canteen. To ensure that the conditions remained the same when switching from price per hg to price per kg, every screen and sign was updated in accordance with the applicable treatment.

Every ingredient across the salad bar and the self-service buffet was equipped with either spoons or tongs (Figs. 3A-B) depending on the specific treatment. The slotted area of the spoons had a length of 9.5 cm, a width of 7.5 cm, and could hold a total volume of 40 ml. The oval area of the tongs had a length of 7.5 cm and a width of 3.5 cm. The spoons with these measurements were used for almost all ingredients, such as ground beef, salad, rice, and salsa. However, a few select ingredients, such as jalapeños and feta cheese, used smaller spoons and tongs for convenience. Further, to enhance realism and maximize the ecological validity of the study setup (Kihlstrom, 2021; Otterbring & Folwarczny, 2024), two deviations from the treatments were consistently implemented: Tongs were always used to grab tortillas and spoons were always used when serving sour cream and other semi-liquid foods.

We took several steps to enhance the internal validity of the study, given the "noisy" nature of field studies. Beyond keeping the food constant across conditions, we collected information about the hourly outdoor temperature at a reliable source (timeanddate.no), which compiles historical weather information like temperature and condition descriptions provided by Custom Weather, the latter of which delivers accurate weather reports from reporting stations worldwide. We then used hourly outdoor temperature as a control variable to account for this potentially confounding factor in our robustness checks. This was deemed important as several studies have documented differential consumer responses as a function of outdoor temperature (Busse, Pope, Pope, & Silva-Risso, 2015; Hirshleifer & Shumway, 2003; Murray, Di Muro, Finn, & Leszczyc, 2010; Schlager, de Bellis, & Hoegg, 2020; Voss, Masuoka, Webber, Scher, & Atkinson, 2013). Moreover, the fact that we relied on highly accurate and reliable scales measuring the amount of food purchased at the precision of grams means that we used objective behavioral data as our focal dependent variable.

4.1. Results and discussion

In direct contrast to the forecasts made by participants in Study 1 A

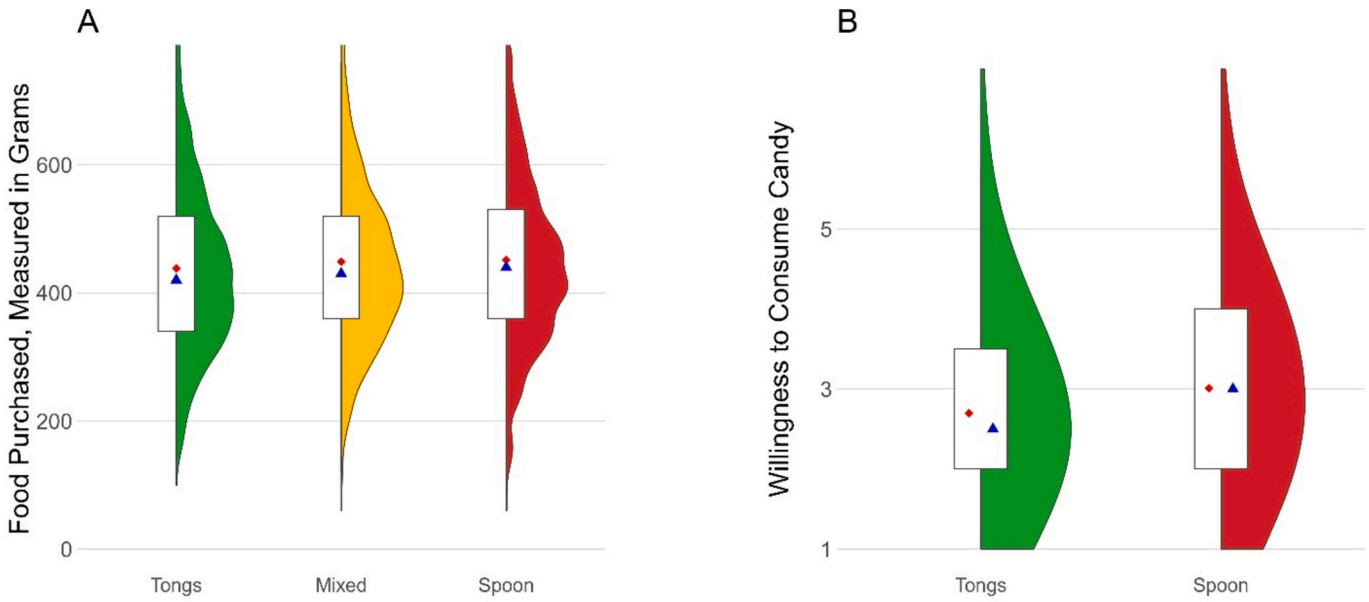

Fig. 1. Results depending on serving utensils across Study 2 (Panel A) and Study 3 (Panel B). *Note.* The figure displays the distribution of food purchased measured in grams (Panel A) and participants' willingness to consume candy (Panel B). The violin plots illustrate the density distribution of the data for each condition, with the wider sections representing a higher concentration of values. The boxplots inside the violins provide additional statistical information, with the 25th percentile reflected by the bottom of the boxplot and the 75th percentile reflected by the top of the boxplot. The black lines extending from the box (whiskers) represent 1.5 times the interquartile range, capturing most of the data excluding extreme values. Red diamonds represent the means for each condition and blue triangles represent the medians. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 2. A: Price sign used in hg conditions. **2B:** Price sign used in kg conditions.

Fig. 3. A: Spoons used in the field study. **3B:** Tongs used in the field study.

and our first preregistered hypothesis, the results of the field study found no significant difference in the amount of food purchased depending on whether the price was featured per hg ($M = 443$ g, $SD = 142$ g) or kg ($M = 447$ g, $SD = 143$ g), as evidenced by a planned contrast analysis, $t(1961) = 0.50$, $p = .616$, $d = 0.02$. Also contrary to the lay beliefs expressed by participants in Study 1B, but consistent with our second preregistered hypothesis, a similar analysis found a significant difference in the amount of food purchased depending on whether customers used tongs ($M = 438$ g, $SD = 141$ g) or spoon ($M = 452$ g, $SD = 145$ g), $t(1961) = 2.02$, $p = .043$, $d = 0.09$.[1] The average reduction of 14 g corresponds to a decrease in the amount of food purchased per meal by 3.1 % compared to when the canteen used spoons as the serving utensils.

Robustness checks with hourly outdoor temperature as a covariate

[1] We consistently preregistered one-sided hypotheses, given the stated direction of our focal predictions; however, as the nature and significance of all our hypothesized results remain unchanged by using two-tailed tests, we report the more conservative two-sided test statistics above and elsewhere in our exposition.

across two one-way analyses of covariance (ANCOVAs), with (1) price display or (2) serving utensils as the between-subjects factor, produced comparable results. The first ANCOVA found no significant effect of price display on the amount of food purchased, $F(1, 1962) = 0.01$, $p = .914$, $\eta^2 < 0.001$. Hourly outdoor temperature was not significantly associated with the amount of food purchased, $F(1, 1962) = 0.17$, $p = .680$, $\eta^2 < 0.001$. The second ANCOVA yielded a significant effect of serving utensils on the amount of food purchased, $F(1, 1962) = 7.38$, $p = .007$, $\eta^2 = 0.004$. Again, hourly outdoor temperature was not significantly associated with the amount of food purchased, $F(1, 1962) = 3.59$, $p = .058$, $\eta^2 = 0.002$.

