Universität Basel

Fakultät für Psychologie

Nudging Healthy Food and Sustainable Choices

Inauguraldissertation zur Erlangung der Würde eines Doktors der Philosophie vorgelegt der Fakultät für Psychologie der Universität Basel von

Swen Jonas Kühne

aus Rieden, SG

Zürich, 2020

Universität
Basel

Fakultät für
Psychologie

Nudging Healthy Food and Sustainable Choices

Inauguraldissertation zur Erlangung der Würde eines Doktors der Philosophie vorgelegt der Fakultät für Psychologie der Universität Basel von

Swen Jonas Kühne

aus Rieden, SG

Zürich, 2020

Universität Basel

Fakultät für Psychologie

Genehmigt von der Fakultät für Psychologie auf Antrag von

Prof. Dr. Klaus Opwis
Prof. Dr. Ester Reijnen

Datum des Doktoratsexamen:

DekanIn der Fakultät für Psychologie

Universität Basel

Fakultät für Psychologie

Erklärung zur wissenschaftlichen Lauterkeit

Ich erkläre hiermit, dass die vorliegende Arbeit ohne die Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel selbstständig verfasst habe. Zu Hilfe genommene Quellen sind als solche gekennzeichnet. Die veröffentlichten oder zur Veröffentlichung in Zeitschriften eingereichten Manuskripte wurden in Zusammenarbeit mit den Koautoren erstellt und von keinem der Beteiligten an anderer Stelle publiziert, zur Publikation eingereicht, oder einer anderen Prüfungsbehörde als Qualifikationsarbeit vorgelegt. Es handelt sich dabei um folgende Manuskripte:

- Reijnen, E., Kühne, S. J., Stöcklin, M., & Wolfe, J. M. (2019). Choosing or rejecting a food item, does framing matter? And what has sugar to do with it!. *Appetite*, *143*. https://doi.org/10.1016/j.appet.2019.104410
- Reijnen, E., Kühne, S. J., von Gugelberg, H. M., & Crameri, A. (2019). Nudged to a menu position: The role of "I'm loving it"! *Journal of Consumer Policy*, *42*(3), 441–453. https://doi.org/10.1007/s10603-019-09413-4
- Kühne, S. J., Reijnen, E., & Crameri, A. (2020). When too few is bad for the environment: Choice set size and default effects for electricity products. *Swiss Journal of Psychology*, *79*(1), 35–41. https://doi.org/10.1024/1421-0185/a000232

Zürich, 19. Mai 2020

Swen Jonas Kühne

Content

Abstract

Reducing overweight and the overuse of natural resources are two of the biggest challenges facing humanity today. The roots of those challenges lie in, amongst other things, our everyday behavior such as the food and electricity products we choose. Although there are attempts to address these challenges, such as information campaigns, these traditional interventions have been ineffective in changing behavior. In recent years, however, the nudge approach has shown that interventions that use so-called "irrelevant factors" (e.g., defaults) have the potential to promote healthy eating and environmentally friendly behavior. This dissertation describes the results of three published manuscripts that investigate different nudges (i.e., frames, defaults, set size, position effects) and the additional factors (e.g., preference, weighting of aspects) that influence their effectiveness in changing behavior in the two targeted fields. The results indicate that nudges can be effective, but they also show the complexity of the interaction of the variables. For example, one manuscript indicates that the preference for a cuisine style mediates whether a center-stage effect (i.e., that we tend to choose a dish in the middle of a menu) occurs or not. Those additional factors have so far barely been investigated or brought into a consistent theory. The dissertation thereby contributes to a theoretical understanding of the studied nudges and discusses current and future challenges such as the lack of theory, freedom of choice and hedonic consequences of the nudge approach.

