Behind the Veil of Choice

Testing the Effect of Decision Information and Decision Structure Interventions in a Self-Service Food Setting

THOMASSEN, ERIK
ØRITSLAND, CASPER

SUPERVISOR
Anders Emil Tobias Otterbring

University of Agder, 2024
School of Business and Law
Department of Strategy and Management

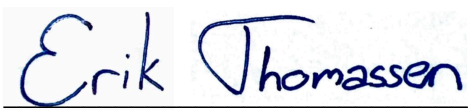
Preface

To SiA Mat og Drikke,

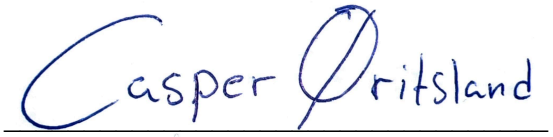
We would like to thank you for allowing us to use your facilities for our research project. Your assistance and cooperation is what allowed us to carry out our field study, which has greatly contributed to the success of our research. Thank you.

To Tobias,

Thank you! Working with you has been a pleasure, and your guidance and inspiration throughout the whole semester has been invaluable. Your excitement and investment in our research project has made the process truly fun, and we have greatly appreciated collaborating with you!



Erik Thomassen



Casper Øritsland

Abstract

Through a series of three high powered studies, this thesis explores and measures responses to two different nudging interventions related to price display and serving utensils in a self-service food environment. Study 1 ($N = 400$) reveals that people intuitively think displaying price in kilograms as opposed to hectograms will lead to a reduction in the amount of food purchased, and that using tongs instead of spoons has no effect on purchasing behavior. In contrast to these findings, a pre-registered field study conducted in study 2 ($N = 1965$) measuring *actual* purchasing behavior revealed that price display did not affect purchasing behavior, but that swapping spoons for tongs resulted in a statistically significant reduction of food purchased. In exploring the underlying mechanisms of these findings, study 3 ($N = 2447$) suggests that the relationship between serving utensils and amount of food served is mediated by both satisfaction and serving effort. Altogether, this paper highlights that different serving utensils can alter purchasing behavior, and that people's intuition regarding behavioral responses does not always correspond to *actual* behavior. Furthermore, study 3 extends the current literature by suggesting that serving effort alone cannot explain the change in amount of food served caused by swapping serving utensils as thought in previous research, and proposes satisfaction as an additional underlying mechanism.

Keywords: Nudge, choice architecture, pre-registration, field study, decision information, decision structure, effort, satisfaction

Table of contents

List of figures and tables

1.0 Introduction

People make suboptimal judgements and decisions all the time, and as a consequence incur massive costs to both themselves and the society as a whole. Encouraging better choices through redesigning the environment in which choices are made has therefore become a popular way to reduce these costs (Callaway et al., 2023, p. 1457). One way to do exactly this is through what is known as *"nudging"*. Nudging involves implementing small interventions in an environment, and while subtle and even often unnoticeable, nudges can have massive effects on changing people's behavior (Callaway et al., 2023, p. 1457). The concept of nudging has been adopted and used to promote better behavior in several different domains, such as health, environmental work, finance, and food (Mertens et al., 2022, p. 5).

Although the performance of a nudge depends on the context in which it is implemented, recent research shows that the domain of food is generally where nudging interventions have had the biggest impact (Mertens et al., 2022, p. 5). A lot of the interventions within the food domain have focused on nudging people towards healthier eating. With a varying degree of success, several creative nudges have been tried, such as reducing plate size (Kallbekken & Sælen, 2013; Kosīte et al., 2019; Libotte et al., 2014; Richardson et al., 2021), altering the position of healthy items in restaurant menus (Dayan & Bar-Hillel, 2011; Kim et al., 2019; Reinholdsson et al., 2023), adding nutritional labels to foods (Byrd & Almanza, 2021; Cioffi et al., 2015; Montagni et al., 2020), informing grocery store customers about other people's fruit and vegetable purchasing habits (Gonçalves et al., 2021; Huitink et al., 2020; Suleman et al., 2022), and even as bizarre as exposing people to slim human-like sculptures in an environment to promote weight loss (Brunner & Siegrist, 2012; Stämpfli et al., 2017).

While many interesting nudges have been studied throughout the years, questionable reporting practices and publication bias have distorted research findings and the true effects of a lot of interventions (Sanders et al., 2018, p. 157; Szaszi et al., 2018, p. 363). The research field is therefore screaming for more high quality research, and with more and more of public policy development being based on behavioral insight and the latest research within the field of nudging, the consequences of poor research involves both wasting valuable resources and damaging people's trust in the behavioral sciences (Sanders et al., 2018, p. 156). With this in mind, this research project aims to contribute to the field of nudging within the food domain by conducting high powered and rigorous research exploring how price display and different

serving utensils affect actual and self-reported behavior in a self-service food environment. Our research objective (RO) is therefore:

RO: To rigorously explore and measure responses to two different nudging interventions related to price display and serving utensils.

This is achieved through a series of three studies, where the first study aims to assess people's intuition regarding the two different interventions, followed by a larger pre-registered field study measuring *actual* behavioral responses to the nudges, while the third study attempts to explain some of the underlying mechanisms behind the findings from the field study.

2.0 Theoretical framework and existing literature

2.1 Decisions in traditional and behavioral economics

2.1.1 The economic man and bounded rationality

Traditional economic theory presupposes that people are rational, and for the most part act in the manner that they *should*. Even though traditional economists do not claim that all people act rationally at all times, they do insist that any deviations from perfect rationality are unsystematic and so small that they are negligible (Angner, 2021, pp. 1–2). Traditional economic theory therefore postulates the idea of the *"economic man"*, who in the course of being economic, also is assumed to be fully rational. This allows the economic man to maximize utility in all situations, and always reach the highest attainable point of his preference scale when making decisions (Simon, 1955, p. 99).

Behavioral scientists have however challenged the traditional economic concept of viewing humans as wholly rational actors (Ruggeri, 2022, p. 36). The american economist and 1978 winner of the Nobel Prize in economics, Herbert Simon (The Nobel Prize, n.d.), raised great doubts about the concept of the economic man, and whether that model provided a suitable foundation to form a theory of how people *do*, or *should* rationally behave, presented in his seminal 1955 paper *"A behavioral model of rational choice"* (Simon, 1955, p. 99). The paper aimed to construct a definition of *"rational choice"* more in line with "*actual"* decision processes (Simon, 1955, pp. 99, 114). Simon later coined the term *"bounded rationality"* to describe rational choices that take into account the human limitations of limited cognitive ability and computational capacity (Simon, 1990, p. 15). The human brain makes decisions

that are subject to the constraints imposed both by its own limitations, as well as limitations tied to the available information (Doucouliagos, 1994, p. 878).

Bounded rationality has since become a central theme within the field of behavioral economics, which is concerned with how decisions are *actually* reached (Simon, 1990, p. 15). An example to illustrate an everyday display of bounded rationality, is that *"[...] if it is raining outside but you have a date in ten minutes, you are likely to buy the most easily available umbrella you can find, ignoring the economic implications of it possibly being overpriced or of poor quality"* (Ruggeri, 2022, p. 38). Bounded rationality can therefore lead to people "*satisfice",* rather than *"optimize"*, where satisficing means that people set a target goal they want to achieve, and then choose the first option that satisfies the criteria of the goal, rather than continuing to evaluate all other available options to make the absolute best decision (Ruggeri, 2022, p. 29).

2.2 Changing behavior

2.2.1 Nudging

By building on the concepts from classical economic theory, the field of *behavioral economics* has evolved by supplementing the traditional economic assumptions with theories, evidence and ideas from the cognitive and social sciences, as well as psychology. The insights from this field of research have proved useful across many different domains, with one of the most prominent tools from the field being *"nudging"*, which can be used to alter people's behavior (Ruggeri, 2022, p. 36).

First introduced by Thaler and Sunstein in their book *"Nudge - Improving Decisions about Health, Wealth, and Happiness"* back in 2008, a nudge is defined as *"[...]any aspect of the choice architecture 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 avoid (Thaler & Sunstein, 2021, p. 8). One example of this can be how one could encourage healthier eating by placing food like fruit at eye level in cafeterias making the healthier food options more salient, rather than outright banning junk food to reach the same goal. Other more intrusive methods of changing behavior, such as bans, fines, taxes, or mandates can not be considered nudges (Thaler & Sunstein, 2021, p. 8).

2.2.2 Choice architecture

Those responsible for introducing and creating nudges are called *"choice architects"*, and a choice architect is the person or collection of individuals who designs the environment in which options are presented (Arno & Thomas, 2016, p. 2). They are responsible for *"[...]organizing the context in which people make decisions"* (Thaler et al., 2012, p. 428), and while this may seem easy, it oftentimes is not, as predicting the actual behavioral responses to different environmental design changes and nudging interventions can be difficult. There is no *"one size fits all"* for nudges, and this means that the best nudging strategies to use varies from case to case.

Engaging in the role as choice architects, several governmental bodies across the world have also established what they call *"nudge units"*. These are units that apply behavioral insights collected from research to the development of public policy and services (Kosters & Van der Heijden, 2015, p. 277). Some of the first examples of this was the *"Behavioral Insight Team"* established by the UK government back in 2010, shortly followed by the United States of America's nudge unit labeled the *"Nudge Squad"* by the US media (Kosters & Van der Heijden, 2015, p. 277). Since then, several other governments around the globe spanning the political landscape have established their own behavioral insights teams, with examples being New Zealand, Germany, France, Japan, India, Qatar, Saudi Arabia and many many more. The popularity of nudging units have also expanded passed governments, and institutions such as the World Bank, United Nations, the European Commission and even the World Health Organization have also begun applying behavioral insights to tackle and change unwanted behavior (Thaler & Sunstein, 2021, pp. 19–20).

2.2.3 Sludge

Nudging is however not a tool limited to governments or big institutions, and corporations, charities, religious organizations, teachers, managers, as well as parents also utilize the power of nudges to encourage and help people to make better decisions (Angner, 2021, p. 263). However, as with every tool, nudges can also be used for *"[...] less benevolent purposes"* (Thaler, 2018, p. 431), and to cause harm. When choice architects design environments in ways that inflicts friction and makes it *harder* for people to achieve an outcome that will make them better off, it is called a *"sludge"* (Thaler & Sunstein, 2021, p. 153). One example of this can be whenever a newspaper or a subscription-based service makes it easy to *subscribe*, but hard to *unsubscribe*. Sludges like this can have a hugely negative impact, and

have even been found to be a source of social inequality affecting the most vulnerable members of society the most (Sunstein, 2022, pp. 654, 665).

2.2.4 Libertarian paternalism

The core concept of nudging lies within a wider construct known as *"libertarian paternalism"*, which is the idea that you can encourage optimal behavior without explicitly mandating preferences or actions, but rather through subtle processes that support autonomy (Ruggeri, 2022, p. 302). The ethical implications behind the concept has however been a source of many debates. Critics of the nudging movement are adamant that nudging conflicts with moral values like autonomy, liberty, respect, and dignity (Schmidt & Engelen, 2020, p. 3). Others argue that nudges are manipulative, and that choice architects should not be responsible for defining what is *"good"* or in others' best interest (Mills, 2015, p. 495; White, 2013, p. XV; Wilkinson, 2013, p. 354).

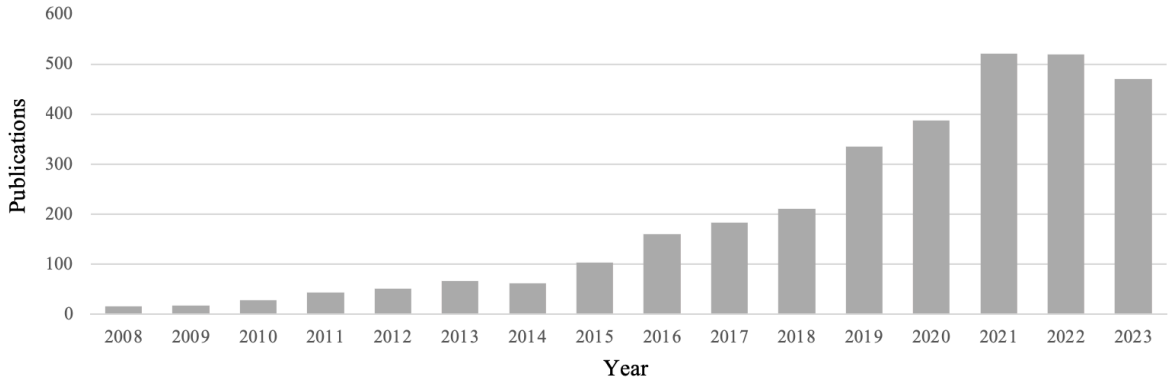
While nudging can be seen as an interesting approach to alter problematic behavior in the society, Goodwin (2012, pp. 85, 90) suggests that nudging is an ineffective tactic to use when tackling big societal issues such as climate change and public health. He argues that these issues should be dealt with through more deliberative processes, encouraging public and democratic engagement. A counterargument against this is however that policymakers have the luxury of choice, as they *"[...] do not have to choose between structural reform and nudging; [they] can do both"* (Schmidt & Engelen, 2020, p. 8).

Although several people have voiced their opposition to, or even condemned the use of nudges, many have also argued in favor of and praised libertarian paternalism as a movement. As nudging neither removes, nor changes options or the economic incentives behind choices, it can be argued that they actually respect freedom of choice (Schmidt & Engelen, 2020, p. 3), and that nudges are mostly cases of *"rational persuasion"* that gently push people towards making better choices while still respecting individual liberty (Hausman & Welch, 2010, pp. 128, 131, 136). This is further supported by the fact that most nudging interventions only target a small percentage of the population and are easily resistible (Schmidt & Engelen, 2020, p. 5).

2.3 Choice architecture framework

Since the concept of nudging was first introduced by Thaler and Sunstein (2008), the research field has seen a steady increase in interest and publications, as shown in Figure 1.

Figure 1
Steady increase in nudging publications since 2008



Note. Publications with the keyword "*nudge*" on Web of Science between 2008 and 2023 as of 24.05.2024 (Web of Science, n.d.).

2.3.1 Choice architecture categories and intervention techniques

All these scientific publications span across several different domains and have covered a lot of different nudges. These nudges all use different intervention techniques, and to better describe the different nudges, Münscher et al. (2016, p. 514) suggested a framework where intervention techniques could be categorized into overarching intervention categories.

This paper specifically focuses on two different nudging interventions, namely displaying prices in different measurement units (hectogram vs. kilogram) and altering the serving utensils in a food self-service environment. These two nudges fall within the intervention categories *"decision information"* and *"decision structure"*, which are displayed in Table 1. Previous research shows that these are the categories that have been the most effective in altering people's behavior, with an average effect size across the categories of Cohen's d = 0.34 and d = 0.54 respectively (Mertens et al., 2022, p. 5).