As the canteen staff typically decides which serving utensils customers can use according to the law of least effort,[2] the canteen normally uses both tongs and spoon depending on individual food items rather than restricting customers to any of these serving utensils. Therefore, we ran another contrast analysis in which we treated the four weeks prior to the intervention as well as the four weeks following the intervention as a baseline period (i.e., the eight Fridays of these weeks when the food served was identical). This analysis enabled us to (1) test whether tongs significantly reduced the amount of food purchased not only compared to the spoon weeks but also when compared against the baseline period, and (2) whether the spoon weeks differed from the baseline period. Planned contrasts revealed that the amount of food purchased was significantly lower when customers used tongs compared to the spoon and baseline weeks, $t(6297) = 2.22$, $p = .027$, $d = 0.08$; however, these latter conditions did not differ significantly, $t(6297) = 0.63$, $p = .529$, $d = 0.01$; see Fig. 1, Panel A. These results were robust to the inclusion of hourly outdoor temperature as a control variable and the results for all our preregistered hypotheses also emerged in comparable nonparametric (Mann-Whitney *U*) tests as well as in analyses whereby outliers that scored more than 3.29 standard deviations beyond the condition-specific means were excluded (Sullivan, Warkentin, & Wallace, 2021; Tabachnik & Fidell, 2019); see the Supplementary Materials for details. The fact that the spoon and baseline weeks did not differ significantly implies that spoon serves as the default option in the canteen, with most food options typically served through spoons unless it is considerably more convenient for customers to use tongs for certain food items.

The accumulated empirical evidence from Study 2 revealed that displaying the price of the food in different measurement units (hg vs. kg) proved ineffective, whereas altering the serving utensils (tongs vs. spoon) significantly influenced the amount of food purchased among thousands of canteen customers. Therefore, we designed a high-powered follow-up experiment characterized by increased rigor and control with the aim of replicating the latter finding for a particularly unhealthy food category (candy), while simultaneously testing for a plausible psychological mechanism that could explain why serving the food through tongs (vs. spoon) should decrease the people's general willingness to consume the applicable food items. Consistent with Tolman's law of least effort, we anticipate that using tongs (vs. spoon) should be perceived as requiring more effort. Considering that effort is inherently aversive, we posit that the presumed effect of serving utensils on people's willingness to consume the available food alternatives—as emerging through our effort manipulation—should be mediated by their satisfaction tied to using said serving utensils, such that tongs (vs. spoon) should decrease satisfaction, thereby reducing their willingness to consume the applicable food items. Study 3 sought to address this possibility.

[2] When we asked the canteen manager how the selection of serving utensils is made, she said that no explicit rule exists depending on individual ingredients but that a decision is usually made based on what is most practically feasible, largely with the point of departure in what makes it easy for customers to grab a given food option.

5. Study 3: Replication and Mediation

This image-supported online experiment was created on Nettskjema, which is a digital and secure survey service run by the University of Oslo that allows for randomization. The survey was distributed via email to a range of recipients whose contact information was obtained from publicly available lists on various Norwegian county municipality websites. These websites provide directories of county employees along with their corresponding contact information. Survey invitations were sent out beginning on August 8, 2024, and responses collected during one week of data collection were used.

The final sample included $N = 1074$ (74.3 % female, 25.7 % male; $M_{age} = 47.03$ years, $SD = 11.96$) community members in the main analyses and a constrained sample of $N = 869$ participants (71.6 % female, 28.4 % male; $M_{age} = 46.94$ years, $SD = 12.03$) with complete BMI information (BMI: $M = 25.53$, $SD = 4.10$; 0.9 % underweight; 50.6 % normal weight; 37.6 % pre-obesity; 8.4 % obesity class I; 1.4 % obesity class II; 1.0 % obesity class III) according to the BMI categorization used by the World Health Organization (World Health Organization, 2024), after excluding two cases whose stated height was below 1 m. The mean BMI score value is very close to the population mean of 25.9 (World Health Organization, 2014), suggesting that the data are relatively representative of the adult Norwegian population, at least in terms of BMI.

The sample size in our main analyses has a statistical power of approximately 95 % to detect small effect sizes equivalent to $d = 0.20$, assuming one-tailed tests and the conventional alpha level of $\alpha = 0.05$ ($N = 1084$ are required to meet the 95 % power criterion given the above stated assumptions). The constrained sample, used for robustness purposes, has a statistical power of roughly 90 % to detect such small effects given the same assumptions ($N = 858$ are required to meet the 90 % power criterion given these assumptions). As such, Study 3 represents yet another high-powered test of our theorizing.

The study originally included 13 additional participants who were excluded for not passing an attention check included in the end of the survey: specifying which serving utensils they were pictorially exposed to at the outset of the study (sugar tongs vs. spoon). Incorrectly stating sugar tongs (spoon) among those who were exposed to spoon (sugar tongs) resulted in a failed attention check and, hence, exclusion from the analyses. Including these participants (1.2 %) in the analyses did not change the nature or significance of the results.

Participants started by reading a brief study description and filling out a digital consent form, after which they were randomly assigned to one of the two experimental conditions in a single-factor between-subjects experiment with serving utensils (sugar tongs vs. spoon) as the between-subjects factor. Next, they replied to demographic details (age and gender) and were pictorially exposed to their assigned serving utensils, after which they were asked to imagine themselves at a store grabbing some pick and mix (candy) using their designated serving utensils (i.e., either sugar tongs or a spoon, depending on the experimental condition). Participants then indicated the amount of candy they would have taken from 12 distinct candy options on a 7-point scale (1 = very little; 7 = very much). We created a composite "willingness to consume candy" index by averaging participants' responses across all 12 items (Cronbach's $\alpha = 0.88$). We used pick and mix items that reflected a balanced selection of chocolates, sweets, and sour candy (Fig. 4), with our chosen items frequently discussed as being among the most popular candy options in the country where the study took place (Ommundsen, 2024).

Following the candy items, we included a single-item measure of our proposed satisfaction mediator, and a manipulation check of effort. Specifically, participants indicated their agreement with the statement, "I am satisfied with picking candy using sugar tongs [a spoon]" (1 = strongly disagree; 7 = strongly agree) as a clear and unambiguous satisfaction measure, with the stated serving utensils reflecting participants' assigned condition. Likewise, they stated their agreement with

Fig. 4. Pick and mix options used as pictorial stimuli in Study 3.

the item, "I find it difficult to pick candy using sugar tongs [a spoon]" (1 = strongly disagree; 7 = strongly agree). We deliberately included the manipulation check (and our proposed mediator) after our focal dependent variable to avoid demand effects associated with hypothesis guessing, considering that manipulation check items that are included prior to the variables of interest can sometime *manipulate* participants' subsequent responses (Ejelöv & Luke, 2020; Frank & Otterbring, 2023; Kühnen, 2010). That said, our mediation results should be interpreted with appropriate caution, given that we did not manipulate the mediator using the causal chain approach (Pirlott & MacKinnon, 2016; Spencer, Zanna, & Fong, 2005). Accordingly, our suggested causal sequence can be called into question on methodological grounds, considering that we measured our proposed mediator after, as opposed to before, our dependent variable. Still, this was a pragmatic choice as we put a greater emphasis on replicating our focal effect instead of providing bulletproof evidence of mediation. Putting our proposed mediator before our dependent variable could have biased the estimates on our focal outcome (Hauser, Ellsworth, & Gonzalez, 2018; Perdue & Summers, 1986; Sigall & Mills, 1998), which made us reach this compromise.