Nudging Healthy Food and Sustainable Choices

To put it bluntly; we eat too much and use too many natural resources. The consequence of our food choices is that 39% of the world's adults are overweight and 13% are even obese, and the trend is rising (WHO, 2020). In addition, overweight and obesity are associated with severe health consequences for the affected people (e.g., heart diseases and diabetes; Bray, 2004; Williams et al., 2015) but it also effects the society as a whole (e.g., financial costs of 8 billion CHF in Switzerland alone, see Schneider & Venetz, 2014). Beside the so-called overweight crisis another crisis is ongoing – the climate crisis. The climate is changing as greenhouse gas emissions are rising and there does not seem to be an end in sight (for an overview see Schellenhuber, 2006). As a consequence, the temperature (Schellenhuber, 2006) and thus the sea levels are rising (Bindoff et al., 2007), and the number of extreme weather events is increasing (National Academies of Sciences, Engineering, and Medicine, 2016). And it is not as if we are not aware of these crises: we know we eat too much, and we know we use too many resources. But why does our behavior not change accordingly? If we are rational human beings – as classical economic theories assume (see, for example, Von Neumann & Morgenstern, 1944) – why do we act so irrationally?

When we look at ourselves there sometimes appears to be two sides to us: one that knows things (we only occasionally act on this knowledge) and another side that generally steers our actions but we do not really know the basis on which we act (we call it gut feeling). The one side knows that the salad is healthier and better for the climate than the burger, but the other side chooses the tasty burger anyway. This real-life example is underpinned by research on decision making which shows that people rely on two cognitive systems (dual process model, see Kahneman, 2011). System 1 is our automatic system, its use is effortless, it works fast, and is unconscious and associative. But there is also System 2, our reflective system. The second system takes effort to use, it is slower, but it is self-aware and deductive, and thereby essential for rational decision making. Because it takes effort to use System 2, System 1 is used for most of our everyday decisions. System 2 should therefore monitor System 1 for tasks that the latter system cannot solve successfully; for example, when you have to decide how to answer a tricky question in a job interview. Therefore, for our everyday decisions (e.g., what to eat for lunch) we do not rationally weight all attributes of different options, calculate the expected utility (or value) and choose the appropriate option, as classical economic theories suggest (e.g., expected utility

theory and its axiomatic characterizations; see Von Neumann & Morgenstern, 1944).[1] Instead, we often have just limited information and make intuitive decisions using heuristics (Gigerenzer, 2015; Gilovich et al., 2002). For example, in the food domain the heuristic "behave like the others" guides us to eat more when we are eating in a large group of friends or the "eat till the plate is empty" heuristic guides us to eat more when we have larger plates and portions (compare with Wansink, 2007). Therefore, the actual reasons why we eat as much as we do are often unclear to us (Wansink, 2007). Since System 1 is mainly activated for those decisions, our behavior is seldom influenced by information campaigns that, for example, simply inform people about the consequences of unhealthy eating behavior (see Hawkes, 2013). The problem does not seem to lie in the processing of information, as information campaigns, for example, have proved effective in changing people's knowledge and attitudes (see, for example, Boles et al., 2014), but the changed attitude does not translate into actual behavior (see Hawkes, 2013). Research could show that in many fields (including healthy food and sustainable choices) there is often a gap between people's knowledge, attitude or even intention and their actual behavior (see Sheeran & Webb, 2016). So, how can we bridge this gap?

In recent years as an alternative to classical information campaigns for changing behavior, a new type of interventions has been established which changes the environment when choices are made (i.e., the choice architecture). They are subsumed under the so-called nudge approach (see Thaler & Sunstein, 2009). A nudge is therefore an intervention that uses so-called "irrelevant factors" of the choice architecture to change behavior in a certain direction (e.g., help people to choose the healthier food option) for the benefit of the individual and society overall (this definition is based on Hansen, 2016; Hansen & Jespersen, 2013; Hausman & Welch, 2010; Thaler & Sunstein, 2009). Irrelevant factors are therefore described as elements of the choice architecture that should – according to classical economic theories – not affect choice, such as the position of a product on a shelf, where the default is set and how a decision is framed (e.g., if something is presented as a loss or gain). Thus, the use of large financial incentives or making options in another way more costly in terms of time, trouble or social sanction and so forth (e.g., install a fast food tax), as well as banning options (e.g., banning fast food from menus), would not be a nudge (see Hausman & Welch, 2010). Nudges therefore play on our System 1 and change the environment so that it is easier to behave relatively effortlessly in a desirable way, for

[1] Note, the dual-process model does not assume that the two systems map to two distinct brain systems but to different types of process (see Stanovich & Toplak, 2012). Furthermore, the key implication of the theory is not that there are exactly two systems but that there is not just one rationale as classic economic theory suggests (see Gilbert, 1999). Therefore, the two systems are fictional constructs to make the processes more understandable (Kahneman, 2011) and appear to be continuous rather than dichotomous (Norman, 2010).