Table 1
An overview of intervention categories and intervention techniques

Intervention category	Intervention technique
Decision information: Increase the availability, comprehensibility, and/or personal relevance of information	*Translate information:* Adapt attributes to facilitate processing of already available information and/or shift decision maker's perspective *Make information visible:* Provide access to relevant information *Provide social reference point:* Provide social normative information to reduce situational ambiguity and behavioral uncertainty
Decision structure: Alter the utility of choice options through their arrangement in the decision environment or the format of decision making	*Change choice defaults:* Set no action default or prompt active choice to address behavioral inertia, loss aversion, and/or perceived endorsement *Change option-related effort:* Adjust physical or financial effort to remove friction from desirable choice option *Change range or composition of options*: Adapt categories or grouping of choice options to facilitate evaluation *Change option consequences:* Adapt social consequences or microincentives to address present bias, bias in probability weighting, and/or loss aversion

Note. This table is adapted from Mertens et al. (2022, p. 3).

The category of decision information is composed of interventions where the choice architect is concerned with altering how the information in an environment is presented (Münscher et al., 2016, p. 514). Techniques in this category include translating information by reframing or simplifying it, making information more visible, or providing social reference points (Münscher et al., 2016, p. 514). The decision structure category covers interventions that affect behavior through the arrangement and composition of alternatives, setting defaults, or alter option-related effort and consequences (Münscher et al., 2016, p. 516). Such interventions take advantage of humans inability to evaluate and compare the best outcome regarding different choice options, as humans instead base decisions on information that is directly available to them (Mertens et al., 2022, p. 2).

While these nudging techniques have been utilized to alter behavior across a variety of different choice environments, one domain in particular stands out as being especially susceptible to nudges. After analyzing more than 200 different studies covering various nudging interventions, Mertens et al. found that nudging tactics were the most effective when targeting *food-related behavior* (2022, pp. 1, 5).

2.4 Nudges within the food domain

2.4.1 Decision information interventions

Within the domain of food, many studies have been conducted covering nudges within the category of decision information. By utilizing the intervention technique of making information more visible, several researchers have for example attempted to nudge people towards choosing healthier food options by exploring how calorie labeling affect food choices (Godden et al., 2024; Jue et al., 2012; Rising & Bol, 2017; VanEpps et al., 2016; Wisdom et al., 2010). This kind of information nudge has however seen varied results, with some reporting that calorie labeling reduces calorie intake (Wisdom et al., 2010, p. 164), and other reporting small but inconsistent effects (Jue et al., 2012, p. 629), or that it only affects people with high self-control (Rising & Bol, 2017, p. 1032). Contrasting this, Godden et al. (2024, p. 1) found no evidence that calorie labels had any effect on nudging people towards healthier snack choices.

In addition to calorie labeling, some researchers have also tried manipulating the position of items on restaurant menus in order to facilitate healthier food choices. This has consistently proven to be effective, suggesting that such a visibility intervention can alter behavior in a predictable way (Dayan & Bar-Hillel, 2011, p. 339; Perez-Cueto, 2021, p. 2; Reinholdsson et al., 2023, p. 557).

Some have also utilized the intervention technique of information translation and tried to promote healthier eating by categorizing food based on nutrient richness with traffic light colors (Erdem et al., 2022; Godden et al., 2024; Seward et al., 2016; Thorndike et al., 2014, 2019). In contrast to calorie labeling, results from several studies on the effect of nutrient labels suggest that these interventions can promote healthier eating behavior (Erdem et al., 2022, p. 37; Godden et al., 2024, p. 1; Thorndike et al., 2014, p. 143, 2019, p. 1). As often in the nudging literature however, also this nudge has seen contradicting evidence and through a

comprehensive study involving approximately 2.6 million servings over the whole research period, Seward et al. (2016, p. 1088) did not find any evidence of traffic lights having an effect on healthier food choice, even when individuals stated that they used the traffic light function.

As for the last intervention technique within the decision information category, namely providing a social reference point, some studies have been conducted where researchers displayed informational materials in grocery stores that informed customers of other people's fruit and vegetable purchasing norms (Gonçalves et al., 2021; Huitink et al., 2020; Suleman et al., 2022). The results from these studies all consistently showed that this social reference nudge worked, suggesting that such a low cost nudge can change purchase behavior and increase the sales of healthy food items (Gonçalves et al., 2021, p. 8; Huitink et al., 2020, p. 5; Suleman et al., 2022, p. 4).

2.4.2 Decision structure interventions

Within the category of decision structure, one of the most prominent intervention techniques is the configuration of new defaults in choice environments, and within the food domain a lot of research has focused on whether changing the plate sizes at restaurants and canteens can reduce food consumption and food waste (Kallbekken & Sælen, 2013; Kosīte et al., 2019; Libotte et al., 2014; Richardson et al., 2021). Both Kalbekken & Sælen (2013, p. 325) and Richardson et al. (2021, p. 1) found significant indications that food consumption and waste would decrease by reducing the plate size at self-service food stations. However, other scholars have found contradicting evidence and propose that reducing plate size have no significant effect on consumption (Kosīte et al., 2019, p. 1; Libotte et al., 2014, p. 91).

In addition to plate size, other default nudges have also been studied such as how pre-portioning of salad affects consumption of vegetables (Friis et al., 2017, p. 1), having apple slices as the default side dish in a fast food restaurant instead of french fries (Wansink & Just, 2016, p. 1), and having whole wheat bread as the default in a sandwich, as opposed to other unhealthier bread choices (van Kleef et al., 2018, p. 179). All of the different default nudges in these studies were effective in producing healthier eating behavior among the participants.

Many have also explored nudging interventions related to altering effort through changing proximity and placement of different food options. In their book about nudging, Thaler and Sunstein described how a director of food services in charge of the cafeterias in hundreds of different schools got kids to eat healthier just by altering the way food was arranged and displayed in the school cafeterias (Thaler & Sunstein, 2021, pp. 1–3). Inspired by this, several other researchers have attempted to alter the way healthy and unhealthy food is arranged, and studied what effect this has on decision-making. Results from many of these studies show that placing healthier foods in a closer proximity and unhealthy food further away generally has a positive influence on healthier eating (Keller et al., 2015, p. 41; Rozin et al., 2011, p. 323; Van Gestel et al., 2018, p. 800). Others have however found that proximity alone does not matter, but rather the position of the healthier item relative to another snack (Knowles et al., 2019, p. 94).

As for the intervention technique of compiling or grouping options, some studies have explored the effect of clearly categorizing and separating healthy and less healthy food from each other (Hanks et al., 2012; Kingham et al., 2023). Hanks et al. found that arranging a cafeteria by separating healthy and unhealthy food into two different lines led to more healthy food consumption (2012, p. 370), while Kingham et al. separated healthy and unhealthy foods in different groups in an online restaurant menu, but did not find this to change people's behavior (2023, p. 4).

Few studies utilized the technique of changing option consequences specifically related to food, however relating to health one study explored how different incentive based approaches could promote weight loss, finding that a group of people incentivized by lottery tickets lost significantly more weight compared to a control group (Volpp et al., 2008, p. 2631). Another study used food as an incentive to increase immunization rates in rural India by offering raw lentils for completed immunization, finding that small incentives had a large positive influence on uptake of immunisation services (Banerjee et al., 2010, p. 1).

As the previous literature shows, nudging interventions comes in many forms, have a varying degree of success, and it is not always easy to predict beforehand whether an intervention will lead to the desired behavioral change or not. Nudges within the food domain does however stand out with an average effect size of Cohen's $d = 0.65$, in contrast to the average of $d = 0.43$ across all domains, although the effectiveness of interventions vary greatly depending on

what nudging technique is being used and the context in which it is implemented (Mertens et al., 2022, pp. 1, 5).

Going forward, this paper will focus on two different intervention techniques within the decision information and decision structure categories, specifically by utilizing nudges related to effort and information translation. Across three different studies, we will test what effect the manipulation of price display and serving utensils has on purchasing behavior.

Altering serving utensils is related to the technique of changing option related effort as it adjusts the physical effort related to self-service of food. Price displayed in different measurement units (hectogram vs. kilogram) is related to the technique of translating information as it adapts the attributes of the way price is presented to shift decision-makers perspective regarding the price of food.

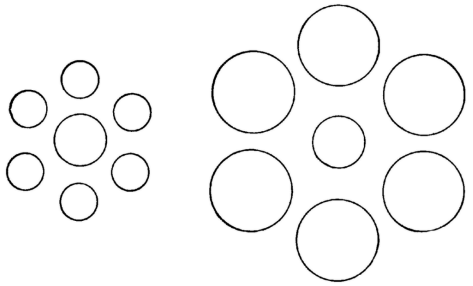
We start by studying people's intuition regarding these two nudges in study 1.

3.0 Study 1 - An intuition study

3.1 Intuition, numerosity and effort

The main objective of study 1 is to gather insight into people's *intuition* regarding how price displayed using different measurement units and how different serving utensils affect purchasing behavior.

Figure 2
"The Titchener Circles"



Note. This figure is from the book *"Experimental Psychology: A Manual of Laboratory Practice"* (Titchener, 1901, p. 169).

The idea of testing people's intuition and beliefs regarding perception of units and sizes is not new, and *"it has long been known that estimates of perceptual magnitude are never made in isolation, but [are] rather made relative to all of the other stimuli which form its context"* (Coren & Miller, 1974, p. 355). One example of this is the *"Titchener Circles"* (Titchener, 1901, p. 169) , which illustrate how we perceive and estimate size based on the context (see Figure 2). When looking at the circles, we see the left circle surrounded by smaller circles as

larger than the right circle surrounded by larger circles, when they in fact are the exact same size.

When it comes to our perceptions of numbers with differing measurement units (like hectograms and kilograms), humans have a tendency to focus more on the total number, rather than the measurement unit in which numbers are presented in. This builds on *"the numerosity heuristic"*, which is a tendency humans have of evaluating higher numbers as more expensive and less attractive than lower numbers, and people tend to overlook the unit in which the numbers are presented (Pandelaere et al., 2011, p. 308). For example, we are likely to overspend when traveling to a country where the currency is a fraction, and underspend when it is a multiple of one's home currency (Raghubir & Srivastava, 2002, p. 335).

As for the second nudge explored in study 1 relating to serving effort, while a few field studies have been conducted with regards to changing the effort of a serving utensil in a self-service food environment (Kanchanachitra et al., 2020; Rozin et al., 2011; Wansink et al., 2006), little research has been done on what people intuitively think will happen when utensils change.

In study 1 we will therefore examine how people intuitively believe measurement units and different serving utensils affect purchasing behavior.

3.2 Study design

Study 1 uses a *single-factor between-subject design*. A single-factor design is used when you want to investigate what effect *one* factor, or a single independent variable, has on a dependent variable (Selvamuthu & Das, 2018, p. 223), while between-subject refers to each participant only being exposed to one treatment (Charness et al., 2012, p. 1). We collect quantitative data for study 1 through a short survey, and in a quantitative study, *"[...] numbers are used directly to represent the characteristics of something"* (Hair et al., 2007, pp. 151–152). Because quantitative data takes the form of numbers, it lends itself ideal to statistical analysis (Hair et al., 2007, p. 152), and the data from this study will be analyzed through two independent samples *t*-tests using *"IBM SPSS Statistics"*, which is a powerful statistical analysis software (IBM, n.d.).

3.2.1 Power analysis

A key element of good research practice is making sure to have adequate statistical power, and this is important as studies with high power increase the likelihood of being able to document both replicable and true effects (Sommet et al., 2023, p. 1). Statistical power *"[...] is the probability that statistical significance will be indicated if it is present"* (Bougie & Sekaran, 2020, p. 280). A statistical power analysis looks at the mathematical relationship between four different variables: the power, the alpha (α), the effect size, and the sample size (*N*). When three of these variables are fixed, the fourth can be calculated (Cohen, 1992, p. 98).

The alpha, or significance level, indicates the maximum allowed limit of type I errors, or false positives (Kang, 2021, p. 2). In most research fields, the standard cutoff has been $\alpha = 0.05$ (Miller & Ulrich, 2019, p. 1). Effect size is a standardized and objective way to measure the size or magnitude of an observed effect (Field, 2013, p. 79). Effect sizes can differ between studies, and the choice of effect size can vary based on the study design, method of measuring the outcome, or the statistical methods used in a study (Kang, 2021, p. 2). Lastly, there is the sample size. Increasing the sample size at a given alpha level can produce more power, which in turn can lead to a higher probability of finding what the researchers are looking for (Bougie & Sekaran, 2020, p. 280).

One of the most useful forms of power analysis is to calculate the sample size needed in a study, and this is typically used for research planning (Cohen, 1992, p. 98). Establishing an appropriate sample size is very important for both analyzing, presenting and drawing conclusions in a research project (In et al., 2020, p. 114). It is generally recommended to establish the power, effect size and alpha first, and then use this information to establish an appropriate sample size (Bougie & Sekaran, 2020, p. 280). A common measurement for effect size is Cohen's *d* (Field, 2013, p. 79).

Typical effect sizes within the marketing literature equate to $d = 0.49$ (Eisend, 2015, p. 23; Lin, n.d.), and within the field of psychology $d = 0.41$ is normal as a measure of a medium effect size (Funder & Ozer, 2019, p. 166; Lin, n.d.). A recent meta-analysis also shows that the average effect size for research tied to behavior/intention is $d = 0.45$ (Eisend et al., 2024, p. 184; Lin, n.d.). Building on these standards, we calculated that a total sample size of $n = 100$ was needed for each group in study 1 (in which we have four in total) to detect a mean

difference corresponding to $d = 0.40$ with a power of 80%, given the standard alpha level of 0.05 and two-tailed tests. The calculations were done using *"G*Power",* a software tool that can calculate sample size and power for various statistical methods (Kang, 2021, p. 1).

3.2.2 Sampling technique and participants

When recruiting participants to a survey, one can use either *random* or *non-random* sampling. Random sampling involves recruiting a random selection of members from a population where each member has an independent and equal probability of being selected (Sedgwick, 2013, p. 1). When collecting data for study 1, we used an approach called *convenience sampling*, by collecting data at the university campus. This sampling method falls under the non-random category, and involves selecting participants based on convenience and accessibility, rather than complete randomness (Sedgwick, 2013, p. 1). With no specific inclusion or exclusion criteria, the participants in study 1 consisted of students physically present at the university campus in which the research project took place during the three days of data collection, where a total of 400 students participated in the study (61.25% female; $M_{age} = 23.74$).

3.2.3 Survey

The study is divided into two parts, A and B. Study 1A tested people's intuition regarding price display (see Appendix A), and study 1B was related to serving utensils (see Appendix B). The surveys contained two demographic questions covering age and gender, followed by one question relating to either price display or serving utensils, measured on a seven-point Likert scale.

To avoid the framing of the questions having an influence on the way people responded, we produced two versions of each main question, where we presented the question from opposite angles. This gave us a total of four questions, split between the two predictions. Study 1A ($n = 200$) asked the question: *"Do you think people will buy more or less food in the buffet in the canteen if the price is shown in kilograms instead of hectograms"* (1 = Less when kilograms, 7 = More when kilograms) or *"Do you think people will buy more or less food in the buffet in the canteen if the price is shown in hectograms instead of kilograms"* (1 = Less when hectogram, 7 = More when hectograms). Study 1B ($n = 200$) asked the question: *"Do you think people will buy more or less food in the buffet in the canteen if the serving utensil used are spoons instead of tongs?"* (1 = Less with spoon, 7 = More with spoon) or *"Do you think*

people will buy more or less food in the buffet in the canteen if the serving utensil used are tongs instead of spoons?" (1 = Less with tongs, 7 = More with tongs). We also provided illustrations of price signs as well as a spoon and a tong to ensure people clearly understood what we were asking.