In the end of the survey, participants indicated their current hunger levels ("How hungry are you right now?" 1 = not at all hungry; 7 = very hungry), while also replying to some other items (liking levels of candy, consumption frequency of candy, and the size of their typical meal portions). Finally, they were encouraged to provide information about their height in cm and weight in kg, which we used to calculate their BMI scores.

In our robustness checks, we used the theoretically relevant control variables of participants' gender (Drewnowski, Mennella, Johnson, & Bellisle, 2012), age (De Graaf & Zandstra, 1999), hunger (Otterbring, 2019) with (constrained sample) or without (full sample) their BMI scores (Bobowski & Mennella, 2017) to isolate the effect of our manipulated factor on our key variables of interest (i.e., willingness to consume candy and satisfaction with using one's assigned serving utensils).

5.1. Results and discussion

An independent samples *t*-test verified the effectiveness of the manipulation in terms of perceived effort, $t(998.81) = -14.92$, $p < .001$, $d = -0.92$. Participants in the tongs condition ($M = 4.07$, $SD = 2.03$) versus spoon condition ($M = 2.39$, $SD = 1.64$) indicated higher effort with their assigned serving utensils.

There was also a significant difference in the opposite direction on participants' satisfaction with using their assigned serving utensils, $t(1072) = 14.70$, $p < .001$, $d = 0.90$. Participants in the tongs condition ($M = 3.47$, $SD = 1.87$) indicated lower satisfaction than those in the spoon condition ($M = 5.11$, $SD = 1.79$).

Of particular importance for the current investigation, and in replication of the main finding from Study 2, there was a significant difference between conditions in participants' willingness to consume candy, $t(1072) = 2.98$, $p = .003$, $d = 0.18$. Participants in the tongs condition ($M = 2.55$, $SD = 1.05$) reported a lower willingness to consume candy than those in the spoon condition ($M = 2.74$, $SD = 1.06$); see Fig. 1, Panel B.

A simple mediation model (PROCESS Model 4; Hayes, 2017) tested whether the effect of serving utensils on participants' willingness to consume candy was mediated by their satisfaction ratings. Participants' assigned condition (0 = tongs; 1 = spoon) served as the predictor, satisfaction ratings (continuous) acted as the mediator, and the "willingness to consume candy" index (continuous) served as the outcome. To isolate the effect of our experimental manipulation, we added the theoretically relevant control variables of age (continuous), sex (female = 0; male = 1), and hunger (continuous). However, the nature and significance of our results remain unchanged if these variables are dropped.

Mirroring the above analyses, the effect of serving utensils on participants' satisfaction was significant, $b = 1.64$, SE = 0.11, $t = 14.69$, $p < .001$. In addition, the total effect of serving utensils on participants' willingness to consume candy was significant, $b = 0.19$, SE = 0.06, $t = 3.02$, $p = .003$. Further, satisfaction ratings were positively associated with participants' willingness to consume candy, $b = 0.04$, SE = 0.02, $t = 2.50$, $p = .013$. Importantly, when the "willingness to consume candy" variable was regressed on both the experimental condition and participants' satisfaction ratings, the size of the serving utensils effect on this variable was reduced, $b = 0.12$, SE = 0.07, $t = 1.74$, $p = .082$. A bootstrapping procedure that generated a sample size of 5000 assessed this mediation effect. The result of a 95 % confidence interval (CI) revealed that the indirect effect through participants' satisfaction ratings was significant (b = 0.07, SE = 0.03, 95 % CI = [0.01; 0.13]).

Because participants could choose whether to report their height and weight, we conducted additional robustness checks on the constrained sample ($N = 869$; 80.9 % of the participants used in our main analyses) who provided reliable estimates on these metrics, wherein we additionally controlled for participants' BMI. The nature and significance of all main findings remained unchanged; see the Supplementary Materials for the full model output.

6. General Discussion

In a series of studies across samples, settings, and study paradigms—including data from over 6000 meal transactions—the current research rigorously examined the impact of two nudging interventions commonly classified as pertaining to the decision information and decision structure categories on customers' food-related decisions. More precisely, we tested the effectiveness of displaying the price of the food served per kg (vs. hg) as a way of conceptualizing a decision information nudge (Mertens et al., 2022), and we investigated the impact of swapping serving utensils from spoons to tongs as an operationalization of a decision structure nudge (Münscher et al., 2016).

Studies 1 A-B revealed that people intuitively predicted that featuring the price of the food served per kg (vs. hg) should substantially reduce the amount of food purchased by canteen customers, but that changing the serving utensils from spoon to tongs should not have any noticeable effect. In contrast to these lay beliefs, Study 2 tested these interventions in an actual field setting at a university canteen, finding that price display (kg vs. hg) had no demonstrable effect on the amount of food purchased among canteen customers, but that serving utensils (tongs vs. spoon) robustly influenced the amount of food purchased. Customers purchased significantly lighter meals when grabbing the food with tongs (vs. spoon), also when compared with baseline weeks (pre- and post-intervention), during which both tongs and spoons were available, and spoons served as the default unless a given food item was considerably easier to grab using tongs. These findings mirror recent meta-analytic evidence (Cadario & Chandon, 2020; Mertens et al., 2022), showing that (behavior-oriented) decision structure nudges, such as those influencing the convenience or effort to grab certain food items, are more effective than (cognition-oriented) decision information nudges, such as communicating information in a more easily understandable way. Study 3 replicated these results for a particularly unhealthy food category (candy) in an image-supported online experiment, while also offering empirical evidence for a plausible psychological mechanism that can explain why this effect emerged. Specifically, using tongs (vs. spoon) was perceived as requiring more effort, which decreased the satisfaction tied to using said serving utensils, thereby reducing people's willingness to consume candy.

The results were robust to the inclusion of several theoretically relevant control variables, held across different analytic approaches, and generalized to objective behavioral outcomes (food purchased in grams) measured at the fine-grained level of individual meals in a real-world setting. Given the simplicity and cost effectiveness of swapping spoons with tongs, combined with the behavioral evidence underscoring the practical relevance of relying on this approach, these findings might be useful to nudge customers to healthier food decisions, ultimately benefiting public health.

6.1. Contribution and Implications

Together, the current research makes three central contributions. First, most former studies in our topic domain have been characterized by small sample sizes, often combined with aggregated data such as total sales volume per day rather than more granular data (e.g., individual meal transactions), thereby restricting the effect sizes of interest to moderate or large effects and posing threats to measurement and statistical conclusion validity (Elbæk, Mitkidis, Aarøe, & Otterbring, 2023). This is particularly problematic considering that recent research and meta-analytic evidence on various nudging interventions have concluded that the typical nudging effects are small by conventional standards, at least when publication bias has been accounted for (Cadario & Chandon, 2020; DellaVigna & Linos, 2022; Mertens et al., 2022). Accordingly, most prior research that has documented nudging effects has been under-powered to detect such small effects. In contrast, we present several studies with large sample sizes coupled with objective and highly granular data, thus enabling us to detect even small effect sizes with high statistical power and addressing issues associated with measurement and statistical conclusion validity.