example eat healthier or use fewer natural resources.[2] In recent years, studies have also shown the effectiveness of nudges in various fields and settings outside the food and use of natural resources context (for an overview see OECD, 2017a).[3] For example, switching the default of organ donation towards donation has led to an increase in organ donors and saved lives (Johnson & Goldstein, 2003), placing warning signs with social norms at a park entrance has been shown to reduce the stealing of petrified wood in a park (Cialdini et al., 2006) and placing labels with information about the healthiness of a food item has been shown to be an effective nudge to foster healthy eating (Cecchini & Warin, 2016). These few examples of nudges already show that a variety of mechanisms (i.e., defaults, social norms, simplify information) have been applied in different fields such as medicine, environmental protection, and food. These examples make it evident that "nudge" is an umbrella term and that further classification is necessary to understand the mechanisms that are effective in changing behavior in certain situations and choosing appropriate interventions.

Therefore, different categorization systems for nudges have been established. For example, the EAST framework developed by the UK government's Behavioral Insights Team – also called the "Nudge Unit" (BIT, 2014; see also Halpern, 2015) – suggests that nudges should be *Easy* (reduce friction for desirable behavior, for example set a default), *Attract* (attract attention, for example highlight a desirable menu option), *Social* (use social norms, for example, indicate that most people behave in a desirable way) and/or *Timely* (consider the timing of interventions, for example time reminders for payments). Nudges in the EAST framework can thus be used in an isolated or combined way. The EAST framework, similar to other frameworks (e.g., Mindspace by Vlaev et al., 2016; 4Ps by Chance et al., 2014; Nudge taxonomy by Ly et al., 2013), clusters nudges based on their practice-based similarities but misses a more general theory about the processes behind the nudges. However, nudges can also be categorized based on their targeted processing mechanism which helps to understand the theoretical background to nudges.

One such approach is the framework postulated by Hansen and Jespersen (2013). They categorize nudges into four groups based on the consciousness of the processing (i.e., conscious or unconscious process – they call them type 1 or type 2 nudges) and the transparency of the nudges for people (i.e., visibility as a nudge). Conscious nudges are nudges that involve some aspects of conscious processing, whereas unconscious nudges are solely driven by automatic processes. Note, although the conscious nudges target a feature that is actually associated with

[2] Note, the nudge approach is not without critics. This is addressed in the discussion part of the dissertation as well as in the discussion of the full-length manuscripts.
[3] Bernartzi et al. (2017) could show that nudges are even more cost-effective than traditional policies (such as monetary incentives).

System 2 processing (i.e., consciousness), other features (like effort or speed) are mainly System 1 driven: automatic and fast through the use of heuristics. It is assumed that the features (e.g., conscious, effort, speed) of the two systems are just loosely correlated (see Pfister et al., 2017). Furthermore, conscious nudges always involve automatic System 1 processes, but unconscious nudges are not processed by System 2.[4] The other dimensions of Hansen and Jespersen's (2013) categorization distinguishes between transparent nudges, which are visible and easy to monitor, and non-transparent nudges, that is, nudges which usually remain invisible to people. The two dimensions, consciousness and the transparency of the framework, can be arranged to form a two-dimensional matrix (see Figure 1). Thus, the two dimensions are understood as continuous rather than dichotomous (i.e., more or less transparent and consciously processed).

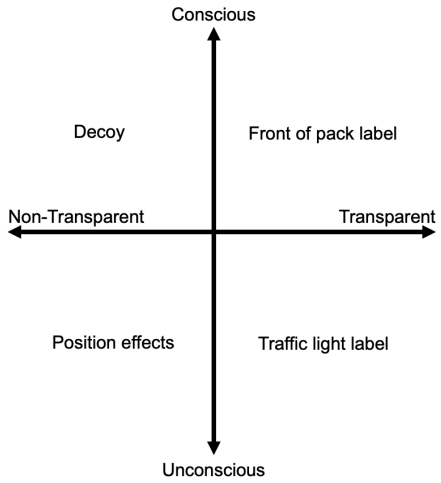


Figure 1. The two-dimensional matrix to categorize nudges (based on Hansen & Jespersen, 2013).