3.2.4 Procedure

The survey for the intuition study was printed out on paper and handed out physically, and we distributed the survey to students we met around the university campus over a total of 3 days. All students were informed that it was entirely voluntary to participate in the study, and that all answers were completely anonymous. With this physical pencil and paper approach, we received an approximated response rate of 99%. To make the survey distribution as random as possible, we shuffled the papers before distributing them so that two of the same versions of the survey would not be handed out twice in a row. This decreased the chance that two people sitting beside each other could get the same version of the survey, and thereby copy each other's answers. All respondents were also told that they could *not* discuss the questions before answering the survey.

3.3 Results and discussion

As we made no assumptions regarding the direction of people's intuition in study 1A and B, two tailed independent samples *t*-tests were used to analyze people's beliefs regarding the amount of food bought when presented with different price units and serving utensils.

Findings from study 1A suggest that people believe that consumers will buy less food in a buffet when price is presented in kilograms ($M = 3.17$, $SD = 1.53$) compared to hectograms ($M = 5.15$, $SD = 1.59$), and that there is a significant difference between the two conditions $t(198) = -9.00$, $p < .001$, $d = -1.27$.

The results from study 1B also produces interesting insights when it comes to people's belief on the impact of using different serving utensils. Through an independent samples *t*-test, our findings show that people think there is no difference in the amount of food bought in a buffet if the serving utensil used is tongs ($M = 4.35$, $SD = 1.37$) as opposed to spoons ($M = 4.36$, $SD = 1.24$). There was no statistically significant difference between the groups $t(198) = -0.05$, $p = .957$, $d = -0.01$.

Study 1 suggests that people intuitively believe that customers will buy substantially less food when price is presented in kilograms (vs. hectograms), and that people do not believe that there is a difference between using spoons or tongs with regards to how much food one would buy. With a vast amount of previous research showing that there is often a gap between what people say they will do and how they *actually* behave (e.g. Allan et al., 2011, p. 635; Grimmer & Miles, 2017, p. 2; Sainsbury et al., 2013, p. 52; Sheeran & Webb, 2016, p. 511; Sniehotta et al., 2005, p. 143), we want to further explore how these interventions work in real life conditions through a quasi-experimental field study in a university canteen.

4.0 Study 2 - A quasi-experimental field study in a university canteen

While studying *"common sense"* and intuition can be useful, these beliefs can have a large and sometimes unfortunate influence on scientific concepts and theories (Kelley, 1992, p. 4). It is therefore important to also test the concepts explored in study 1 in the real world, by utilizing a different study design to get a better understanding on how these nudges work. The purpose of study 2 is therefore to examine how price display and changes in serving utensils influence consumer's *actual* purchasing behavior through a field study in a real canteen environment. The field study was conducted in a Norwegian university canteen with 500-750 transactions per day, across four consecutive Fridays with the food kept constant across all days.

4.1 Hypothesis development

4.1.1 Translation

Because changes in perspective can lead to changes in preferences, framing or transforming information in a given context can have an effect on the way information is perceived. Several scholars have studied the impact of transforming numerical information (Bagchi & Davis, 2016; Fecher et al., 2019; Liu et al., 2023; Pandelaere et al., 2011; Raghubir & Srivastava, 2002), and in one study where respondents were asked to rate objects as expensive on a scale from 1-7 (1 = Strongly disagree, 7 = Strongly agree), Fecher et al. (2019, pp. 132–133) found that the same products presented in price per kilograms, were viewed as significantly more expensive than when presented in hectograms. Similar findings are found in a study by Liu et al. (2023, p. 411) where fruits and vegetables were viewed as more expensive when price was presented in kilograms, compared to hectograms. The

numerosity theory also suggests that prices presented in large measurement units, hence higher numerical value, will result in a perception of higher prices (Bagchi & Davis, 2016, pp. 89–90). Based on this, we therefore hypothesize that:

H1: Displaying price in NOK/kg (vs. NOK/hg) will reduce the amount of food purchased measured in weight.

4.1.2 Effort

The body of research regarding how different serving utensils impact food served is less extensive, however a 2006 study looked at how increasing the size of an ice cream scoop at an ice cream social affected the amount of ice cream a group of nutrition experts served themselves. This study found that increasing the size of the ice cream scoop from 2 oz to 3 oz increased the amount of ice cream people served themselves by 14.5%, however this result was not statistically significant[1] (Wansink et al., 2006, pp. 240–242).

Another study found that swapping spoons with tongs at a salad bar over an 11 week period reduced average food intake by 16.5% (Rozin et al., 2011, p. 328). Similar results were found when Kanchanachitra et al. (2020) attempted to reduce intake of fish sauce at a university canteen by placing the sauce in a bowl together with a *"special"* spoon which had a hole in it. This increased the serving effort and decreased the intake of fish sauce among the students (Kanchanachitra et al., 2020, p. 13). Rozin et al. (2011, p. 330) suggest that the most plausible mechanism behind this behavior is *"Tolman's Law of Least Effort"*, and refers to Tolman (1932) who argues that choices tend to be made with a *"[...] minimum expenditure of physical energy"* (Tolman, 1932, p. 448), and that the observed behavior change is due to people minimizing their expended energy when obtaining food (Rozin et al., 2011, p. 330). Conducting a similar study as Rozin et al. (2011) of spoons vs. tongs, we hypothesize that:

H2: Using tongs (vs. spoons) as the serving utensil in the self-service canteen will reduce the amount of food purchased measured in weight.

[1] 5.77 oz vs. 5.04 oz of ice cream served, $F(1, 80) = 2.70$, $p = 0.10$ (Wansink et al., 2006, p. 240).

4.1.3 Combining choice architecture techniques

Some scholars have also tested two or more nudges at the same time, producing results suggesting nudges can be effective when combined (Brandon et al., 2019, p. 5293; Dannenberg & Weingärtner, 2023, p. 6; Hanks et al., 2013, p. 868; Thorndike, 2020, p. 1). For example, in one study Brandon et al. (2019, p. 5293) tested two nudging strategies both in isolation, and combined, and found that when isolated the two nudges reduced consumption of energy use in a normal household with 2 to 4% and 7% when combined. This could indicate an additive function by combining multiple nudges at the same time, and we therefore hypothesize that:

H3: The display of price in NOK/kg (vs. NOK/hg) in combination with using tongs (vs. spoons) as the serving utensil in the self-service canteen will further enhance the reduction of food purchased measured in weight compared to all other conditions.

All three hypotheses were pre-registered before beginning any data collection (see https://aspredicted.org/C9W_CM4 or Appendix C).

4.2 Field study

Study 2 is a quasi-experimental field study conducted in a university canteen where different nudging interventions are introduced in the canteen environment. The effects of the nudges are then measured through analyzing the canteen's transactional data. In a quantitative research setting, the researchers should be as detached from the research as possible, using methods that minimize their involvement to maximize objectivity (Sukamolson, 2007, p. 5). The transactional data is downloaded in the days subsequent to the study taking place, and is therefore completely objective, anonymous and detached from the researchers.

The main strength of the transactional data used in this study, is that it measures the customers *actual* behavior, in contrast to how a survey only can measure what people *say* they would do. Having access to objective data that measures actual behavior has been crucial in this study, as studies based on self-reporting often suffer from bias, with participants responding in the way that makes them look as good as possible. People tend to under-report *"inappropriate"* behavior and over-report behavior that is seen as *"appropriate"* (Donaldson & Grant-Vallone, 2002, p. 247), and using quantitative

transactional data therefore allows us to see how each intervention *truly* affects purchasing behavior without being affected by this reporting bias.

4.2.1 Quasi-experimental design

The canteen in which the field study takes place offers a wide variety of food and beverages, including a self-service buffet and a salad bar. It is a hotspot among students, faculty members and other guests at the university, serving as a place to eat, socialize, and work. The sample in this study is therefore all users of the canteen on the days the research project takes place. As a result of the high activity, several hundred transactions are completed each day, making it an attractive arena for testing different nudging interventions.

The most practical design to utilize in the environment in which this study takes place was a quasi-experimental design, where individual subjects are not randomly assigned to groups (Creswell, 1994, p. 130). As to allow the canteen to operate as normal during the field study, it was not practical to randomize treatments on an individual level. We did however randomize on a group level, by randomly assigning each of the four conditions to different weeks using random.org, which is a random number generator that uses atmospheric noise to create true randomness (random.org, n.d.).

4.2.2 Power analysis

A systematic scoping review from 2018 that set out to analyze when and why nudges actually work, found that amongst the 156 studies included in their review, only 7% had applied a power analysis (Szaszi et al., 2018, p. 355). This is unfortunate for the field of research, as power analysis and sample size calculation is essential to ensure that the research is both replicable and rigorous (Baldwin, 2017, p. 13), both hallmarks of scientific research (Bougie & Sekaran, 2020, p. 17).

In the planning phases of study 2, we therefore conducted a power analysis to determine the appropriate sample size for the field study. Based on historical sales data from the canteen in which the study takes place, we expected a total sample size of 2000-3000 transactions during the full course of the study, or approximately 500-750 transactions per condition. Our sample size calculations using G*Power shows that $N = 620$ transactions were needed to detect small effect sizes of $d = 0.20$ at a power level of 80%. This assumes independent samples *t*-tests, the standard alpha level of 0.05, and one-tailed predictions.

4.2.3 Conditions

The study took place across four weeks, and included four different conditions. Each of the conditions were run on separate Fridays, and the food was kept constant across all days.

Condition 1) Using spoons as the serving utensil and displaying the price in NOK/hg.
Condition 2) Using tongs as the serving utensil and displaying the price in NOK/hg.
Condition 3) Using spoons as the serving utensil and displaying the price in NOK/kg.
Condition 4) Using tongs as the serving utensil, and displaying the price in NOK/kg.

Price was displayed both on wall mounted TV-screens, and using signs (see Figure 3 and 4) at the food counter of the self-serving buffet and salad bar. In total, the price was shown on four individual TV-screens and two signs across the canteen. To ensure that the conditions remained the same when switching from price highlighted in NOK/hg to NOK/kg, every screen and sign was updated between the different conditions.

Figure 3
Price sign used in condition 1 and 2

Figure 4
Price sign used in condition 3 and 4

Every ingredient in both the salad bar and the self-service buffet, was equipped with either spoons (Figure 5) or tongs (Figure 6) during the respective conditions. 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. A few select ingredients such as jalapeños and feta cheese however used smaller spoons and tongs for convenience. The only exceptions where deviations from the four conditions were made, were always using tongs for the tortillas, and spoons when serving sour cream and other semi-liquid foods. The exceptions stayed constant across all four conditions.

Figure 5
Spoons used in the field study

Figure 6
Tongs used in the field study

4.2.4 Dealing with several food orders

In the transaction data, there were some receipts that contained more than one food order. When receipts contain more than one food order, we treated each food order as individual orders, and not as one aggregated. This is because we assumed that each weighted food option corresponded 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 corresponded to food that was purchased to be consumed by more than one individual.

4.2.5 Pre-registration

The field study was also pre-registered prior to starting any data collection. As governments are now using insights from research on decisions and behavior in the development of new public policies (Sanders et al., 2018, p. 144), it is more important than ever that the research has a high standard. An issue within the field of nudging has however been poor reporting practices, and this is problematic as it may hinder the development and accumulation of evidence. In addition to this, biases in publication and reporting may distort the true effects of different interventions (Szaszi et al., 2018, pp. 355, 363). Some have even gone as far as claiming that when adjusting for publication bias, *"[...] no evidence for the effectiveness of nudges remains"* (Maier et al., 2022, p. 1).

One measure that has gained traction as one of the most promising solutions to improving the replicability of scientific effects, is the *pre-registration* of research projects (van den Akker et al., 2023, p. 1). Pre-registration allows researchers to describe the details of their project in a public registry before starting the data collection. This serves several purposes, such as allowing collegial input on the project prior to starting the study, as well as allowing journal editors to accept papers conditionally, regardless of the results of the research project (American Psychological Association, 2021).

Pre-registration is however uncommon in nudging-studies, and Szaszi et al. (2018, p. 355) found in their scoping review based on 156 empirical studies about nudging that *none* were pre-registered. Furthermore, questionable reliability of research findings has negative effects for both theory and practice, for example as public policy development is based on intervention ideas from lab studies and field experiments (Sanders et al., 2018, p. 156).
To ensure our study follows the high standards required in scientific research, and to make sure our findings hold the quality needed to contribute to the research field, we pre-registered our study at *"https://aspredicted.org/"* during the planning phases. Researchers in the field must all work together on topics and measures such as pre-registration, data sharing, replication consortia, identification of statistical irregularities and accepting that some trials may discover one-off findings, as these are all steps in the right direction to solving issues of replication and publication bias that exist in the field of nudging today (Sanders et al., 2018, p. 157).

4.2.6 Reliability and validity

At the very core of scientific research is the pursuit of truth and limitation of error, and reliability and validity are ways of demonstrating and communicating the trustworthiness of the findings and the rigor of the research process (Roberts et al., 2006, p. 41). External validity is related to the generalizability of the results from a study to other settings, and internal validity refers to how confident one can be that it was the manipulation of the independent variable that caused the change in the dependent variable. Studies conducted in the field tend to have a higher external validity, with the trade-off of having a lower internal validity compared to for example lab experiments (Bougie & Sekaran, 2020, pp. 170–171). Reliability concerns the consistency of observations, and whether different observers will be able to attain the same results (Bougie & Sekaran, 2020, p. 139). Study 2 has the benefit of measuring completely objective transactional data from a large sample, and looks at the canteen customer's *actual* purchasing behavior, which leaves no room for creative interpretation from the researchers, and thus increasing the reliability.

As the field study also takes place in a noisy real world environment, internal validity can be somewhat compromised with many factors being hard to control for. We did however design the study with measures in mind to help increase the internal validity. These measures were having a large sample size calculated beforehand, the study taking place over several different weeks, and keeping the food constant across all days.

4.2.7 Ethical considerations

Collecting informed consent from participants involved in research projects is generally considered to be a central part of ethical research practice within the social sciences (Wiles et al., 2007, p. 1), however there are some conditions that that can render the requirement of informed consent truly impractical (Laurijssen et al., 2022, p. 545). One of these conditions is when *"[...] obtaining informed consent leads to invalid study outcomes"* (Laurijssen et al., 2022, p. 545). The true behavioral responses to the nudging interventions in this field study could only be measured by the participants being unaware of the small changes in the canteen environment. Informed consent was not collected in this study, because it was truly impractical and would invalidate the study outcomes. However, as the interventions introduced in the canteen were non-intrusive, almost unnoticeable, and did not in any way hinder people from serving themselves as much food as they wanted, we deemed the decision of not collecting informed consent in this particular case to be ethically justifiable.