Second, and relatedly, no prior studies in our topic domain seem to have included any preregistrations specifying the target sample size, exclusion criteria, focal hypotheses, and statistical analyses prior to data collection. A scoping review comprising 156 nudging studies found that none of them were preregistered and only 7 % included a power analysis (Szaszi et al., 2018). As the fields of research that have largely developed the nudging literature struggle with relatively low replicability rates (Camerer et al., 2018; Klein et al., 2018) and given that academic nudging results are increasingly used to shape public policy, ensuring high-quality research that can be trusted is paramount (Cologna et al., 2025). Moreover, a recent nudging meta-analysis by Mertens et al. (2022) found clear evidence of publication bias. Similarly, DellaVigna and Linos (2022) compared interventions conducted either by people from nudging units or academic researchers, with the latter group having stronger pressures to publish their work in prestigious outlets. They found that nudges that appeared in academic journals, on average, outperformed a (no nudge) control condition by 8.7 percentage points (33.4 % increased effectiveness), whereas the average impact by the nudging units was substantially smaller, outperforming the control condition by only 1.4 percentage points (8.0 % increased effectiveness). This striking difference was largely attributed to publication bias, with selective publication—amplified by low statistical power—explaining almost 70 % of the difference in effect sizes between the two samples.

Our high-powered and preregistered field study suggests that certain nudges indeed appear to work as intended, even in "noisy" real-world settings, but that their impact might be materially smaller than what has been previously assumed and admittedly smaller than what we originally anticipated. Indeed, our mean and median differences between the tongs and spoon conditions correspond to a 3.1 % or 14 g (mean) and 4.5 % or 20 g (median) reduction in the amount of food purchased per meal, or a modest effect size of $d = 0.09$, making the current research much closer to the estimates from the nudging units, as reported in DellaVigna and Linos (2022), and other recent estimates (Maier et al., 2022; Szaszi et al., 2022; see also Gandhi, Manning, & Duckworth, 2024).

Still, even small effects by arbitrary conventional standards can have massive cumulative consequences at the population level if they are robust, replicable, and scalable (Götz, Gosling, & Rentfrow, 2022). As a case in point, a frequent visitor at the university canteen where our field study took place, who might go there for lunch 4–5 working days a week across an entire year of, say, 44 weeks (weekends, public holidays, and vacation excluded) will purchase roughly 2.8 (mean) to 4.0 (median) kg less food there annually, assuming our estimates are accurate and stable. This decrease might well have certain health implications already at an individual level. If the decrease is rather calculated based on the daily number of meal purchases at the canteen (around 500 a day or 110,000 annually), it would correspond to a reduction of 1.5 (mean) to 2.2 (median) tons of food annually, underscoring the notion that seemingly small effects might be ultimately consequential when implemented at scale (Funder and Ozer, 2019). Moreover, our small effects are to be expected given the smaller size of our serving utensils (i.e., those normally used at the canteen) compared to those used by Rozin et al. (2011), with these scholars finding a decrease in the amount of food purchased by 16.5 % when changing spoons to tongs, but with their spoons being substantially larger (volume capacity: 120 ml) than those used in the current research (volume capacity: 40 ml).

Third, while several studies have speculated about potential mechanisms assumed to be responsible for differences in food consumption occurring as a function of various serving utensils (Brunner, 2013; Rozin et al., 2011; Sugimoto et al., 2023), they have not tested such tentative claims empirically. In contrast, we included a formal manipulation check of effort in our final experimental study, theorizing that as effort is inherently aversive (David et al., 2024), our serving utensils manipulation should influence customers' satisfaction with using their assigned

serving utensils, subsequently driving their willingness to consume candy in a congruent way; an assumption that was verified empirically. Thus, the current research holds effort and the resulting change in satisfaction responsible for the effectiveness of reducing food consumption when swapping spoons with tongs.

From a methodological standpoint, this research also indicates that although certain nudging interventions might be thought of as obvious in their effectiveness (Study 1 A) while others appear to be counterproductive (Study 1B), people fail to forecast which of our tested interventions managed to influence customers' actual behavior in "the real world, where it all starts" (Rozin, 2001, p. 5). This is important from a policy perspective, as our results highlight the need for field evidence before firm conclusions can be drawn regarding the effectiveness of various nudging interventions.

6.2. Limitations and Future Research

From a practical perspective, the current research contributes to the literature by showing the effectiveness of a nudging approach to discourage consumption of food, in general, and unhealthy food, in particular (Harbers et al., 2020). Our findings suggest that regulations requiring the use of specific serving utensils for unhealthy food categories (e.g., at candy shops or for certain unhealthy ingredients at self-service buffets) may contribute to reducing consumption and subsequent intake of such food. Further studies should test whether combining this approach with other strategies, such as information provisioning (e.g., warning labels at the point of purchase), might further boost effectiveness (Ares, Antúnez, Curutchet, & Giménez, 2023).

In our intuition studies, we framed the questions around the extent to which participants expected *others* to behave in a certain way instead of asking how they *themselves* would act. This decision was originally intended to mitigate response bias and provide more accurate estimates, as our selected framing approach often yields more truthful and less self-serving responses (John, Loewenstein, & Prelec, 2012; Klein & Epley, 2016; Zhang & Alicke, 2021). However, we acknowledge that this framing may have inflated our intuition-based results. Indeed, recent research has demonstrated that people generally believe various external cues (e.g., physical objects, consumption goods) to have a greater effect on others than on themselves, while simultaneously inferring that such cues are more useful to others (Polman, Ziano, Wu, & Van Kerckhove, 2022; Ziano & Villanova, 2023). Considering the value of examining consumer lay beliefs (Friestad & Wright, 1995; Kramer, Irmak, Block, & Ilyuk, 2012), future studies should test whether the question framing (others vs. oneself) influences participants' responses.

Much research about measurement units—tapping into our tested decision information nudge—has been conducted in artificial settings through hypothetical scenarios and self-report measures (Fecher et al., 2019; Wertenbroch et al., 2007; Yao & Oppewal, 2016). While such studies have yielded profound insights regarding some cause-effect relationships under conditions characterized by high rigor and control, the contrasting findings from our forecasting experiments and our field study also highlight the importance of methodological pluralism to better understand the true practical implications of nudges. After all, our tested decision information nudge of price display was assumed to have a massive effect yet proved ineffective, while our decision structure nudge was thought to have no noticeable effect at all yet significantly influenced the amount of food purchased by thousands of canteen customers. This is not to question the existence of the numerosity effect or other decision information nudges. However, their impact in real-world settings might be weaker or more context-specific than what is commonly assumed, at least in habitual consumption venues (e.g., university canteens or staff restaurants). As such, while our field study should be viewed as a conservative test of the theorizing underlying measurement unit and numerosity effects, our null findings related to the price display manipulation do not necessarily invalidate these theories.

Given that much of consumers' decision-making in food contexts follows a habitual process (Machín et al., 2020), it could be that the decision information nudge of price display in our field study was ineffective simply because most canteen customers are so used to this context that they did not pay sufficient attention to the external informational cue of whether the food price was featured per kg or hg.[3] Stated differently, those who noticed this cue might well have modified the amount of food purchased, although most canteen customers possibly just made routine purchases without considering this cue. This interpretation might also explain the superiority of our decision structure nudge, as manipulations of the specific serving utensils used do not require customers' overt attention to influence their purchase behavior. As such, our results underscore the need for nudging strategies that rely on automatic behavior, without requiring attention and cognitive processing.