An example of a conscious and transparent nudge would be the use of a front-of-pack label indicating the number of calories a product contains. This nudge needs to be consciously processed by consumers and the goal of such labels is clearly visible; to help people make healthier food choices. On the other hand, highlighting this calorie label using a red (as in a traffic light label where red implies stop or bad) background would be an example of an unconscious but transparent nudge; the color red is processed automatically (i.e., draws attention and associations with the color red such as stop) and although this effect is at first unconscious (i.e., people will not report the association with stop), the intention of the nudge may be recognized (is transparent) by the person (i.e., reduce unhealthy product choices). An example of a conscious and non-transparent nudge would be when there is a choice of two products (an

[4] For example, when a person has to estimate the number of inhabitants of a city and an anchor is presented (e.g., more or less than 2,000 inhabitants?), the environmental cue is first processed by System 1 as a reference point. Even later, when conscious processes start, the anchor influences the conscious (i.e., System 2) processes (see Cartwright, 2018; and Hansen & Jespersen, 2013).

unhealthy and cheap vs. a healthy and expensive product) and an irrelevant third alternative (a so-called decoy – as healthy as the second product but even more pricy) is added, which makes people pick the middle option. The additional third product is therefore consciously processed and weighted (e.g., people could state that the third product was equally healthy and more expensive) but is not usually monitored as an attempt to persuade people to choose the middle option. The fourth group of nudges are the unconscious and non-transparent nudges. An example of such a nudge would be the change of the order of a menu at a restaurant in such a way that people choose the healthy option more often (e.g., healthy entrées in the middle; see Manuscript 2). People are not usually able to recall the exact position (e.g., was it the third, fifth or seventh option on the menu?) of the dish they choose as the position is processed unconsciously and, furthermore, they are usually unaware of ordering effects (i.e., the nudge is not transparent).

Overall, the framework by Hansen and Jespersen (2013) can help to understand the mechanism of nudges and develop further theories in the nudging approach. This is important, considering that today there is no consistent theory or conceptual framework that guides a choice architect (i.e., the person that plans the nudge intervention, for example a politician or restaurant owner) towards a successful nudging intervention in a given context. Although, there have been attempts to develop a consistent theory on some nudges (e.g., framing effects, see Maule & Villejoubert, 2007), for many nudges there is a hotchpotch of theories or effects that are completely contradictory. For example, products (e.g., on a shelf, on a website or on a buffet) are brought into an order and, by placing a product in a preferred spot, one may increase the sale of the product. In the literature, a variety of such position effects have been reported; there is the top-choice effect (or primacy effect; the item on top or the first one is picked most often), the center-stage effect (the item in the middle is picked most often) and the recency effect (the last item is picked most often). However, there is no consistent theory that predicts under which circumstances which effect will occur (although in Manuscript 2 we tried to establish such a framework for restaurant menus). So, when there is no consistent theory, how does one plan a nudge intervention?

Today, when planning a nudge intervention, choice architects use existing interventions that have been proven to be effective in similar situations and then adapt them to the specific context. For example, the EAST framework, which describes examples of such effective nudges (see Halpern, 2015), can be used here. Usually, interventions are first tested on a small scale and when proven to be effective, the interventions are scaled up and their effect tested. Thus, even scientists with years of experience have difficulty in predicting the outcome of studies and are sometimes surprised by it (see Tripp, 2017). Nudge interventions often simply do not work as expected.

So, we have on the one hand evidence of the effectiveness of real-world nudge interventions by the practitioners. On the other hand, however, we have a lack of theoretical models that can guide the practitioners (i.e., choice architects) from the general theory (e.g., dual-process model) to planning an intervention. There is no model that describes how an intervention has to be designed in specific contexts, for example which variables have to be considered and which nudge could be applicable. Therefore, this field, where the quest for understanding (i.e., the general theory) and the consideration of use (i.e., the real-world interventions) meet, the so-called use-inspired basic research (see Pasteur's quadrant in Stokes, 1997), is where this dissertation is positioned. Use-inspired basic research thus tries to develop more general principles that have relevance for a variety of problems but also investigates under which circumstances the principles are valid (see Rogers et al., 2007).[5]