In addition to this, the transactional data from the field study was completely anonymous, and contained no personal or sensitive information such as credit card information or banking references. With the transactional data being downloaded subsequent to the field study taking place, it was also completely impossible to tie any transactions to individual people[2].

4.2.8 Parametric and non-parametric tests

The transactional data is analyzed using independent samples *t*-tests, which is a type of *"parametric test"*. When using parametric tests, the data must meet a set of predefined assumptions (Field & Hole, 2003, p. 234; Zikmund et al., 2010, p. 517). According to Field (2013, p. 165), there are four assumptions that must be met in order to perform parametric tests, namely the assumption of additivity and linearity, normality, homogeneity of variance, and independence. If the data being analyzed does not meet these assumptions, one could use non-parametric tests instead. These are tests that are used when *" [...] the numbers do not conform to a known distribution"* (Zikmund et al., 2010, p. 517). The non-parametric version of an independent *t*-test is the Mann-Whitney test (Field & Hole, 2003, p. 235), and will be used depending on whether or not the data fulfills all four assumptions.

[2]Both the data protection officer at the University of Agder and Sikt confirmed that the project did not have to be registered and assessed by Sikt, as all data collected was completely anonymous.

4.3 Results and discussion

As pre-registered, all hypotheses in study 2 are tested through a series of three independent samples *t*-tests in SPSS, with the comparisons that were specified in the hypotheses. Given the directional predictions in the hypotheses, we will use one-tailed tests. Furthermore, the means and standard deviations are presented in grams of food purchased.

4.3.1 Pre-registered hypotheses

The first hypothesis predicts that people will buy less food measured in weight if price is presented in kilograms instead of hectograms. An independent samples *t*-test shows that there is no statistically significant effect $t(1963) = 0.61$, $p = .271$, $d = 0.03$ on food purchased measured in weight when price is presented in kilograms ($n = 965$, $M = 447.18$, $SD = 143.28$) compared to hectograms ($n = 1000$, $M = 443.25$, $SD = 142.47$). In the context of this study, price display had no effect on the amount of food purchased. This finding does not support hypothesis 1.

The second hypothesis predicted a decrease in food purchased measured in weight when using tongs as serving utensil instead of spoons at the self-service canteen. An independent samples *t*-test show that there is a statistically significant effect $t(1963) = -2.08$, $p = .019$, on the amount of food purchased in weight when using tongs ($n = 961$, $M = 438.35$, $SD = 140.65$) as serving utensil compared to spoons ($n = 1004$, $M = 451.72$, $SD = 144.69$). Although statistically significant, Cohen's $d = 0.09$ indicates that the effect was small. The finding does however support hypothesis 2.

The third hypothesis predicted that displaying both price in kilograms and using tongs as serving utensil in combination would further decrease the amount of food purchased measured in weight compared to all other conditions. An independent samples *t*-test show no statistically significant effect $t(1963) = -0.34$, $p = .369$, $d = -0.02$ when combining tongs and displaying price in kilograms ($n = 441$, $M = 443.17$, $SD = 145.39$) compared to the other conditions ($n = 1524$, $M = 445.76$, $SD = 142.15$). This finding does not support hypothesis 3.

Our data did not meet the criteria of normality, and we therefore also conducted the non-parametric version of an independent t-test, namely the Mann-Whitney test, however this

did not change the results in any of the hypotheses[3]. As pre-registered, no outliers were excluded in the main analyses, but we also ran tests where we excluded outliers scoring beyond plus or minus 3.29 standard deviations away from the condition-specific means regarding the weight of food purchased (n = 24), but neither this yielded any different results.

4.3.2 Exploratory analysis

In addition to testing our three pre-registered hypotheses, we also conducted a few exploratory analyses. Firstly, we checked if people bought more beverages when the price was displayed in NOK/kg vs. NOK/hg, but found no change in purchasing behavior (Appendix D). A week after the field study ended, the canteen also raised the prices in the self-service buffet and the salad bar from 16.90 NOK/hg to 17.90 NOK/hg. This was beyond our control and knowledge, and therefore not part of our pre-registration. We did however want to test if the increased prices had an effect on the amount of food purchased measured in weight through an exploratory analysis, and our analysis showed that the increase in price had no effect on the amount of food purchased at all[4].

We also looked at the transactional data before and after the field study, when no interventions were implemented, and performed a planned contrast test to check the robustness regarding the findings from hypothesis 2, indicating that people buy less food when serving themselves with tongs as opposed to spoons. When comparing the 16.90 NOK/hg and 17.90 NOK/hg price conditions to the tong-conditions, the contrast test showed that there was a statistically significant decrease in amount of food bought measured in weight between the tongs-conditions and the two other price conditions, further strengthening the robustness of the findings from hypothesis 2[5].

---

[3] H1 (z = -.713, p = .238), H2 (z = -2.363, p = .009), H3 (z = -.367, p = .357).

[4] An independent samples t-test comparing 16.90 NOK/hg (n = 887, M = 449.15, SD = 143.23) and 17.90 NOK/hg (n = 1156, M = 449.32, SD = 153.77) price conditions with no interventions showed no significant difference between the two groups $t(2040)$ = -0.03, p = .490, d = 0.00.

[5] When comparing the 16.90 NOK/hg and 17.90 NOK/hg price conditions (n = 2043, M = 449.25, SD = 149.25) to the tong-conditions (n = 961, M = 438.35, SD = 140.65), a planned contrast test showed that there was a statistically significant decrease in food bought measured in weight between the tong-conditions and the two other price conditions $t(3001)$ = 1.89, p = .029, d = 0.07 with a one tailed prediction.

4.3.3 Discussion

The results from the field study, revealed no evidence to support our hypothesis that displaying price per kilograms reduces the amount of food measured in weight. A lot of the previous research regarding measurement units have relied on study designs that involve surveys and hypothetical scenarios (e.g. Bagchi & Davis, 2012; Bagchi & Li, 2011; Burson et al., 2009; Fecher et al., 2019; Gourville, 1998; Kwortnik et al., 2006; Liu et al., 2023; Pandelaere et al., 2011; Raghubir & Srivastava, 2002; Shirai, 2017; Siddiqui et al., 2018; Wertenbroch et al., 2007; Yao & Oppewal, 2016a, 2016b), as opposed to testing predictions in real life settings. While it is well documented through previous research that changing the measurement units can alter behavior, it may be the case that noisy real world environments could reduce the previously observed effect in some contexts, or as in our case, even nullify it completely.

One explanation may be that since the food being sold in the self-service part of the canteen on the days the field study took place was taco, customers attention to the measurement unit was lowered, as taco is a tremendously popular dish among Norwegians where people have *very* specific habits regarding both the order and amount in which they serve themselves. Another possible explanation for the null-findings, may be that in the environment in which the study took place, there were few other alternatives to buy food leading people to buy regardless of how expensive they perceived the price to be. This thought was further strengthened when we saw that a 6% price increase from 16.90 NOK/hg to 17.90 NOK/hg did not change the amount of food purchased in weight, also indicating a low sensitivity and attention to price.

Furthermore, the findings from our field study support hypothesis 2, and is also in line with previous findings indicating that increasing serving effort also decreases the amount of food people buy or consume (Kanchanachitra et al., 2020; Rozin et al., 2011; Wansink et al., 2006). Our field study did however not find this effect to be particularly big. Measured in weight, Rozin et al. (2011, p. 328) saw an average reduction of 16.5% when people served themselves with tongs as opposed to spoons in a salad bar. In our field study, taking place in a canteen serving taco consisting of several different ingredients ranging from ground beef to tortilla chips, as well as a salad bar, we only saw an average reduction of food bought of 3%.

While the observed effect of the intervention in our study is noticeably smaller than the one found by Rozin et al. (2011), it is worth to note that the spoons used in their study were three times bigger than those used in our. The spoons used in their study could hold a volume of 120 ml, compared to the spoons we used that could hold 40 ml. Their tongs were however similar in size to the ones we used, with theirs being 7.62 cm long and 2.54 cm wide, and ours being 7.50 cm long and 3.50 cm wide. This may help explain how both studies found the same effect, namely that increasing serving effort slightly led to a reduction of food bought, but also why their study saw a much bigger difference between the two utensils. Furthermore, although a 3% decrease might seem small, it is not uncommon for nudging interventions outside of academia to have an impact close to this, with an average intervention percentage point impact of 1.4 (DellaVigna & Linos, 2022, p. 81).

As price display proved to have no effect in the context of this field study, the results regarding the third hypothesis were of no surprise as it was based on the notion that nudging interventions work well in tandem as suggested by previous findings (Brandon et al., 2019; Dannenberg & Weingärtner, 2023; Hanks et al., 2013; Thorndike, 2020).

5.0 Study 3 - Examining mediation effects

5.1 Effort and satisfaction

The results from study 2 indicate that people buy less food when serving themselves with tongs as opposed to spoons. This effect has also been found in previous research by Rozin et al. (2011, p. 328), who suggested that the most plausible explanation for the observed behavioral change is due to people trying to minimize the energy expended when obtaining food following Tolman's Law of Least Effort (Tolman, 1932, p. 448), and stating that the simplest hypothesis to explain the results is that people find it more effortful to extract food with tongs as opposed to spoons (Rozin et al., 2011, p. 330). Wanting to further explore the underlying mechanisms of the change in purchasing behavior observed in study 2 and building on the arguments presented here, we hypothesize that:

H4A: The effect of serving utensil on food served is mediated by serving effort. Specifically, spoons (vs. tongs) reduces serving effort and reduced serving effort increases the amount of food served.

Moreover, studies measuring the effect of satisfaction on purchasing behavior in the context of food also show that several factors such as perceived quality, visual cues, service, and atmosphere can impact food consumption and spending behaviors (Chow et al., 2007, p. 698; Ha & Jang, 2010, p. 520; Jang & Namkung, 2009, p. 451; Sulek & Hensley, 2004, p. 235; Wadhera & Capaldi-Phillips, 2014, p. 132). In addition to this, several scholars have studied the relationship between spending and satisfaction, and previous research shows that there is a positive relationship between the two, meaning that when satisfaction increases, so does spending (Anderson & Sullivan, 1993, p. 141; Fornell et al., 2010, p. 28; Kamakura et al., 2002, p. 294; Yeung et al., 2013, p. 406; Yeung & Ennew, 2001, p. 114).

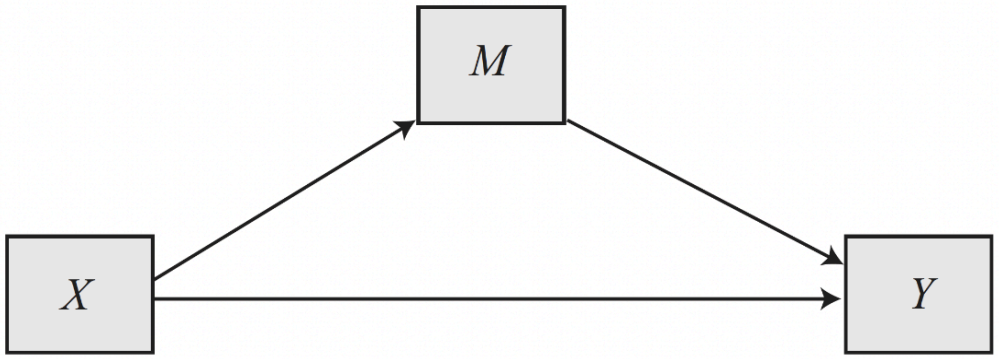
Few studies have however looked at how satisfaction relates to amount of food served by customers in self-service settings. It makes sense logically that spending and consumption should be closely related to amount of food served, and as satisfaction is closely related to spending and consumption, this could also be the case for food served. Based on this, and in an attempt to further expand the literature and explore more mechanisms that can explain the behavioral change observed in the field study relating to amount of food bought we hypothesize that:

H4B: The effect of serving utensil on food served is mediated by satisfaction. Specifically, spoons (vs. tongs) increases satisfaction and increased satisfaction increases the amount of food served.

5.2 Objective and study design

The main objective of study 3 is to investigate if serving effort and satisfaction has a mediating effect on the relationship between serving utensil and food served. Mediation refers to when a relationship between a predictor X (serving utensil) and outcome variable Y (food served) can be explained by their relationship with a mediator M (serving effort or satisfaction) (Field, 2013, p. 408; Hayes, 2013, p. 87), (see Figure 7). The mediation effects presented in the hypotheses above will be tested with data collected through an online survey, and analyzed using the PROCESS macro for SPSS (Hayes, 2013).

Figure 7
A simple mediation model with one mediator variable



Note. This figure is from Hayes (2013, p. 7).

We designed study 3 as a single-factor between-subjects experiment with serving utensil (tongs vs. spoon) as the manipulated factor. Data was collected through a survey where participants were randomly assigned to one of two different scenarios. In one of the scenarios spoons were presented as the serving utensil in a self-service buffet, while tongs were used in the other.

5.2.1 Power

A power analysis was also performed in the planning phase of study 3 to determine the appropriate sample size for the study. Using G*Power, we calculated that a total sample size of $N = 2476$ was needed in order to detect effect sizes of $d = 0.10$ with a power level of 80%, given a one-tailed prediction and the standard alpha level of 0.05.

5.2.2 Procedure and sample

The survey was created using "*Nettskjema"*, a digital and secure survey service run by the University of Oslo (Universitetet i Oslo, n.d.-a, n.d.-b) which is also approved as a research tool for collecting data by the University of Agder (Universitetet i Agder, n.d.). Furthermore, this service allows respondents to either conduct the survey on their computers or mobile phones, making it more accessible and convenient for all respondents (Universitetet i Oslo, n.d.-a). In addition to this, Nettskjema allows for randomization, which was a key point with the survey in study 3, as we had two different scenarios that needed to be randomized between participants.

When collecting data for our online survey in study 3, which had no specific inclusion or exclusion criteria, we used a convenience sampling technique, namely first contacting friends and family and then a more widespread electronic e-mail distribution. While e-mail distribution has its advantages such as ease of access, there are also some major disadvantages such as *extremely* low response rate, and that many people consider such invitations as spam which can lead to the researchers being flooded with angry e-mails (Bougie & Sekaran, 2020, p. 144). With this in mind, and us not wanting to annoy anyone, we only sent out one request to participate in the survey, without any follow up mails. To get the respondents we needed, we instead opted for distributing it to a larger sample of people.

Due to the low predicted response rate through e-mail distribution, we collected a long list of e-mail addresses. The e-mail list contained a combination of students we have had courses with which were accessible directly through the university e-mail systems, while the rest and the absolute majority of the e-mail addresses came from publicly available addresses listed at the different Norwegian county municipality websites. At these websites, every person who works in the county municipality is listed, with contact information publicly available.

We kept the survey online for a total of 8 days, and during this period 3030 people started the survey and 82.71% of them completed it giving us a total of 2506 survey responses. The high percentage of people completing the survey indicates that it was not too long.

5.3 Survey design

The survey for study 3 (Appendix E) consisted of a total of 35 questions, where some of the questions were related to another research project. The questions covered general demographics, control questions, manipulation checks and finally an attention check. The demographic questions covered gender, weight, height and whether or not the participant was vegan or vegetarian. The questions relating to weight and height were *not* obligatory to answer, and all answers were totally anonymous. To make sure as many people as possible could participate, we created both a Norwegian and an English version where the participants' preferred language could be chosen at the top of the survey. The survey questions were based both on well established scales from literature[6], as well as observations from our own previously presented intuition study and field study.