Some studies have documented licensing effects on consumer choice (Khan & Dhar, 2006; Rishika, Feurer, & Haws, 2022), whereby consumers who initially choose or consume virtuous products (e.g., healthy) use those decisions to subsequently justify the consumption of more vice or indulgent products (e.g., tasty but unhealthy). Indeed, this compensatory licensing accout has been frequently observed in the food domain. For instance, simply sampling a healthy (vs. unhealthy) food item seems to boost subsequent purchases and choices of unhealthy food products (Biswas, Abell, Lim, Inman, & Held, 2024). Moreover, behavioral interventions designed to promote two distinct healthy behaviors (meal logging and meditation) paradoxically reduce completion rates of the opposite behavior by 19–29 % (Trachtman, 2024). Even low-fat nutrition labels can amplify food intake, particularly among overweight consumers, by increasing perceptions of what constitutes an appropriate serving size and decreasing guilt tied to consumption (Wansink & Chandon, 2006). Additionally, references to health can create a "health halo," influencing consumer behavior (Ares et al., 2023). For example, relative to fast-food chains that do not make health claims (e.g., McDonald's), comparable chains claimed to be healthy (e.g., Subway) leads consumers to underestimate the calorie content of main dishes, ultimately resulting in a greater choice likelihood of higher-calorie side dishes, drinks, or desserts (Chandon & Wansink, 2007). As the current research focused on one-time consumption occasions, we cannot determine how the serving utensils manipulation operates longitudinally or sequentially. Although prior studies suggest that effort nudges can significantly reduce food intake without the risk of subsequent compensatory behaviors in terms of, for instance, increased food cravings (Maas, de Ridder, de Vet, & De Wit, 2012), future research should seek to ascertain that our tested interventions do not generate unintended (unhealthy) rebound effects on customers' subsequent purchase or choice behavior.

A critic might argue that few if any restaurant managers would willingly introduce interventions likely to increase customer complaints, as indicated by our decrease in serving satisfaction when swapping spoons with tongs. Yet, we contend that subtle nudges can be successfully incorporated without knowingly annoying customers even if said nudges slightly increase the effort required to grab certain food items. In support of this notion, several other effort nudges have already been implemented by restaurants across the globe, such as placing healthy

[3] In support of this notion, the canteen increased the price of the food right after our intervention from approximately £11.89 to £12.59 per kg without any significant impact on the average amount of food purchased by customers, implying that most customers might not have noticed the raised price. Indeed, the average amount of food purchased during the four baseline weeks prior to the intervention (M = 552 g, SD = 145 g), when the price of the food was £11.89 per kg, did not differ from the four baseline weeks following the intervention (M = 446 g, SD = 146 g), when the price of the food was increased to £12.59 per kg, t (4333) = 1.25, p = .211, d = 0.04.

options closer and unhealthy farther away (Alinia et al., 2011; Garnett, Balmford, Sandbrook, Pilling, & Marteau, 2019; Rozin et al., 2011; Wansink & Hanks, 2013). This suggests that restaurant managers are, in fact, willing to make certain food items more difficult to reach or obtain, even if it implies increased effort from target customers.

In our field study, the transaction data were restricted to the total quantity of the food purchased, meaning that we cannot know with certainty how our focal effect operates at a level of even higher data granularity (i.e., individual ingredients). Although most prior related studies have found effort nudges to be equally effective across healthy and unhealthy food categories (Brunner, 2013; Rozin et al., 2011; Sugimoto et al., 2023), high-powered preregistered studies are warranted before definite conclusions can be reached regarding the generality of our tested interventions across healthy and unhealthy food items. Assuming that our results are applicable irrespective of the health status of the food categories offered, as indicated by our final study, it might be that serving utensils can be considered for financial optimization and waste reduction purposes by restaurant managers (e.g., at all-you-can-eat buffets with a fixed price regardless of the amount of food consumed) in a way that simultaneously benefits public or planetary health (Folwarczny, Otterbring, Sigurdsson, Tan, & Li, 2023; Hagmann, Ho, & Loewenstein, 2019). With a more detailed cash register system than that used in the current research, we could have tested and designed nudges tailored at specific ingredients or food categories, enabling us to develop more precise and targeted interventions than those allowed by our transaction data. For example, spoons and tongs could have been strategically allocated across different food categories and ingredients to encourage consumers to serve themselves more sustainable and healthy food options with high profit margins (e.g., certain locally or regionally produced vegetarian food items served through simple spoons), while discouraging consumption of unsustainable and unhealthy food alternatives with lower profit margins (e.g., some meat options originating from a distant continent served through tricky tongs). Future research should address this possibility.

Ethics statement

This research complied with the local and national guidelines for nonmedical research involving human subjects.

CRediT authorship contribution statement

Tobias Otterbring: Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Formal analysis, Conceptualization. **Erik Thomassen:** Writing – review & editing, Writing – original draft, Investigation, Data curation, Conceptualization. **Casper Solli Øritsland:** Writing – review & editing, Writing – original draft, Investigation, Data curation, Conceptualization. **Gastón Ares:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors are grateful to Michał Folwarczny for helping with Fig. 1 and to Sia Mat og Drikke for allowing us to use your facilities and for assistance and cooperation in the field study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodqual.2025.105435.

Data availability

Data will be made available on request.

References

Alinia, S., Lassen, A. D., Krogholm, K. S., Christensen, T., Hels, O. H., & Tetens, I. (2011). A workplace feasibility study of the effect of a minimal fruit intervention on fruit intake. *Public Health Nutrition, 14*(8), 1382–1387.

Ares, G., Antúnez, L., Curutchet, M. R., & Giménez, A. (2023). Warning labels as a policy tool to encourage healthier eating habits. *Current Opinion in Food Science, 51*, Article 101011.

Ares, G., Natero, V., Gugliucci, V., Machín, L., Alcaire, F., de León, C., & Otterbring, T. (2023). Health-washing of Ultraprocessed products on Instagram: Prevalence and strategies in an emerging market. *Journal of Nutrition Education and Behavior, 55*(11), 815–822.

Bagchi, R., & Davis, D. F. (2016). The role of numerosity in judgments and decision-making. *Current Opinion in Psychology, 10*, 89–93.

Biswas, D., Abell, A., Lim, M., Inman, J. J., & Held, J. (2024). Effects of sampling healthy versus unhealthy foods on subsequent food purchases. *Journal of the Academy of Marketing Science*, 1–16.

Bobowski, N., & Mennella, J. A. (2017). Personal variation in preference for sweetness: Effects of age and obesity. *Childhood Obesity, 13*(5), 369–376.

Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology, 100*(2), 431–449.

Brunner, T. A. (2013). It takes some effort. How minimal physical effort reduces consumption volume. *Appetite, 71*, 89–94.

Busse, M. R., Pope, D. G., Pope, J. C., & Silva-Risso, J. (2015). The psychological effect of weather on car purchases. *The Quarterly Journal of Economics, 130*(1), 371–414.

Cadario, R., & Chandon, P. (2020). Which healthy eating nudges work best? A meta-analysis of field experiments. *Marketing Science, 39*(3), 465–486.

Cadario, R., Parguel, B., & Benoit-Moreau, F. (2016). Is bigger always better? The unit effect in carbon emissions information. *International Journal of Research in Marketing, 33*(1), 204–207.