Therefore, the goal of this dissertation is twofold: first to understand how choices are affected by environmental cues in specific settings, and second, to develop frameworks for circumstances in which a specific effect more generally occurs. Such frameworks should then help to plan future interventions. Therefore, this dissertation focusses on two of the main problems of our society, overweight and climate change. In those two fields behavior change is urgent and well-designed nudging interventions could have a major impact. That nudging can be effective in those fields has been shown by previous studies, for example meta-analysis in the food domain could show that nudges on average increase healthier dietary or nutritional choices by 15.3% (Arno & Thomas, 2016). Thus, different types of nudges have been shown to be more or less effective (Cadario & Chandon, 2019b). To study how choices are affected by environmental cues we used different nudges in different settings.[6] We accordingly chose nudges that we estimated to have a large impact in the specific situation based on existing literature. The nudges used in our studies varied in their transparency and degree of conscious processing.

Manuscript 1 builds on the findings of basic research about decision frames (i.e., whether you choose or reject an item) and transfers it to the food context. Using traffic light labels – often used to label foods in supermarkets – as stimulus material, we investigate how different aspects (sugar, energy content, etc.) of the product interact with the decision frame (i.e., whether you

[5] Rogers et al. (2007) use this definition for applied cognitive research but they also state that, in their opinion, applied psychology "should be thought of as use inspired basic research" (p. 4). They thus mix up applied and use inspired basic research, which are actually two different things as defined by Stokes (1997).

[6] Note, not every nudge is applicable to or effective in every setting; it may even be that a nudge works in one direction in a setting and in the opposite direction in another setting. For example, adding a small gift (e.g., a CD) to a product increases sales of this product (Darke & Chung, 2005) but a small gift (e.g., a pen) to encourage a charity donation reduces donations (see Newman & Shen, 2012).

choose or reject one of two food items) to influence your decision. Furthermore, different theories about decision frames are tested and real-world implications are discussed.

Manuscript 2 attempts to bring order into the already mentioned hotchpotch of position effects regarding restaurant menu choices. Therefore, moderator variables previously ignored in the food literature, such as preference, were identified and tested in combination with position effects. Based on the results, a framework for position effects was established that should help to guide choice architects to plan future interventions.

Manuscript 3 studies how defaults influence the choice of green electricity products. We accordingly used – compared to earlier studies – an enlarged choice set and compared our findings to the findings of those earlier studies. Theoretical and practical implications are discussed and the results could help electricity providers to design their product layout.

Overall, we use a variety of nudges (more and less conscious and transparent ones) in two different fields (food and electricity) in the manuscripts. The central theme in all manuscripts is the desire to contribute to a more general understanding and theory of the nudge approach. We start with a well-known example, the influence of decision frames.

Summary of the Manuscripts

Manuscript 1: Choosing or Rejecting a Food Item, does Framing Matter? And what has sugar to do with it!

To target overweight and obesity, governments in different countries have launched so-called front-of-pack labels to inform people about the calories and/or the nutrients that a product contains (OECD, 2017b). Unlike the nutrition facts usually placed on the back of an item, these front-of-pack labels are more conspicuous and usually only provide information about a selection of nutrients and the calorie content. One of the most common labels is the “traffic light label” (TLL) introduced by the Food Standards Agency (2016). The TLL is placed on food items and indicates the amount of fat, saturates, sugar and salt a product contains in numbers but also in an appropriate color code (the label rates the nutrients based on the recommendation of daily intake, for example low = green). The TLL should thus simplify the information and help consumers pick the healthy food options. Research could show that labels lead consumers to buy healthier items (see Franckle et al., 2018), but so far little is known about how consumers compare and weight the different values of the nutrients on the label. The study by Hieke and Wilczynski

(2012) indicates that some nutrients are perceived as more important (e.g., sugar) than others (e.g., salt). So, do consumers perceive a product that has 2 red and 2 green nutrient lights (e.g., high in fat and saturates but low in sugar and salt) as more, less, or equal healthy as a product that has an all-orange (moderate values in all nutrients) label?

Research outside the food domain, for example the study by Shafir (1993), indicates that people are looking for stimulus features that are compatible with the task at hand (compatibility hypothesis). This would mean that when consumers *choose* between an item that has an all-orange TLL and an item with a 2 red/2 green TLL, they would choose the 2 red/2 green item more often due to the positive attributes (2 green lights) of the item. On the other hand, when consumers do not choose between two items but *reject* one of them, they focus more on the negative aspects of the items (e.g., looking which has more red lights and then reject it). This would lead to irrational behavior where participants choose the 2 red/2 green item more often but also more often reject the 2 red/2 green item. So, how do these decision frames in combination with specific values of some nutrients (e.g., sugar) influence the decision to choose one of the two products?