[6] Question 6-14 in the survey were adapted from Steptoe et al. (1995, p. 272), and question 19-24 were adapted from Govern & Marsch (2001, p. 369).

5.3.1 Measures

To assess the respondents perception of effort and satisfaction regarding the utensils, a scenario was presented where respondents were asked to envision themselves standing in a buffet about to serve themselves taco with either spoons or tongs. After the scenario was presented, we asked the following question: *"How much of each ingredient below would you take when serving yourself with a spoon/tong"* where the respondents were then tasked to evaluate a total of eight different ingredients, namely *"ground beef"*, *"cheese"*, *"nacho chips"*, *salad"*, *"corn"*, *"tomato"*, *"cucumber"*, and *"onion",* measured on a seven point Likert-scale (1 = A lot less than an average person, 7 = A lot more than an average person). Responses were averaged to compute the index *food served* (Cronbach's α = .60). Although a low Cronbach's alpha could indicate that the items in the index poorly captures the construct, a Cronbach's α = .60 is deemed as acceptable in early stages of research such as this (Churchill, 1979, p. 68).

Furthermore, we asked the questions: *"I find it difficult to pick up food with a spoon/tong"*, and *"I am satisfied with picking up food with a spoon/tong"* measured on a seven-point Likert scale (1 = Totally disagree, 7 = Totally agree). Although we used single-item measures, it has been demonstrated that such measures can have equally high predictive validity as multiple-item measures (Bergkvist & Rossiter, 2007, p. 182). At the end of the survey, there was also an attention check asking: *"Earlier in the survey you were shown a picture and presented a scenario. Did it involve a spoon or a tong?"*.

5.3.2 Data cleaning

Some responses were removed from our data before the analyses. First, we removed all respondents with a BMI of >40 (n = 27), which are categorized as *extremely* or *morbidly* obese (Kitahara et al., 2014, p. 14), and falls within the group of *"obesity class III"*, which is the highest BMI category by the World Health Organization (World Health Organization, 2010). On the other side of the scale, we also removed all respondents with a BMI of <16.5 (n = 7), as these respondents fell into the category of being *severely underweight* (Weir & Jan, 2023). We removed the respondents within these BMI categories as we assumed these respondents had conducted mischievous or faulty reporting (e.g. a height of 17 cm and a weight of 82 kg). We also removed all respondents that failed the final attention check (n = 26), and with only a few respondents failing the attention check, there is an indication that the data is of good quality. The total sample size after data cleaning was N = 2446 for study 3.

5.4 Manipulation check

To check if the participants responded to the manipulation, an independent samples *t*-test was conducted. The manipulation check showed that there was a statistically significant difference between the spoon and tong conditions in terms of effort $t(2442) = 3.64$, $p < 0.001$. The tong condition ($M = 2.20$, $SD = 1.51$) was perceived as significantly more effortful than the spoon condition ($M = 1.99$, $SD = 1.31$).

5.5 Results and discussion

In order to examine the indirect effect of serving utensil on food served, mediated by serving effort (H4A) or satisfaction (H4B), we conducted a mediation analysis using Hayes' (2013, pp. 86–87) simple mediation model. This was done using model 4 in the PROCESS Macro version 4.2 for SPSS, where serving utensil (spoon vs. tongs) was the predictor, serving effort and satisfaction were the mediators and food served was the outcome variable. Standardized coefficients are used to compare the predictor and outcome variable as these were expressed in different units of measurement on a Likert scale (Bougie & Sekaran, 2020, p. 292).

Hypothesis 4A predicts that the relationship between serving utensil and food served is mediated by serving effort, more specifically that spoons (vs. tongs) reduces serving effort and reduced serving effort increases the amount of food served. As shown in Figure 8, we find that serving utensil has a significant influence on serving effort ($\beta = -.15$, $t = -3.64$, $p < .001$), and that serving effort has a significant influence on food served ($\beta = .09$, $t = 3.92$, $p < .001$). A bootstrapping procedure based on 5000 bootstrap samples was used to assess the mediation effect. The result of a 95% confidence interval (CI) indicated that the indirect effect through serving effort was significantly different from zero ($\beta = -0.01$, 95% CI = [-0.024, -0.005]). These findings indicate that using spoons as the serving utensil reduces serving effort, however we do not find proof that reduced serving effort increases the amount of food served, hence, no support for hypothesis 4A.

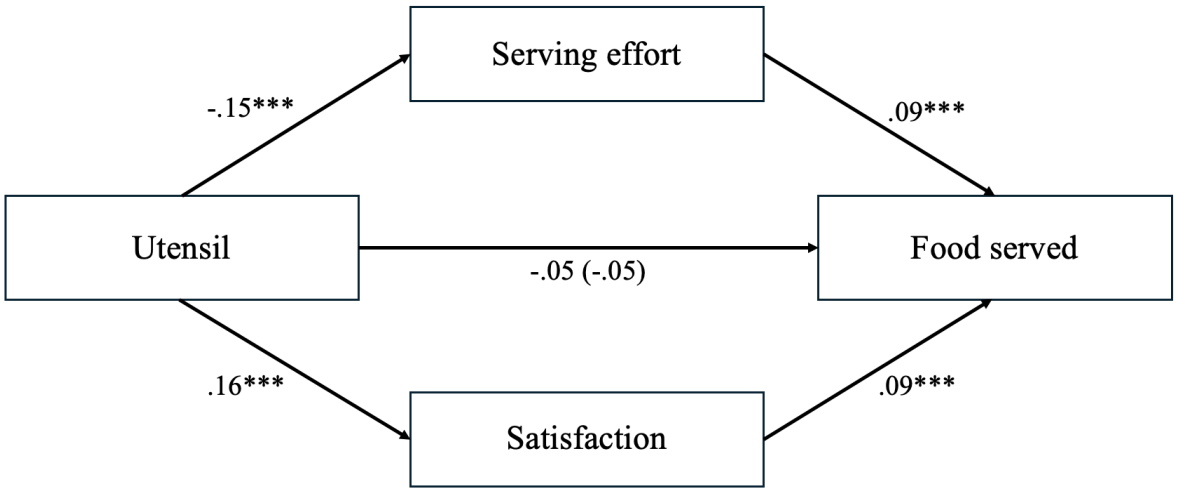
Furthermore, hypothesis 4B predicts that the relationship between serving utensil and food served is mediated by satisfaction, specifically that spoons (vs. tongs) increases satisfaction and increased satisfaction increases the amount of food served. Our findings suggest that serving utensil has an influence on satisfaction ($\beta = .16$, $t = 3.91$, $p < .001$), and satisfaction has a statistically significant effect on food served ($\beta = .09$, $t = 3.83$, $p < .001$). A 95% CI

indicates a statistically significant indirect effect through satisfaction (β = 0.01, 95% CI = [0.005, 0.025]). These findings support hypothesis 4B.

The total effect of serving utensil on food served without any mediators was $F(3, 2440)$ = 7.30, β = -.05, t = -1.31, p = .190, and regressing food served on both serving effort and satisfaction increased the effect of serving utensil with β = -.05, t = -1.33, p = .185.

Figure 8

A simple mediation model with utensil as the predictor, amount of food served as the outcome variable and serving effort and satisfaction as the mediators



*** p < .001.

5.5.1 Discussion

The results from study 3 indicate that both serving effort and satisfaction has a mediating effect on the relationship between serving utensil and amount of food served in a self-service setting. In line with previously hypothesized explanations by Rozin et al. (2011), we do find a relationship between serving utensil and serving effort, however we do not find support for the notion that increased effort decreases amount of food served, with our findings suggesting that people buy *more*, not less food when serving effort increases.

One potential explanation for this could be that people get angry or annoyed when the serving effort increases, and that a reaction to this anger is that they serve themselves more food. This explanation can be supported by previous research, which has found negative emotions and anger to be associated with increased food consumption (Devonport et al., 2019, p. 16; Evers et al., 2010, p. 792; Musaiger et al., 2016, p. 2).

Furthermore, our findings also suggest that serving utensil has an effect on satisfaction, and that when satisfaction increases, so does the amount of food people serve themselves. Satisfaction has not been discussed in previous studies looking at how serving utensil affect food served, and while different to what has been proposed before, it is similar to previous findings indicating that satisfaction has an impact on both spending and food consumption.

6.0 General discussion

Through a series of three different studies with a combined total sample of $N = 4871$, this paper has rigorously explored and measured people's responses to two different nudging interventions related to price display and serving utensils.

The intuition study shows that people expect others to buy less food when price is presented in kilograms as opposed to hectograms, which supports previous findings on the topic of measurement units and the numerosity effect. When testing this in the field however, we do not find an actual behavioral response to the different price displays. In addition to this, findings from the field study conducted in a self-service food environment suggest that people buy less food when serving themselves with tongs, as opposed to spoons. This is in line with what previous studies have found, but contradictory to what people intuitively believe will happen, highlighting that intuitively expected behavior is not always the same as the actual behavioral response observed in a real world environment. Furthermore, findings from our last study indicate that both serving effort and satisfaction mediates the relationship between serving utensil and amount of food served in a self-service setting.

As many others have before, this paper finds that there is a disconnect between how people think others will behave and how people *actually* behave (e.g. Allan et al., 2011, p. 636; Grimmer & Miles, 2017, p. 2; Sainsbury et al., 2013, p. 52; Sheeran & Webb, 2016, p. 511; Sniehotta et al., 2005, p. 143), presenting a compelling case for the importance of constructing nudging interventions, tactics and policies based on real life behavior and highlighting the need for field studies to assess true *practical* effects of nudges.

6.1 Theoretical implications

This paper makes five central theoretical contributions to the nudging literature. First, our work presents people's intuition regarding how different serving utensils affect food

purchasing behavior, with the findings suggesting that people do not intuitively think serving utensil will have an effect on how much food people buy. As nudges are supposed to be non-intrusive, understanding how people experience different nudging interventions, in addition to their actual behavioral reaction, is important.

The second and main theoretical contribution of this paper, is that it adds to the small body of literature in the food nudging domain concerning how serving utensil affects purchasing behavior. This is a niche still in its infancy, with only a few research projects having had this focus in the past, and theoretical value is added by replicating the effect found by Rozin et al. (2011, p. 328) and showing that swapping spoons for tongs is a nudge that can lead to a reduction in food bought in a self-service canteen environment. This finding adds value by suggesting that the effects found in the past, and replicated now, are more than just one-off findings.

Research done by both Wansink et al. (2006) and Rozin et al. (2011) have previously shown that changing serving utensil in a food self-service environment affects amount of food served, and it has been suggested that the most plausible mechanism driving this behavioral change is that when a serving utensil increases the effort it takes to serve food, the amount of food served is reduced (Rozin et al., 2011, p. 330). This claim has however not been scientifically tested in previous studies, and our third theoretical contribution is therefore the suggestion that while there is a relationship between serving utensil and serving effort, increased serving effort leads to an increase in the amount of food served, contrasting previously hypothesised explanations. Our fourth contribution is that we expand the current literature by offering a new potential mechanism, suggesting that satisfaction also mediates the relationship between serving utensil and food served, meaning people are more satisfied when serving themselves with spoons as opposed to tongs, and that increased satisfaction also leads to an increase in food served.

Finally, this paper contributes to the research field by publishing the null-findings from hypothesis 1 and hypothesis 3 (that hg vs. kg price display had no effect on purchasing behavior, and that price display and serving utensil had no additive effect when used in combination), helping counteract publication bias in the nudging research field. To understand the true effects of different nudges, it is important that researchers publish results even when they are unspectacular, as this is the only way to truly build an understanding of

how and when nudges work and when they do not. A recent study comparing findings from nudging interventions that were and were *not* part of the academic literature found that the the average impact of nudges in the literature was very large at 8.7 percentage points, compared to the statistically significant but way lower average impact of 1.4 percentage points of nudges from various nudging units outside of academia (DellaVigna & Linos, 2022, p. 81). This further highlights the publication bias within the field of nudging, and the importance of good research and publication practices.

6.2 Methodological implications

This paper also makes three key methodological contributions to the nudging field. Firstly by having conducted a power analyses for all three studies, and secondly by presenting one of very few pre-registered field studies within the nudging literature. A scoping review consisting of 156 nudging studies found that none of the studies in the review were pre-registered and only 7% had applied a power analysis (Szaszi et al., 2018, p. 355). As the field of behavioral research has been struggling with a replication crisis for some time, and with the nudging literature being used more and more in the development of public policies, ensuring high quality research and reliable findings have become more important than ever. The challenges of replication and publication bias can only be overcome if researchers within the field work together by producing high quality research (Sanders et al., 2018, pp. 156–157), and pre-registration is one solution that could contribute to this (van den Akker et al., 2023, p. 1).

Furthermore, the findings from the intuition based survey conducted in study 1A (kg vs hg) stands in direct opposition to the objective data measuring actual purchasing behavior from our field study. This indicates that although assumptions that seem intuitively *obvious,* like that displaying price in higher measurement units makes them look more expensive and will in turn reduce customer spending, may not always have the same effect in world conditions. A lot of research that has been conducted regarding measurement units has been done in artificial conditions through imaginative scenarios and surveys (e.g. Bagchi & Davis, 2012; Bagchi & Li, 2011; Burson et al., 2009; Fecher et al., 2019; Gourville, 1998; Kwortnik et al., 2006; Liu et al., 2023; Pandelaere et al., 2011; Raghubir & Srivastava, 2002; Shirai, 2017; Siddiqui et al., 2018; Wertenbroch et al., 2007; Yao & Oppewal, 2016a, 2016b), and while this does provide useful insight into how people are affected by measurement units, the

opposing results from the intuition study and the field study in this paper also makes a third methodological contribution by making the case for the importance of researchers investigating effects using different methodological approaches to understand the true *practical* implications of nudges on behavior.

6.3 Practical implications

The results from the field study indicating that people buy less food when serving themselves food with tongs as opposed to spoons may also have some interesting practical implications for a certain kind of food establishment group, namely the all-you-can-eat buffets. In such a buffet, the restaurants sets *"[...] a fixed price for access to food, and they allow customers to consume as much food as wanted at no additional charge"* (Just & Wansink, 2011, p. 193). In such a setting, all reduction in the amount of food bought by the customers would be positive for the establishment, and even though the effect observed in our field study was small, only translating into a 3% decrease of food bought, this could potentially have large financial implications when applied at scale across a whole all-you-can-eat franchise[7].

6.4 Limitations and future research

Our research is not without limitations, and the main limitation of the field study in this research project is that it was not possible to measure the amount of food people bought on an ingredient level through the transaction data. Although our study measured food bought on an individual level, without knowing exactly which ingredients people took less of when using tongs, as opposed to spoons, the practical implications of the current study are somewhat limited. Knowing the effect per individual on an ingredient level could possibly allow these interventions to be used to nudge people towards serving themselves more of healthy ingredients and less unhealthy ones. In addition to this, serving utensil could be optimized on an ingredient level for financial optimization, making people serve themselves more of the ingredients with higher margins and vice versa. We suggest that future research try to replicate the effect, but measure amount of food bought on an ingredient level for each individual purchase.