Callaway, F., Hardy, M., & Griffiths, T. L. (2023). Optimal nudging for cognitively bounded agents: A framework for modeling, predicting, and controlling the effects of choice architectures. *Psychological Review, 130*(6), 1457–1491.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., & Wu, H. (2018). Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. *Nature Human Behaviour, 2*(9), 637–644.

Chandon, P., & Wansink, B. (2007). The biasing health halos of fast-food restaurant health claims: Lower calorie estimates and higher side-dish consumption intentions. *Journal of Consumer Research, 34*(3), 301–314.

Chaput, J. P., Klingenberg, L., Astrup, A., & Sjödin, A. M. (2011). Modern sedentary activities promote overconsumption of food in our current obesogenic environment. *Obesity Reviews, 12*(5), e12–e20.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284–290.

Collins, E. I., Thomas, J. M., Robinson, E., Aveyard, P., Jebb, S. A., Herman, C. P., & Higgs, S. (2019). Two observational studies examining the effect of a social norm and a health message on the purchase of vegetables in student canteen settings. *Appetite, 132*, 122–130.

Cologna, V., Mede, N. G., Berger, S., et al. (2025). Trust in scientists and their role in society across 68 countries. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-024-02090-5

David, L., Vassena, E., & Bijleveld, E. (2024). The unpleasantness of thinking: A meta-analytic review of the association between mental effort and negative affect. *Psychological Bulletin, 150*(9), 1070–1093.

De Graaf, C., & Zandstra, E. H. (1999). Sweetness intensity and pleasantness in children, adolescents, and adults. *Physiology & Behavior, 67*(4), 513–520.

DellaVigna, S., & Linos, E. (2022). RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica, 90*(1), 81–116.

DellaVigna, S., Pope, D., & Vivalt, E. (2019). Predict science to improve science. *Science, 366*(6464), 428–429.

Drewnowski, A., Mennella, J. A., Johnson, S. L., & Bellisle, F. (2012). Sweetness and food preference. *The Journal of Nutrition, 142*(6), 1142S–1148S.

Eisend, M., Pol, G., Niewiadomska, D., Riley, J., & Wedgeworth, R. (2024). How much have we learned about consumer research? *A Meta-Meta-Analysis. Journal of Consumer Research, 51*(1), 180–190.

Ejelöv, E., & Luke, T. J. (2020). "Rarely safe to assume": Evaluating the use and interpretation of manipulation checks in experimental social psychology. *Journal of Experimental Social Psychology, 87*, Article 103937.

Elbæk, C. T., Mitkidis, P., Aarøe, L., & Otterbring, T. (2023). Subjective socioeconomic status and income inequality are associated with self-reported morality across 67 countries. *Nature Communications, 14*(1), 5453.

Fecher, A., Robbert, T., & Roth, S. (2019). Same price, different perception: Measurement-unit effects on price-level perceptions and purchase intentions. *Journal of Retailing and Consumer Services, 49*, 129–142.

Folwarczny, M., Otterbring, T., Sigurdsson, V., Tan, L. K. L., & Li, N. P. (2023). Old minds, new marketplaces: How evolved psychological mechanisms trigger mismatched food preferences. *Evolutionary Behavioral Sciences, 17*(1), 93–101.