Another limitation is that we only had one set of sizes for the spoons and tongs. Although we find statistically significant evidence that swapping tongs for spoon decreases the amount of

[7]Assuming the effect observed in the field study is both replicable and stable over time.

food people buy measured in weight, some might question its practical significance due to the low effect size of $d = 0.09$. The low effect size could possibly be explained by slotted spoons used in the field study being relatively small (at least compared to the three times larger spoons used by Rozin et al. (2011, p. 328) in a similar field study), and that the difference in serving effort between the two utensils was not that big. We therefore suggest that future research focus on testing this mechanism in more detail, by conducting field studies with varying sizes of tongs and spoons to better understand the relationship between serving utensil and purchasing behavior.

7.0 Conclusion

In this paper, we have measured responses to two different nudging interventions through three separate high-powered studies, including one pre-registered field study. We find that people have a clear intuitive belief that displaying the price in a self-service canteen in kilograms, as opposed to hectograms, would lead to a decrease in food purchased. The objective data from our field study measuring true *actual* behavior, does however show that in the context of this study, price display had no significant effect on purchasing behavior. In addition to this, we find that people intuitively believe swapping serving utensils would have no effect on purchasing behavior, however findings from our field study indicate that swapping spoons for tongs leads to a significant reduction in the amount of food bought. Furthermore, this paper presents new evidence indicating that the relationship between serving utensil and food served is mediated by both serving effort and satisfaction.

In conclusion, these findings contribute to the nudging literature by thoroughly testing and reporting the effect of two different nudging interventions related to price display and serving utensil, and thereby extending the literature within the the nudging categories of decision information and decision structure in the context of food.

References

Allan, J. L., Johnston, M., & Campbell, N. (2011). Missed by an inch or a mile? Predicting the size of intention-behaviour gap from measures of executive control. *Psychology & Health*, *26*(6), 635–650. https://doi.org/10.1080/08870441003681307

American Psychological Association. (2021). *Preregistration*. American Psychological Association. https://www.apa.org/pubs/journals/resources/preregistration

Anderson, E. W., & Sullivan, M. W. (1993). The Antecedents and Consequences of Customer Satisfaction for Firms. *Marketing Science*, *12*(2), 125–143. https://doi.org/10.1287/mksc.12.2.125

Angner, E. (2021). *A Course in Behavioral Economics* (3rd ed.). Bloomsbury Academic.

Arno, A., & Thomas, S. (2016). The efficacy of nudge theory strategies in influencing adult dietary behaviour: A systematic review and meta-analysis. *BMC Public Health*, *16*. https://doi.org/10.1186/s12889-016-3272-x

Bagchi, R., & Davis, D. F. (2012). $29 for 70 Items or 70 Items for $29? How Presentation Order Affects Package Perceptions. *The Journal of Consumer Research*, *39*(1), 62–73. https://doi.org/10.1086/661893

Bagchi, R., & Davis, D. F. (2016). The role of numerosity in judgments and decision-making. *Current Opinion in Psychology*, *10*, 89–93. https://doi.org/10.1016/j.copsyc.2015.12.010

Bagchi, R., & Li, X. (2011). Illusionary Progress in Loyalty Programs: Magnitudes, Reward Distances, and Step-Size Ambiguity. *The Journal of Consumer Research*, *37*(5), 888–901. https://doi.org/10.1086/656392

Baldwin, S. A. (2017). Improving the rigor of psychophysiology research. *International Journal of Psychophysiology*, *111*, 5–16. https://doi.org/10.1016/j.ijpsycho.2016.04.006

Banerjee, A. V., Duflo, E., Glennerster, R., & Kothari, D. (2010). Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation

campaigns with and without incentives. *BMJ* , *340*. https://doi.org/10.1136/bmj.c2220

Bergkvist, L., & Rossiter, J. R. (2007). The Predictive Validity of Multiple-Item versus Single-Item Measures of the Same Constructs. *Journal of Marketing Research*, *44*(2), 175–184. https://doi.org/10.1509/jmkr.44.2.175

Bougie, R., & Sekaran, U. (2020). *Research methods for business: A skill-building approach* (8th ed.). John Wiley & Sons.

Brandon, A., List, J. A., Metcalfe, R. D., Price, M. K., & Rundhammer, F. (2019). Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(12), 5293–5298. https://doi.org/10.1073/pnas.1802874115

Brunner, T. A., & Siegrist, M. (2012). Reduced food intake after exposure to subtle weight-related cues. *Appetite*, *58*(3), 1109–1112. https://doi.org/10.1016/j.appet.2012.03.010

Burson, K. A., Larrick, R. P., & Lynch, J. G., Jr. (2009). Six of One, Half Dozen of the Other: Expanding and Contracting Numerical Dimensions Produces Preference Reversals. *Psychological Science*, *20*(9), 1074–1078. https://doi.org/10.1111/j.1467-9280.2009.02394.x

Byrd, K., & Almanza, B. (2021). Restaurant menu labeling for calories and sodium: Effect of consumer mindset of immediate versus future consequences. *Journal of Foodservice Business Research*, *24*(3), 310–347. https://doi.org/10.1080/15378020.2020.1849765

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. https://doi.org/10.1037/rev0000445

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject

and within-subject design. *Journal of Economic Behavior & Organization*, *81*(1), 1–8. https://doi.org/10.1016/j.jebo.2011.08.009

Chow, I. H.-S., Lau, V. P., Lo, T. W.-C., Sha, Z., & Yun, H. (2007). Service quality in restaurant operations in China: Decision- and experiential-oriented perspectives. *International Journal of Hospitality Management*, *26*(3), 698–710. https://doi.org/10.1016/j.ijhm.2006.07.001

Churchill, G. A. (1979). A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research*, *16*(1), 64–73. https://doi.org/10.1177/002224377901600110

Cioffi, C. E., Levitsky, D. A., Pacanowski, C. R., & Bertz, F. (2015). A nudge in a healthy direction. The effect of nutrition labels on food purchasing behaviors in university dining facilities. *Appetite*, *92*, 7–14. https://doi.org/10.1016/j.appet.2015.04.053

Cohen, J. (1992). Statistical Power Analysis. *Current Directions in Psychological Science*, *1*(3), 98–101. https://doi.org/10.1111/1467-8721.ep10768783

Coren, S., & Miller, J. (1974). Size contrast as a function of figural similarity. *Perception & Psychophysics*, *16*(2), 355–357. https://doi.org/10.3758/BF03203955

Creswell, J. W. (1994). *Research design: Qualitative & quantitative approaches*. SAGE.

Dannenberg, A., & Weingärtner, E. (2023). The effects of observability and an information nudge on food choice. *Journal of Environmental Economics and Management*, *120*. https://doi.org/10.1016/j.jeem.2023.102829

Dayan, E., & Bar-Hillel, M. (2011). Nudge to nobesity II: Menu positions influence food orders. *Judgment and Decision Making*, *6*(4), 333–342. https://doi.org/10.1017/S1930297500001947

DellaVigna, S., & Linos, E. (2022). RCTs to Scale: Comprehensive Evidence from Two Nudge Units. *Econometrica: Journal of the Econometric Society*, *90*(1), 81–116.

https://doi.org/10.3982/ecta18709

Devonport, T. J., Nicholls, W., & Fullerton, C. (2019). A systematic review of the association between emotions and eating behaviour in normal and overweight adult populations. *Journal of Health Psychology*, *24*(1), 3–24. https://doi.org/10.1177/1359105317697813

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding Self-Report Bias in Organizational Behavior Research. *Journal of Business and Psychology*, *17*(2), 245–260. https://doi.org/10.1023/A:1019637632584

Doucouliagos, C. (1994). A Note on the Evolution of Homo Economicus. *Journal of Economic Issues*, *28*(3), 877–883. https://doi.org/10.1080/00213624.1994.11505586

Eisend, M. (2015). Have We Progressed Marketing Knowledge? A Meta-Meta-Analysis of Effect Sizes in Marketing Research. *Journal of Marketing*, *79*(3), 23–40. https://doi.org/10.1509/jm.14.0288

Eisend, M., Pol, G., Niewiadomska, D., Riley, J., & Wedgeworth, R. (2024). How Much Have We Learned about Consumer Research? A Meta-Meta-Analysis. *The Journal of Consumer Research*, *51*(1), 180–190. https://doi.org/10.1093/jcr/ucad062

Erdem, S., Campbell, D., & Mc Carthy, T. (2022). Can traffic light labelling nudge heuristical decision processes? *Journal of Behavioral Economics for Policy*, *6*(1), 37–45.

Evers, C., Marijn Stok, F., & de Ridder, D. T. D. (2010). Feeding Your Feelings: Emotion Regulation Strategies and Emotional Eating. *Personality & Social Psychology Bulletin*, *36*(6), 792–804. https://doi.org/10.1177/0146167210371383

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. https://doi.org/10.1016/j.jretconser.2019.03.017

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and*

rock "n" roll (4th ed.). SAGE.

Field, A. P., & Hole, G. (2003). *How to Design and Report Experiments*. SAGE.

Fornell, C., Rust, R. T., & Dekimpe, M. G. (2010). The Effect of Customer Satisfaction on Consumer Spending Growth. *Journal of Marketing Research*, *47*(1), 28–35. https://doi.org/10.1509/jmkr.47.1.28

Friis, R., Skov, L. R., Olsen, A., Appleton, K. M., Saulais, L., Dinnella, C., Hartwell, H., Depezay, L., Monteleone, E., Giboreau, A., & Perez-Cueto, F. J. A. (2017). Comparison of three nudge interventions (priming, default option, and perceived variety) to promote vegetable consumption in a self-service buffet setting. *PLoS ONE*, *12*(5). https://doi.org/10.1371/journal.pone.0176028

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. https://doi.org/10.1177/2515245919847202

Godden, E., Cutello, C. A., & Dens, N. (2024). The impact of nutritional labeling on adult snack choices: A controlled field experiment in a non-commercial professional setting. *Appetite*, *193*. https://doi.org/10.1016/j.appet.2023.107167

Gonçalves, D., Coelho, P., Martinez, L. F., & Monteiro, P. (2021). Nudging Consumers toward Healthier Food Choices: A Field Study on the Effect of Social Norms. *Sustainability*, *13*(4). https://doi.org/10.3390/su13041660

Goodwin, T. (2012). Why We Should Reject "Nudge." *Politics* , *32*(2), 85–92. https://doi.org/10.1111/j.1467-9256.2012.01430.x

Gourville, J. T. (1998). Pennies-a-Day: The Effect of Temporal Reframing on Transaction Evaluation. *The Journal of Consumer Research*, *24*(4), 395–408. https://doi.org/10.1086/209517

Govern, J. M., & Marsch, L. A. (2001). Development and Validation of the Situational

Self-Awareness Scale. *Consciousness and Cognition*, *10*(3), 366–378. https://doi.org/10.1006/ccog.2001.0506

Grimmer, M., & Miles, M. P. (2017). With the best of intentions: a large sample test of the intention-behaviour gap in pro-environmental consumer behaviour. *International Journal of Consumer Studies*, *41*(1), 2–10. https://doi.org/10.1111/ijcs.12290

Hair, J. F., Money, A. H., Samouel, P., & Page, M. (2007). *Research methods for business*. Wiley.

Ha, J., & Jang, S. (2010). Effects of service quality and food quality: The moderating role of atmospherics in an ethnic restaurant segment. *International Journal of Hospitality Management*, *29*(3), 520–529. https://doi.org/10.1016/j.ijhm.2009.12.005

Hanks, A. S., Just, D. R., Smith, L. E., & Wansink, B. (2012). Healthy convenience: nudging students toward healthier choices in the lunchroom. *Journal of Public Health*, *34*(3), 370–376. https://doi.org/10.1093/pubmed/fds003

Hanks, A. S., Just, D. R., & Wansink, B. (2013). Smarter Lunchrooms Can Address New School Lunchroom Guidelines and Childhood Obesity. *The Journal of Pediatrics*, *162*(4), 867–869. https://doi.org/10.1016/j.jpeds.2012.12.031

Hausman, D. M., & Welch, B. (2010). Debate: To Nudge or Not to Nudge. *The Journal of Political Philosophy*, *18*(1), 123–136. https://doi.org/10.1111/j.1467-9760.2009.00351.x

Hayes, A. (2013). *Introduction to mediation, moderation, and Conditional Process Analysis: A Regression-based approach*. The Guilford Press.

Huitink, M., Poelman, M. P., van den Eynde, E., Seidell, J. C., & Dijkstra, S. C. (2020). Social norm nudges in shopping trolleys to promote vegetable purchases: A quasi-experimental study in a supermarket in a deprived urban area in the Netherlands. *Appetite*, *151*. https://doi.org/10.1016/j.appet.2020.104655

IBM. (n.d.). *IBM SPSS Statistics*. Retrieved April 25, 2024, from

https://www.ibm.com/products/spss-statistics

In, J., Kang, H., Kim, J. H., Kim, T. K., Ahn, E. J., Lee, D. K., Lee, S., & Park, J. H. (2020). Tips for troublesome sample-size calculation. *Korean Journal of Anesthesiology*, *73*(2), 114–120. https://doi.org/10.4097/kja.19497

Jang, S., & Namkung, Y. (2009). Perceived quality, emotions, and behavioral intentions: Application of an extended Mehrabian–Russell model to restaurants. *Journal of Business Research*, *62*(4), 451–460. https://doi.org/10.1016/j.jbusres.2008.01.038

Jue, J. J. S., Press, M. J., McDonald, D., Volpp, K. G., Asch, D. A., Mitra, N., Stanowski, A. C., & Loewenstein, G. (2012). The impact of price discounts and calorie messaging on beverage consumption: A multi-site field study. *Preventive Medicine*, *55*(6), 629–633. https://doi.org/10.1016/j.ypmed.2012.10.009

Just, D., & Wansink, B. (2011). The Flat-Rate Pricing Paradox: Conflicting Effects of "All-You-Can-Eat" Buffet Pricing. *The Review of Economics and Statistics*, *93*(1), 193–200. https://doi.org/10.1162/REST_a_00057

Kallbekken, S., & Sælen, H. (2013). "Nudging" hotel guests to reduce food waste as a win–win environmental measure. *Economics Letters*, *119*(3), 325–327. https://doi.org/10.1016/j.econlet.2013.03.019

Kamakura, W. A., Mittal, V., de Rosa, F., & Mazzon, J. A. (2002). Assessing the Service-Profit Chain. *Marketing Science*, *21*(3), 294–317. https://doi.org/10.1287/mksc.21.3.294.140

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). https://doi.org/10.1371/journal.pone.0238642

Kang, H. (2021). Sample size determination and power analysis using the G*Power software. *Journal of Educational Evaluation for Health Professions*, *18*.

https://doi.org/10.3352/jeehp.2021.18.17

Keller, C., Markert, F., & Bucher, T. (2015). Nudging product choices: The effect of position change on snack bar choice. *Food Quality and Preference*, *41*, 41–43. https://doi.org/10.1016/j.foodqual.2014.11.005