Frank, D. A., & Otterbring, T. (2023). Being seen… by human or machine? Acknowledgment effects on customer responses differ between human and robotic service workers. *Technological Forecasting and Social Change, 189*, Article 122345.
Friestad, M., & Wright, P. (1995). Persuasion knowledge: Lay people's and researchers' beliefs about the psychology of advertising. *Journal of Consumer Research, 22*(1), 62–74.
Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science, 2*(2), 156–168.
Galizzi, M. M., & Whitmarsh, L. (2019). How to measure behavioral spillovers: A methodological review and checklist. *Frontiers in Psychology, 10*, 342.
Gandhi, L., Manning, B. S., & Duckworth, A. L. (2024). Effect size magnification: No variable is as important as the one you're thinking about—While you're thinking about it. *Current Directions in Psychological Science, 09637214241268222*.
Garnett, E. E., Balmford, A., Sandbrook, C., Pilling, M. A., & Marteau, T. M. (2019). Impact of increasing vegetarian availability on meal selection and sales in cafeterias. *Proceedings of the National Academy of Sciences, 116*(42), 20923–20929.
Garnett, E. E., Marteau, T. M., Sandbrook, C., Pilling, M. A., & Balmford, A. (2020). Order of meals at the counter and distance between options affect student cafeteria vegetarian sales. *Nature Food, 1*(8), 485–488.
Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research, 35*(3), 472–482.
Götz, F. M., Gosling, S. D., & Rentfrow, P. J. (2022). Small effects: The indispensable foundation for a cumulative psychological science. *Perspectives on Psychological Science, 17*(1), 205–215.
Hagmann, D., Ho, E. H., & Loewenstein, G. (2019). Nudging out support for a carbon tax. *Nature Climate Change, 9*(6), 484–489.
Harbers, M. C., Beulens, J. W., Rutters, F., De Boer, F., Gillebaart, M., Sluijs, I., & Van Der Schouw, Y. T. (2020). The effects of nudges on purchases, food choice, and energy intake or content of purchases in real-life food purchasing environments: A systematic review and evidence synthesis. *Nutrition Journal, 19*, 1–27.
Hauser, D. J., Ellsworth, P. C., & Gonzalez, R. (2018). Are manipulation checks necessary? *Frontiers in Psychology, 9*, 998.
Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. Guilford publications*.
Hirshleifer, D., & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *Journal of Finance, 58*(3), 1009–1032.
Hummel, D., & Maedche, A. (2019). How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics, 80*, 47–58.
John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*(5), 524–532.
Kanchanachitra, M., Chamchan, C., Kanchanachitra, C., Suttikasem, K., Gunn, L., & Vlaev, I. (2020). Nudge interventions to reduce fish sauce consumption in Thailand. *PloS One, 15*(9), Article e0238642.
Kardes, F. R. (2006). When should consumers and managers trust their intuition? *Journal of Consumer Psychology, 16*(1), 20–24.
Khan, U., & Dhar, R. (2006). Licensing effect in consumer choice. *Journal of Marketing Research, 43*(2), 259–266.
Kihlstrom, J. F. (2021). Ecological validity and "ecological validity". *Perspectives on Psychological Science, 16*(2), 466–471.
Kleber, J., Florack, A., & Chladek, A. (2016). How to present donations: The moderating role of numeracy in cause-related marketing. *Journal of Consumer Marketing, 33*(3), 153–161.
Klein, N., & Epley, N. (2016). Maybe holier, but definitely less evil, than you: Bounded self-righteousness in social judgment. *Journal of Personality and Social Psychology, 110*(5), 660–674.
Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Jr., Alper, S., & Sowden, W. (2018). Many labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science, 1*(4), 443–490.
Kramer, T., Irmak, C., Block, L. G., & Ilyuk, V. (2012). The effect of a no-pain, no-gain lay theory on product efficacy perceptions. *Marketing Letters, 23*, 517–529.
Kühnen, U. (2010). Manipulation checks as manipulation: Another look at the ease-of-retrieval heuristic. *Personality and Social Psychology Bulletin, 36*(1), 47–58.
Lally, P., Van Jaarsveld, C. H., Potts, H. W., & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology, 40*(6), 998–1009.
Lane, M. M., Davis, J. A., Beattie, S., Gómez-Donoso, C., Loughman, A., O'Neil, A., & Rocks, T. (2021). Ultraprocessed food and chronic noncommunicable diseases: A systematic review and meta-analysis of 43 observational studies. *Obesity Reviews, 22*(3), Article e13146.
Lembregts, C., & Pandelaere, M. (2013). Are all units created equal? The effect of default units on product evaluations. *Journal of Consumer Research, 39*(6), 1275–1289.
Lin, Y. C., & Fang, S. H. (2013). The face value of foreign currency on consumer price perception—The moderating effect of product substitution. *Journal of Business Research, 66*(6), 745–751.
Liu, Y., Thakor, M., & Chen, R. (2023). Is 8.8 per kilogram more expensive than 3.99 per pound? An investigation of the effect of measurement units on price perceptions. *Canadian Journal of Administrative Sciences, 40*(4), 407–424.
Luo, Y., Li, A., Soman, D., & Zhao, J. (2023). A meta-analytic cognitive framework of nudge and sludge. *Royal Society Open Science, 10*(11), Article 230053.
Maas, J., de Ridder, D. T., de Vet, E., & De Wit, J. B. (2012). Do distant foods decrease intake? The effect of food accessibility on consumption. *Psychology & Health, 27*(sup2), 59–73.
Machín, L., Curutchet, M. R., Gugliucci, V., Vitola, A., Otterbring, T., de Alcantara, M., & Ares, G. (2020). The habitual nature of food purchases at the supermarket: Implications for policy making. *Appetite, 155*, Article 104844.
Mai, R., Hoffmann, S., Lasarov, W., & Buhs, A. (2019). Ethical products = less strong: How explicit and implicit reliance on the lay theory affects consumption behaviors. *Journal of Business Ethics, 158*, 659–677.
Maier, M., Bartoš, F., Stanley, T. D., Shanks, D. R., Harris, A. J., & Wagenmakers, E. J. (2022). No evidence for nudging after adjusting for publication bias. *Proceedings of the National Academy of Sciences, 119*(31), Article e2200300119.
Mayo, D. G. (2018). *Statistical inference as severe testing: How to get beyond the statistics wars*. Cambridge University Press.
Meiselman, H. L., Hedderley, D., Staddon, S. L., Pierson, B. J., & Symonds, C. R. (1994). Effect of effort on meal selection and meal acceptability in a student cafeteria. *Appetite, 23*(1), 43–55.
Mertens, S., Herberz, M., Hahnel, U. J., & Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences, 119*(1), Article e2107346118.
Migliavada, R., Ricci, F. Z., & Torri, L. (2021). A three-year longitudinal study on the use of pre-ordering in a university canteen. *Appetite, 163*, Article 105203.
Monga, A., & Bagchi, R. (2012). Years, months, and days versus 1, 12, and 365: The influence of units versus numbers. *Journal of Consumer Research, 39*(1), 185–198.
Morales, A. C., Amir, O., & Lee, L. (2017). Keeping it real in experimental research—Understanding when, where, and how to enhance realism and measure consumer behavior. *Journal of Consumer Research, 44*(2), 465–476.
Münscher, R., Vetter, M., & Scheuerle, T. (2016). A review and taxonomy of choice architecture techniques. *Journal of Behavioral Decision Making, 29*(5), 511–524.
Murray, K. B., Di Muro, F., Finn, A., & Leszczyc, P. P. (2010). The effect of weather on consumer spending. *Journal of Retailing and Consumer Services, 17*(6), 512–520.
Nyhus, E. K., Frank, D. A., Król, M. K., & Otterbring, T. (2024). Crypto cravings: Gender differences in crypto investment intentions and the mediating roles of financial overconfidence and personality. *Psychology & Marketing, 41*(3), 447–464.
Ofir, C., Raghubir, P., Brosh, G., Monroe, K. B., & Heiman, A. (2008). Memory-based store price judgments: the role of knowledge and shopping experience. *Journal of Retailing, 84*(4), 414–423.
Ohlwein, M. (2023). Preferences as a determinant of the measurement-unit effect for unit prices. *International Journal of Market Research, 65*(6), 677–686.
Ommundsen, A. (2024). The candy key: These are the Norwegians' favorites. *Godt*, March 25, 2024 Available at: https://www.godt.no/aktuelt/i/BW5bo0/smaagodt-dette-er-nordmenns-favoritter(accessed August 1, 2024).
Otterbring, T. (2019). Time orientation mediates the link between hunger and hedonic choices across domains. *Food Research International, 120*, 124–129.
Otterbring, T., Arsenovic, J., Samuelsson, P., Malodia, S., & Dhir, A. (2024). Going the extra mile, now or after a while: The impact of employee proactivity in retail service encounters on customers' shopping responses. *British Journal of Management, 35*(3), 1425–1448.
Otterbring, T., & Folwarczny, M. (2024). Social validation, reciprocation, and sustainable orientation: Cultivating "clean" codes of conduct through social influence. *Journal of Retailing and Consumer Services, 76*, Article 103612.
Otterbring, T., Folwarczny, M., & Gidlöf, K. (2023). Hunger effects on option quality for hedonic and utilitarian food products. *Food Quality and Preference, 103*, Article 104693.
Otterbring, T., & Rolschau, K. (2021). Beauty is in the eye of the beer holder but rarely because of the beer. *Personality and Individual Differences, 179*, Article 110921.
Pagliai, G., Dinu, M., Madarena, M. P., Bonaccio, M., Iacoviello, L., & Sofi, F. (2021). Consumption of ultra-processed foods and health status: A systematic review and meta-analysis. *British Journal of Nutrition, 125*(3), 308–318.
Pandelaere, M., Briers, B., & Lembregts, C. (2011). How to make a 29% increase look bigger: The unit effect in option comparisons. *Journal of Consumer Research, 38*(2), 308–322.
Pelham, B. W., Sumarta, T. T., & Myaskovsky, L. (1994). The easy path from many to much: The numerosity heuristic. *Cognitive Psychology, 26*(2), 103–133.
Perdue, B. C., & Summers, J. O. (1986). Checking the success of manipulations in marketing experiments. *Journal of Marketing Research, 23*(4), 317–326.
Perkovic, S., Otterbring, T., Schärli, C., & Pachur, T. (2022). The perception of food products in adolescents, lay adults, and experts: A psychometric approach. *Journal of Experimental Psychology: Applied, 28*(3), 555–575.
Philipp-Muller, A., Costello, J. P., & Reczek, R. W. (2023). Get your science out of here: When does invoking science in the marketing of consumer products backfire? *Journal of Consumer Research, 49*(5), 721–740.
Pirlott, A. G., & MacKinnon, D. P. (2016). Design approaches to experimental mediation. *Journal of Experimental Social Psychology, 66*, 29–38.
Polman, E., Ziano, I., Wu, K., & Van Kerckhove, A. (2022). Consumers believe that products work better for others. *Journal of Consumer Research, 49*(1), 132–153.
Prusaczyk, E., Earle, M., & Hodson, G. (2021). A brief nudge or education intervention delivered online can increase willingness to order a beef-mushroom burger. *Food Quality and Preference, 87*, Article 104045.
Raghubir, P., & Srivastava, J. (2002). Effect of face value on product valuation in foreign currencies. *Journal of Consumer Research, 29*(3), 335–347.
Raufeisen, X., Wulf, L., Köcher, S., Faupel, U., & Holzmüller, H. H. (2019). Spillover effects in marketing: Integrating core research domains. *AMS Review, 9*, 249–267.
Raulio, S., Roos, E., & Prättälä, R. (2010). School and workplace meals promote healthy food habits. *Public Health Nutrition, 13*(6A), 987–992.

Rishika, R., Feurer, S., & Haws, K. L. (2022). Really rewarding rewards: Strategic licensing in long-term healthy food consumption. *Journal of Consumer Research, 49*(2), 268–287.
Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641.
Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review, 5*(1), 2–14.
Rozin, P., Scott, S., Dingley, M., Urbanek, J. K., Jiang, H., & Kaltenbach, M. (2011). Nudge to nobesity I: Minor changes in accessibility decrease food intake. *Judgment and Decision Making, 6*(4), 323–332.
Sanders, M., Snijders, V., & Hallsworth, M. (2018). Behavioural science and policy: Where are we now and where are we going? *Behavioural Public Policy, 2*(2), 144–167.
Schaerer, M., Du Plessis, C., Nguyen, M. H. B., Van Aert, R. C., Tiokhin, L., Lakens, D., & Gender Audits Forecasting Collaboration. (2023). On the trajectory of discrimination: A meta-analysis and forecasting survey capturing 44 years of field experiments on gender and hiring decisions. *Organizational Behavior and Human Decision Processes, 179*, Article 104280.
Schlager, T., de Bellis, E., & Hoegg, J. (2020). How and when weather boosts consumer product valuation. *Journal of the Academy of Marketing Science, 48*, 695–711.
Sigall, H., & Mills, J. (1998). Measures of independent variables and mediators are useful in social psychology experiments: But are they necessary? *Personality and Social Psychology Review, 2*(3), 218–226.
Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology, 89*(6), 845–851.
Stämpfli, A. E., Stöckli, S., & Brunner, T. A. (2017). A nudge in a healthier direction: How environmental cues help restrained eaters pursue their weight-control goal. *Appetite, 110*, 94–102.
Stylianou, K. S., Fulgoni, V. L., III, & Jolliet, O. (2021). Small targeted dietary changes can yield substantial gains for human health and the environment. *Nature Food, 2*(8), 616–627.
Sugimoto, M., Tajiri, E., Nakashima, N., & Sakamoto, T. (2023). Effects of Using a Perforated Spoon on Salt Reduction When Consuming Ramen Noodles: A Randomized Crossover Study of Japanese Male University Students. *Nutrients, 15*(13), 2864.
Sullivan, J. H., Warkentin, M., & Wallace, L. (2021). So many ways for assessing outliers: What really works and does it matter? *Journal of Business Research, 132*, 530–543.
Szaszi, B., Goldstein, D. G., Soman, D., & Michie, S. (2024). Generalizability and effectiveness of choice architecture interventions. *PsyArXiv. Available at:*. https://doi.org/10.31234/osf.io/8p4vq (January 12, 2025)
Szaszi, B., Higney, A., Charlton, A., Gelman, A., Ziano, I., Aczel, B., & Tipton, E. (2022). No reason to expect large and consistent effects of nudge interventions. *Proceedings of the National Academy of Sciences, 119*(31), Article e2200732119.
Szaszi, B., Palinkas, A., Palfi, B., Szollosi, A., & Aczel, B. (2018). A systematic scoping review of the choice architecture movement: Toward understanding when and why nudges work. *Journal of Behavioral Decision Making, 31*(3), 355–366.
Tabachnik, B., & Fidell, L. (2019). *Using multivariate statistics*. Pearson.
Tangari, A. H., Burton, S., & Davis, C. (2014). Do they have your number? Understanding the moderating role of format effects and consumer numeracy for quantitative front-of-package nutrition claims. *Journal of Consumer Affairs, 48*(3), 620–633.
Taylor, S. A., & Baker, T. L. (1994). An assessment of the relationship between service quality and customer satisfaction in the formation of consumers' purchase intentions. *Journal of Retailing, 70*(2), 163–178.
Thaler, R. H., & Sunstein, C. R. (2021). *Nudge: The* (final ed.). Yale University Press.
The Forecasting Collaborative. (2023). Insights into the accuracy of social scientists' forecasts of societal change. *Nature Human Behaviour, 7*(4), 484–501.
Tolman, E. C. (1932). *Purposive behavior in animals and men*. University of California Press.
Trachtman, H. (2024). Does promoting one healthy behavior detract from others? Evidence from a field experiment. *American Economic Journal: Applied Economics, 16*(2), 249–277.
Vecchio, R., & Cavallo, C. (2019). Increasing healthy food choices through nudges: A systematic review. *Food Quality and Preference, 78*, Article 103714.
Voss, J. D., Masuoka, P., Webber, B. J., Scher, A. I., & Atkinson, R. L. (2013). Association of elevation, urbanization and ambient temperature with obesity prevalence in the United States. *International Journal of Obesity, 37*(10), 1407–1412.
Wansink, B., & Chandon, P. (2006). Can "low-fat" nutrition labels lead to obesity? *Journal of Marketing Research, 43*(4), 605–617.
Wansink, B., & Hanks, A. S. (2013). Slim by design: Serving healthy foods first in buffet lines improves overall meal selection. *PLoS One, 8*(10), Article e77055.
Wansink, B., Van Ittersum, K., & Painter, J. E. (2006). Ice cream illusions: Bowls, spoons, and self-served portion sizes. *American Journal of Preventive Medicine, 31*(3), 240–243.
Wertenbroch, K., Soman, D., & Chattopadhyay, A. (2007). On the perceived value of money: The reference dependence of currency numerosity effects. *Journal of Consumer Research, 34*(1), 1–10.
Wilcox, K., & Prokopec, S. (2019). Restraint that blinds: Attention narrowing and consumers' response to numerosity in self-control decisions. *Journal of Consumer Research, 46*(2), 371–387.
World Health Organization. (2014). Global status report on noncommunicable diseases 2014. Available at: https://www.who.int/publications/i/item/9789241564854 (accessed October 26, 2024).
World Health Organization. (2024). A healthy lifestyle - WHO recommendations. Available at: https://www.who.int/europe/news-room/fact-sheets/item/a-healthy-lifestyle—who-recommendations (accessed October 23, 2024).
Yao, J., & Oppewal, H. (2016). Unit pricing increases price sensitivity even when products are of identical size. *Journal of Retailing, 92*(1), 109–121.
Yi, S., Kanetkar, V., & Brauer, P. (2022). Nudging food service users to choose fruit-and vegetable-rich items: Five field studies. *Appetite, 173*, 105978.
Zech, H. G., van Dijk, W. W., & van Dillen, L. F. (2023). Food approach dynamics in daily life: Speed and force of food approach movements fluctuate with hunger, but less so for people with high BMI. *Journal of Experimental Psychology: General, 152*(8), 2300–2317.
Zhang, Y., & Alicke, M. (2021). My true self is better than yours: Comparative bias in true self judgments. *Personality and Social Psychology Bulletin, 47*(2), 216–231.
Ziano, I., & Villanova, D. (2023). More useful to you: Believing that others find the same objects more useful. *Journal of Experimental Social Psychology, 106*, Article 104460.
Zumthurm, S., & Stämpfli, A. (2024). A diet-related health prompt with the Swiss food pyramid as a nudge to reduce meat consumption. *Food Quality and Preference, 115*, Article 105105.