Kelley, H. H. (1992). Common-Sense Psychology and Scientific Psychology. *Annual Review of Psychology*, *43*, 1–24. https://doi.org/10.1146/annurev.ps.43.020192.000245

Kim, J., Hwang, E., Park, J., Lee, J. C., & Park, J. (2019). Position Effects of Menu Item Displays in Consumer Choices: Comparisons of Horizontal Versus Vertical Displays. *Cornell Hospitality Quarterly*, *60*(2), 116–124. https://doi.org/10.1177/1938965518778234

Kingham, A., Kemps, E., Prichard, I., & Tiggemann, M. (2023). The effect of spatial separation on food and drink choices from an online menu. *Eating Behaviors*, *51*. https://doi.org/10.1016/j.eatbeh.2023.101816

Kitahara, C. M., Flint, A. J., Berrington de Gonzalez, A., Bernstein, L., Brotzman, M., MacInnis, R. J., Moore, S. C., Robien, K., Rosenberg, P. S., Singh, P. N., Weiderpass, E., Adami, H. O., Anton-Culver, H., Ballard-Barbash, R., Buring, J. E., Freedman, D. M., Fraser, G. E., Beane Freeman, L. E., Gapstur, S. M., … Hartge, P. (2014). Association between Class III Obesity (BMI of 40–59 kg/m2) and Mortality: A Pooled Analysis of 20 Prospective Studies. *PLoS Medicine*, *11*(7). https://doi.org/10.1371/journal.pmed.1001673

Knowles, D., Brown, K., & Aldrovandi, S. (2019). Exploring the underpinning mechanisms of the proximity effect within a competitive food environment. *Appetite*, *134*, 94–102. https://doi.org/10.1016/j.appet.2018.12.005

Kosīte, D., König, L. M., De-loyde, K., Lee, I., Pechey, E., Clarke, N., Maynard, O., Morris, R. W., Munafò, M. R., Marteau, T. M., Fletcher, P. C., & Hollands, G. J. (2019). Plate

size and food consumption: a pre-registered experimental study in a general population sample. *The International Journal of Behavioral Nutrition and Physical Activity*, *16*. https://doi.org/10.1186/s12966-019-0826-1

Kosters, M., & Van der Heijden, J. (2015). From mechanism to virtue: Evaluating Nudge theory. *Evaluation*, *21*(3), 276–291. https://doi.org/10.1177/1356389015590218

Kwortnik, R. J., Creyer, E. H., & Ross, W. T. (2006). Usage-Based versus Measure-Based Unit Pricing: Is there a Better Index of Value? *Journal of Consumer Policy*, *29*, 37–66. https://doi.org/10.1007/s10603-005-6054-x

Laurijssen, S. J., van der Graaf, R., van Dijk, W. B., Schuit, E., Groenwold, R. H., Grobbee, D. E., & de Vries, M. C. (2022). When is it impractical to ask informed consent? A systematic review. *Clinical Trials* , *19*(5), 545–560. https://doi.org/10.1177/17407745221103567

Libotte, E., Siegrist, M., & Bucher, T. (2014). The influence of plate size on meal composition. Literature review and experiment. *Appetite*, *82*, 91–96. https://doi.org/10.1016/j.appet.2014.07.010

Lin, H. (n.d.). *Convert between different effect sizes*. Retrieved April 25, 2024, from https://www.escal.site/

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. https://doi.org/10.1002/cjas.1709

Maier, M., Bartoš, F., Stanley, T. D., Shanks, D. R., Harris, A. J. L., & Wagenmakers, E.-J. (2022). No evidence for nudging after adjusting for publication bias. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(31). https://doi.org/10.1073/pnas.2200300119

Mertens, S., Herberz, M., Hahnel, U. J. 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 of the United States of America*, *119*(1). https://doi.org/10.1073/pnas.2107346118

Miller, J., & Ulrich, R. (2019). The quest for an optimal alpha. *PloS One*, *14*(1). https://doi.org/10.1371/journal.pone.0208631

Mills, C. (2015). The Heteronomy of Choice Architecture. *Review of Philosophy and Psychology*, *6*, 495–509. https://doi.org/10.1007/s13164-015-0242-7

Montagni, I., Prevot, F., Castro, Z., Goubel, B., Perrin, L., Oppert, J.-M., & Fontvieille, A.-M. (2020). Using Positive Nudge to Promote Healthy Eating at Worksite: A Food Labeling Intervention. *Journal of Occupational and Environmental Medicine*, *62*(6). https://doi.org/10.1097/JOM.0000000000001861

Münscher, R., Vetter, M., & Scheuerle, T. (2016). A Review and Taxonomy of Choice Architecture Techniques. *Journal of Behavioral Decision Making*, *29*(5), 511–524. https://doi.org/10.1002/bdm.1897

Musaiger, A. O., Nabag, F. O., & Al-Mannai, M. A. (2016). Is emotional eating a problem of concern among adolescents in poor countries? An example from Sudan. *International Journal of Adolescent Medicine and Health*, *29*(5). https://doi.org/10.1515/ijamh-2016-0049

Pandelaere, M., Briers, B., & Lembregts, C. (2011). How to Make a 29% Increase Look Bigger: The Unit Effect in Option Comparisons. *The Journal of Consumer Research*, *38*(2), 308–322. https://doi.org/10.1086/659000

Perez-Cueto, F. J. A. (2021). Nudging plant-based meals through the menu. *International Journal of Gastronomy and Food Science*, *24*. https://doi.org/10.1016/j.ijgfs.2021.100346

Raghubir, P., & Srivastava, J. (2002). Effect of Face Value on Product Valuation in Foreign Currencies. *The Journal of Consumer Research*, *29*(3), 335–347. https://doi.org/10.1086/344430

random.org. (n.d.). *What's this fuss about true randomness?* Retrieved March 18, 2024, from https://www.random.org/

Reinholdsson, T., Hedesström, M., Ejelöv, E., Hansla, A., Bergquist, M., Svenfelt, Å., & Nilsson, A. (2023). Nudging green food: The effects of a hedonic cue, menu position, a warm‐glow cue, and a descriptive norm. *Journal of Consumer Behaviour*, *22*(3), 557–568. https://doi.org/10.1002/cb.2129

Richardson, R., Prescott, M. P., & Ellison, B. (2021). Impact of plate shape and size on individual food waste in a university dining hall. *Resources, Conservation and Recycling*, *168*. https://doi.org/10.1016/j.resconrec.2020.105293

Rising, C. J., & Bol, N. (2017). Nudging Our Way to a Healthier Population: The Effect of Calorie Labeling and Self-Control on Menu Choices of Emerging Adults. *Health Communication*, *32*(8), 1032–1038. https://doi.org/10.1080/10410236.2016.1217452

Roberts, P., Priest, H., & Traynor, M. (2006). Reliability and validity in research. *Nursing Standard*, *20*(44), 41–45.

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. https://doi.org/10.1017/S1930297500001935

Ruggeri, K. (Ed.). (2022). *Psychology and behavioral economics: Applications for public policy* (2nd ed.). Routledge.

Sainsbury, K., Mullan, B., & Sharpe, L. (2013). Gluten free diet adherence in coeliac disease. The role of psychological symptoms in bridging the intention-behaviour gap. *Appetite*, *61*, 52–58. https://doi.org/10.1016/j.appet.2012.11.001

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. https://doi.org/10.1017/bpp.2018.17

Schmidt, A. T., & Engelen, B. (2020). The ethics of nudging: An overview. *Philosophy Compass*, *15*(4). https://doi.org/10.1111/phc3.12658

Sedgwick, P. (2013). Convenience sampling. *BMJ* , *347*. https://doi.org/10.1136/bmj.f6304

Selvamuthu, D., & Das, D. (2018). *Introduction to Statistical Methods, Design of Experiments and Statistical Quality Control*. Springer Nature Singapore. https://doi.org/10.1007/978-981-13-1736-1

Seward, M. W., Block, J. P., & Chatterjee, A. (2016). A Traffic-Light Label Intervention and Dietary Choices in College Cafeterias. *American Journal of Public Health*, *106*(10), 1808–1814. https://doi.org/10.2105/AJPH.2016.303301

Sheeran, P., & Webb, T. L. (2016). The Intention–Behavior Gap. *Social and Personality Psychology Compass*, *10*(9), 503–518. https://doi.org/10.1111/spc3.12265

Shirai, M. (2017). Effects of price reframing tactics on consumer perceptions. *Journal of Retailing and Consumer Services*, *34*, 82–87. https://doi.org/10.1016/j.jretconser.2016.09.009

Siddiqui, R. A., Monga, A., & Buechel, E. C. (2018). When Intertemporal Rewards Are Hedonic, Larger Units of Wait Time Boost Patience. *Journal of Consumer Psychology: The Official Journal of the Society for Consumer Psychology*, *28*(4), 612–628. https://doi.org/10.1002/jcpy.1019

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Simon, H. A. (1990). Bounded Rationality. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *Utility and Probability* (pp. 15–18). Palgrave Macmillan.

https://doi.org/10.1007/978-1-349-20568-4_5

Sniehotta, F. F., Scholz, U., & Schwarzer, R. (2005). Bridging the intention–behaviour gap: Planning, self-efficacy, and action control in the adoption and maintenance of physical exercise. *Psychology & Health*, *20*(2), 143–160. https://doi.org/10.1080/08870440512331317670

Sommet, N., Weissman, D. L., Cheutin, N., & Elliot, A. J. (2023). How Many Participants Do I Need to Test an Interaction? Conducting an Appropriate Power Analysis and Achieving Sufficient Power to Detect an Interaction. *Advances in Methods and Practices in Psychological Science*, *6*(3). https://doi.org/10.1177/25152459231178728

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. https://doi.org/10.1016/j.appet.2016.11.037

Steptoe, A., Pollard, T. M., & Wardle, J. (1995). Development of a Measure of the Motives Underlying the Selection of Food: the Food Choice Questionnaire. *Appetite*, *25*(3), 267–284. https://doi.org/10.1006/appe.1995.0061

Sukamolson, S. (2007). Fundamentals of quantitative research. *Language Institute Chulalongkorn University*, *1*(3), 1–20.

Sulek, J. M., & Hensley, R. L. (2004). The Relative Importance of Food, Atmosphere, and Fairness of Wait: The Case of a Full-service Restaurant. *The Cornell Hotel and Restaurant Administration Quarterly*, *45*(3), 235–247. https://doi.org/10.1177/0010880404265345

Suleman, S., Sweeney-Magee, M., Pinkney, S., Charbonneau, K., Banh, K., Hale, I., & Amed, S. (2022). Evaluation of two social norms nudge interventions to promote healthier food choices in a Canadian grocery store. *BMC Public Health*, *22*. https://doi.org/10.1186/s12889-022-14370-8

Sunstein, C. R. (2022). Sludge Audits. *Behavioural Public Policy* , *6*(4), 654–673. https://doi.org/10.1017/bpp.2019.32

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. https://doi.org/10.1002/bdm.2035

Thaler, R. H. (2018). Nudge, not sludge. *Science*, *361*(6401), 431. https://doi.org/10.1126/science.aau9241

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Thaler, R. H., & Sunstein, C. R. (2021). *Nudge: Improving decisions about health, wealth and happiness* (Final edition.). Yale University Press.

Thaler, R. H., Sunstein, C. R., & Balz, J. P. (2012). Choice Architecture. In E. Shafir (Ed.), *The Behavioral Foundations of Public Policy* (pp. 428–439). Princeton University Press. https://doi.org/10.13140/2.1.4195.2321

The Nobel Prize. (n.d.). *Herbert Simon: Facts*. NobelPrize.org. Retrieved January 16, 2024, from https://www.nobelprize.org/prizes/economic-sciences/1978/simon/facts/

Thorndike, A. N. (2020). Healthy choice architecture in the supermarket: Does it work? *Social Science & Medicine*, *266*. https://doi.org/10.1016/j.socscimed.2020.113459

Thorndike, A. N., Gelsomin, E. D., McCurley, J. L., & Levy, D. E. (2019). Calories Purchased by Hospital Employees After Implementation of a Cafeteria Traffic Light-Labeling and Choice Architecture Program. *JAMA Network Open*, *2*(7). https://doi.org/10.1001/jamanetworkopen.2019.6789

Thorndike, A. N., Riis, J., Sonnenberg, L. M., & Levy, D. E. (2014). Traffic-Light Labels and Choice Architecture: Promoting Healthy Food Choices. *American Journal of Preventive*

Medicine, *46*(2), 143–149. https://doi.org/10.1016/j.amepre.2013.10.002

Titchener, E. B. (1901). *Experimental Psychology: A Manual of Laboratory Practice* (Vol. 1). The MacMillan Company.

Tolman, E. C. (1932). *Purposive Behavior in Animals and Men*. University of California Press.

Universitetet i Agder. (n.d.). *Nettskjema*. Retrieved April 23, 2024, from https://old.uia.no/bibliotek/forskning-og-publisering2/nettskjema

Universitetet i Oslo. (n.d.-a). *Hva er Nettskjema*. Retrieved April 23, 2024, from https://www.uio.no/tjenester/it/adm-app/nettskjema/mer-om/

Universitetet i Oslo. (n.d.-b). *Nettskjema*. Retrieved April 23, 2024, from https://www.uio.no/tjenester/it/adm-app/nettskjema/

van den Akker, O. R., van Assen, M. A. L. M., Bakker, M., Elsherif, M., Wong, T. K., & Wicherts, J. M. (2023). Preregistration in practice: A comparison of preregistered and non-preregistered studies in psychology. *Behavior Research Methods*. https://doi.org/10.3758/s13428-023-02277-0

VanEpps, E. M., Downs, J. S., & Loewenstein, G. (2016). Calorie Label Formats: Using Numeric and Traffic Light Calorie Labels to Reduce Lunch Calories. *Journal of Public Policy & Marketing*, *35*(1), 26–36. https://doi.org/10.1509/jppm.14.112

Van Gestel, L. C., Kroese, F. M., & De Ridder, D. T. D. (2018). Nudging at the checkout counter - A longitudinal study of the effect of a food repositioning nudge on healthy food choice. *Psychology & Health*, *33*(6), 800–809. https://doi.org/10.1080/08870446.2017.1416116

van Kleef, E., Seijdell, K., Vingerhoeds, M. H., de Wijk, R. A., & van Trijp, H. C. M. (2018). The effect of a default-based nudge on the choice of whole wheat bread. *Appetite*, *121*, 179–185. https://doi.org/10.1016/j.appet.2017.11.091

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial Incentive–Based Approaches for Weight Loss: A Randomized Trial. *The Journal of the American Medical Association*, *300*(22), 2631–2637. https://doi.org/10.1001/jama.2008.804

Wadhera, D., & Capaldi-Phillips, E. D. (2014). A review of visual cues associated with food on food acceptance and consumption. *Eating Behaviors*, *15*(1), 132–143. https://doi.org/10.1016/j.eatbeh.2013.11.003

Wansink, B., & Just, D. R. (2016). The limits of defaults: why french fries trump apple slices. *BMC Research Notes*, *9*, 263. https://doi.org/10.1186/s13104-016-2061-z

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. https://doi.org/10.1016/j.amepre.2006.04.003

Web of Science. (n.d.). *Citation report*. Retrieved May 24, 2024, from https://www.webofscience.com/wos/woscc/citation-report/73705c93-d8f4-4c94-8647-44116f71e848-de7184cf

Weir, C. B., & Jan, A. (2023, June 26). *BMI Classification Percentile And Cut Off Points*. National Library of Medicine; StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK541070/

Wertenbroch, K., Soman, D., & Chattopadhyay, A. (2007). On the Perceived Value of Money: The Reference Dependence of Currency Numerosity Effects. *The Journal of Consumer Research*, *34*(1), 1–10. https://doi.org/10.1086/513041

White, M. (2013). *The Manipulation of Choice: Ethics and Libertarian Paternalism*. Palgrave Macmillan.

Wiles, R., Crow, G., Charles, V., & Heath, S. (2007). Informed Consent and the Research Process: Following Rules or Striking Balances? *Sociological Research Online*, *12*(2),

99–110. https://doi.org/10.5153/sro.1208

Wilkinson, T. M. (2013). Nudging and Manipulation. *Political Studies*, *61*(2), 341–355. https://doi.org/10.1111/j.1467-9248.2012.00974.x

Wisdom, J., Downs, J. S., & Loewenstein, G. (2010). Promoting Healthy Choices: Information versus Convenience. *American Economic Journal: Applied Economics*, *2*(2), 164–178. https://doi.org/10.1257/app.2.2.164

World Health Organization. (2010, May 6). *A healthy lifestyle - WHO recommendations*. https://www.who.int/europe/news-room/fact-sheets/item/a-healthy-lifestyle---who-recommendations

Yao, J., & Oppewal, H. (2016a). Unit Pricing Increases Price Sensitivity Even When Products are of Identical Size. *Journal of Retailing*, *92*(1), 109–121. https://doi.org/10.1016/j.jretai.2015.09.002

Yao, J., & Oppewal, H. (2016b). Unit pricing matters more when consumers are under time pressure. *European Journal of Marketing*, *50*(5/6), 1094–1114. https://doi.org/10.1108/EJM-03-2015-0122

Yeung, M. C. H., & Ennew, C. T. (2001). Measuring the impact of customer satisfaction on profitability: A sectoral analysis. *Journal of Targeting, Measurement and Analysis for Marketing*, *10*(2), 106–116. https://doi.org/10.1057/palgrave.jt.5740038

Yeung, M. C. H., Ramasamy, B., Chen, J., & Paliwoda, S. (2013). Customer satisfaction and consumer expenditure in selected European countries. *International Journal of Research in Marketing*, *30*(4), 406–416. https://doi.org/10.1016/j.ijresmar.2013.06.001

Zikmund, W. G., Babin, B. J., Carr, J. C., & Griffin, M. (2010). *Business research methods* (8th ed.). South Western Cengage Learning.

Appendices

Appendix A1
Survey study 1A - Kilogram

Alder: _____

Kjønn:
- ☐ Kvinne
- ☐ Mann
- ☐ Annet
- ☐ Ønsker ikke å oppgi

Spørsmål:

Tror du folk vil kjøpe mer eller mindre mat i bufféen i kantinen dersom pris blir vist i **kilogram** istedenfor hektogram (*se illustrasjonsbilde*)?

Illustrasjonsbilde:

SiA MatogDrikke PLUKK & MIKS 169,- pr. kg	SiA MatogDrikke PLUKK & MIKS 16.90,- pr. hg

Mindre med kilogram			**Ingen forskjell**		**Mer med kilogram**	
1	2	3	4	5	6	7

(Sett ring rundt svaret)

Appendix A2

Survey study 1A - Hectogram

Alder: _____

Kjønn:

- ☐ Kvinne
- ☐ Mann
- ☐ Annet
- ☐ Ønsker ikke å oppgi

Spørsmål:

Tror du folk vil kjøpe mer eller mindre mat i bufféen i kantinen dersom pris blir vist i **hektogram** istedenfor kilogram (*se illustrasjonsbilde*)?

Illustrasjonsbilde:

SIA MatogDrikke **PLUKK & MIKS** **16.90,- pr. hg**	SIA MatogDrikke **PLUKK & MIKS** **169,- pr. kg**

Mindre med hektogram			**Ingen forskjell**		**Mer med hektogram**	
1	2	3	4	5	6	7

(Sett ring rundt svaret)

Appendix B1
Survey study 1B - Spoons

Alder: _____

Kjønn:
- ☐ Kvinne
- ☐ Mann
- ☐ Annet
- ☐ Ønsker ikke å oppgi

Spørsmål:

Tror du folk vil kjøpe mer eller mindre mat i bufféen i kantinen dersom servingsredskapet er **skje** istedenfor klype (*se illustrasjonsbilde*)?

Illustrasjonsbilde:

Mindre med skje			**Ingen forskjell**			**Mer med skje**
1	2	3	4	5	6	7

(Sett ring rundt svaret)

Appendix B2
Survey study 1B - Tongs

Alder: _____

Kjønn:

- ☐ Kvinne
- ☐ Mann
- ☐ Annet
- ☐ Ønsker ikke å oppgi

Spørsmål:

Tror du folk vil kjøpe mer eller mindre mat i bufféen i kantinen dersom servingsredskapet er **klype** istedenfor skje (*se illustrasjonsbilde*)?

Illustrasjonsbilde:

Mindre med klype			**Ingen forskjell**			**Mer med klype**
1	2	3	4	5	6	7

(Sett ring rundt svaret)

Appendix C

Pre-registration of field study

CONFIDENTIAL - FOR PEER-REVIEW ONLY

University cafeteria field study: Nudging people toward eating less (#164189)

Created: 02/29/2024 02:59 AM (PT)

This is an anonymized copy (without author names) of the pre-registration. It was created by the author(s) to use during peer-review.
A non-anonymized version (containing author names) should be made available by the authors when the work it supports is made public.

---

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
Hypothesis 1: Displaying price in NOK/kg (vs. NOK/hg) will reduce the amount of food purchased measured in weight.

Hypothesis 2: Using tongs (vs. spoons) as the serving utensil in the self-service canteen will reduce the amount of food purchased measured in weight.

Hypothesis 3: The display of price in NOK/kg (vs. NOK/hg) in combination with using tongs (vs. spoons) as the serving utensil in the self-service canteen will further enhance the reduction of food purchased measured in weight compared to all other conditions.

3) Describe the key dependent variable(s) specifying how they will be measured.
The dependent variable is food purchased in weight at an individual level, measured through transaction data.

4) How many and which conditions will participants be assigned to?
There will be four different conditions, which have been randomly assigned to four different weeks. Each of the conditions will be run on a separate Friday, keeping food constant across all days/weeks.

The first condition involves using spoons as serving utensil and display price in NOK/hg.
The second condition involves using tongs as serving utensil and display price in NOK/hg.
The third condition involves using spoons as serving utensil and display price in NOK/kg.
The fourth condition involves using tongs as serving utensil and display price in NOK/kg.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
The hypotheses will be tested through a series of three independent samples t-tests using planned contrasts corresponding to the comparisons specified as part of the hypotheses. Given the one-sided predictions, we will use one-tailed tests.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
In the main analyses, no outliers will be excluded. However, as the hypotheses are restricted to the amount of food purchased, we will exclude all other purchases (e.g., hot and cold beverages) and only use such additional purchases for exploratory purposes.

Subsequently, as robustness checks, we will perform similar statistical tests as those used to examine our key hypotheses, whereby we will exclude outliers scoring beyond plus or minus 3.29 standard deviations away from the condition-specific means regarding the weight of food purchased.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
We estimate a total sample size of 2000-3000 transactions during the course of the study period (approximately 500-750 transactions per condition). This estimate is based on historical sales data from the canteen in which the study will take place. This sample size has a statistical power considerably greater than 80% to detect small effect sizes equivalent to d = 0.20, assuming the regular alpha level of 0.05 and one-tailed tests (N = 620 transactions are needed to detect such subtle effects with 80% power, given the alpha level stated above and one-tailed tests).

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
When analyzing a receipt that contains more than one food order, we will treat such a receipt as several individual food orders and not as an aggregated order, as we assume that each weighted food option corresponds to one plate to be consumed by one individual (and, accordingly, that a transaction that contains two or more weighted food options correspond to food purchased for more than one individual).

Appendix D

Exploratory study - Beverages bought between price display conditions NOK/hg vs. NOK/kg

Table D1
Chi-square test of beverages bought between the conditions of price displayed in NOK/hg (n = 1000) vs. NOK/kg (n = 965)

	NOK/hg		NOK/kg			
Item	*n*	*%*	*n*	*%*	χ^2 (1)	*p*
Beverages	213	21	205	21	0.00	.98

Appendix E1

Survey study 3 - Spoon

Gender *

- ○ Woman
- ○ Man
- ○ Other
- ○ Prefer not to answer

Weight (**kg**)

Not obligatory, and if you do not wish to answer you can leave the field empty.

Height (**cm**)

IMPORTANT: State your height in **centimeters**.

Not obligatory, and if you do not wish to answer you can leave the field empty.

Are you vegetarian / vegan? *

- ○ Yes
- ○ No
- ○ Prefer not to answer

It is important to me that the food I eat on a typical day:

1 = Not important at all // 7 = Very important

	1	2	3	4	5	6	7
Contains a lot of vitamins and minerals *	○	○	○	○	○	○	○
Keeps me healthy *	○	○	○	○	○	○	○
Is nutritious *	○	○	○	○	○	○	○
Is high in protein *	○	○	○	○	○	○	○
Is good for my skin/teeth/hair/nails etc. *	○	○	○	○	○	○	○
Is high in fibre *	○	○	○	○	○	○	○
Is low in calories *	○	○	○	○	○	○	○
Helps me control my weight *	○	○	○	○	○	○	○
Is low in fat *	○	○	○	○	○	○	○

Do you like tacos? *

Not at all **Love tacos**

1 2 3 4 5 6 7

Value

How often do you eat tacos? *

Never **Very often**

1 2 3 4 5 6 7

Value

How much do you generally eat? *

Very litte A lot

1 2 3 4 5 6 7

Value

How hungry are you now? *

Not hungry Very hungry

1 2 3 4 5 6 7

Value

Scenario:

Imagine that you are at a buffet, and you are about to serve yourself some food. Today, tacos are served for lunch at the buffet. The serving utensil used for the various ingredients is a spoon.

Based on the scenario, answer the following:

1 = Totaly disagree // 7 = Totaly agree

	1	2	3	4	5	6	7
When I am at the buffet, I am conscious of what is going on around me *	○	○	○	○	○	○	○
When I am at the buffet, I am keenly aware of everything in my environment *	○	○	○	○	○	○	○
When I am at the buffet, I am conscious of all objects around me *	○	○	○	○	○	○	○
When I am at the buffet, I am self-conscious about the way I look *	○	○	○	○	○	○	○
When I am at the buffet, I am concerned about the way I present myself *	○	○	○	○	○	○	○
When I am at the buffet, I am concerned about what other people think of me *	○	○	○	○	○	○	○

How much of each ingredient below would you take when serving yourself with a spoon?

1 = A lot less than an avarage person // 7 = A lot more than an avarage person

	1	2	3	4	5	6	7
Ground beef *	○	○	○	○	○	○	○
Cheese *	○	○	○	○	○	○	○
Nacho chips *	○	○	○	○	○	○	○
Salad *	○	○	○	○	○	○	○
Corn *	○	○	○	○	○	○	○
Tomato *	○	○	○	○	○	○	○
Cucumber *	○	○	○	○	○	○	○
Onion *	○	○	○	○	○	○	○

How embarrassing do you find serving yourself the same ingredient multiple times at a buffet? *

Value

I find it difficult to pick up food with a spoon *

Value

I am satisfied with picking up food with a spoon *

Value

Earlier in the survey you were shown a picture and presented a scenario. Did it involve a spoon or a tong? *

- ○ Spoon
- ○ Tong

Appendix E2

Survey study 3 - Tongs

Gender *

- ○ Woman
- ○ Man
- ○ Other
- ○ Prefer not to answer

Weight (**kg**)

Not obligatory, and if you do not wish to answer you can leave the field empty.

Height (**cm**)

IMPORTANT: State your height in **centimeters**.

Not obligatory, and if you do not wish to answer you can leave the field empty.

Are you vegetarian / vegan? *

- ○ Yes
- ○ No
- ○ Prefer not to answer

It is important to me that the food I eat on a typical day:

1 = Not important at all // 7 = Very important

	1	2	3	4	5	6	7
Contains a lot of vitamins and minerals *	○	○	○	○	○	○	○
Keeps me healthy *	○	○	○	○	○	○	○
Is nutritious *	○	○	○	○	○	○	○
Is high in protein *	○	○	○	○	○	○	○
Is good for my skin/teeth/hair/nails etc. *	○	○	○	○	○	○	○
Is high in fibre *	○	○	○	○	○	○	○
Is low in calories *	○	○	○	○	○	○	○
Helps me control my weight *	○	○	○	○	○	○	○
Is low in fat *	○	○	○	○	○	○	○

Do you like tacos? *

Not at all **Love tacos**

1 2 3 4 5 6 7

Value

How often do you eat tacos? *

Value

How much do you generally eat? *

Value

How hungry are you now? *

Value

Scenario:

Imagine that you are at a buffet, and you are about to serve yourself some food. Today, tacos are served for lunch at the buffet. The serving utensil used for the various ingredients is a tong.

Based on the scenario, answer the following:

1 = Totaly disagree // 7 = Totaly agree

	1	2	3	4	5	6	7
When I am at the buffet, I am conscious of what is going on around me *	○	○	○	○	○	○	○
When I am at the buffet, I am keenly aware of everything in my environment *	○	○	○	○	○	○	○
When I am at the buffet, I am conscious of all objects around me *	○	○	○	○	○	○	○
When I am at the buffet, I am self-conscious about the way I look *	○	○	○	○	○	○	○
When I am at the buffet, I am concerned about the way I present myself *	○	○	○	○	○	○	○
When I am at the buffet, I am concerned about what other people think of me *	○	○	○	○	○	○	○

How much of each ingredient below would you take when serving yourself with tongs?

1 = A lot less than an avarage person // 7 = A lot more than an avarage person

	1	2	3	4	5	6	7
Ground beef *	○	○	○	○	○	○	○
Cheese *	○	○	○	○	○	○	○
Nacho chips *	○	○	○	○	○	○	○
Salad *	○	○	○	○	○	○	○
Corn *	○	○	○	○	○	○	○
Tomato *	○	○	○	○	○	○	○
Cucumber *	○	○	○	○	○	○	○
Onion *	○	○	○	○	○	○	○

How embarrassing do you find serving yourself the same ingredient multiple times at a buffet? *

Not embarrassing at all — **Extremely embarrassing**

1 2 3 4 5 6 7

Value

I find it difficult to pick up food with tongs *

Totally disagree — **Totally agree**

1 2 3 4 5 6 7

Value

I am satisfied with picking up food with tongs *

Value

Earlier in the survey you were shown a picture. Did the picture show a spoon or a tong? *

- Spoon
- Tong