Study 1. In the first study we investigated 580 participants in an online survey where they were confronted with two cereal bars,[7] in respect of their TLLs. One of the two options had an all-orange TLL and the other option a 2 red/2 green TLL. To check for the importance of different nutrients we created different versions of the 2 red/2 green label whereby the mean of the recommended reference intake was kept constant. An example of the decision task is displayed in Figure 2. Each participant saw one choice task.

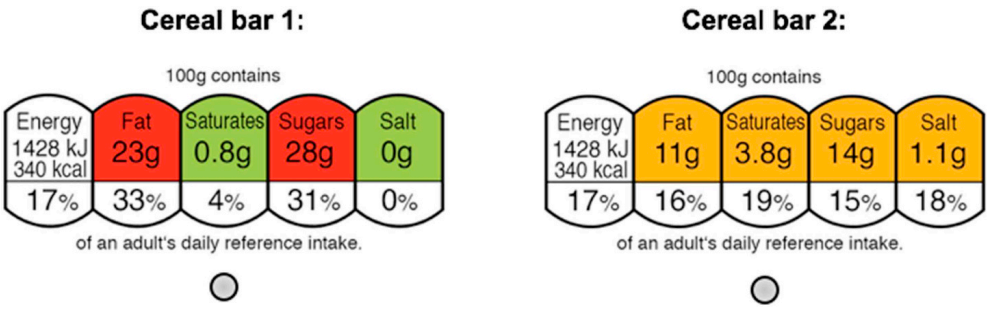


Figure 2. Participants' choice situation (left: 2 red/2 green label; right: all-orange label).

[7] Cereal bars were selected owing to their large variety of nutrient profiles.

The results of the binary logistic regression[8] showed that participants went home more often with a 2 red/2 green bar when sugar was green than when it was red. This indicates that participants gave more weight to sugar than other nutrients. Furthermore, this difference was stronger in the choose condition than in the reject condition (interaction of sugar with the decision frame). When sugar was green, we found results that are in line with the compatibility hypothesis (Shafir, 1993), but not when sugar was red. When sugar was red participants went home less often with the 2 red/2 green bar than with the all-orange bar. This tendency was stronger in the choose than in the reject condition. Similar results outside the food context have been found by other authors (Chen & Proctor, 2017; Colombo et al., 2002; Wedell, 1997) who claim that the difference in the choose/reject frame cannot be explained by compatibility with the task but by the accentuation of the overall preference of one option over the other (accentuation hypothesis). The accentuation hypothesis thereby assumes that consumers are more discriminating under a choose frame than under a reject frame.

Study 2. In the second study, we tried to replicate the results of the first study and, furthermore, test the accentuation hypothesis. The accentuation hypothesis would assume that when the 2 red/2 green bar would be mostly preferred (e.g., when sugar is green), consumers would go home with the 2 red/2 green bar more often under a choose frame than under a reject frame. The opposite would be true when the all-orange bar is preferred overall over the 2 red/2 green bar (e.g., when sugar is red). We therefore tested 520 participants using similar stimulus material and a similar procedure to the first study. The differences were that participants had to make decisions among three products (cereal bars, chips and cereals)[9] and we varied the energy content of the products. The calculated regression slopes[10] indicated that when the item with the 2 red/2 green label is most preferred, participants go home more often with it when the decision is framed as a choice rather than a rejection task. However, if the item with the 2 red/2 green label is not the preferred one, more consumers go home with it when they had to reject one of the items. As can be seen in Figure 3, sugar discriminates if the 2 red/2 green bar or the all-orange bar is preferred overall and calories seem to have a minor influence on choice. Our results are thereby in line with the accentuation hypothesis. Chen and Proctor (2017) explain the occurrence of this effect as being a result of the greater cognitive effort needed in a rejection task than in a

[8] The binary logistic regression showed a main effect for sugar (red vs. green), $\chi^2(1, N = 580) = 45.10$, $p < .001$ and a Sugar x Condition interaction; $\chi^2(1, N = 580) = 8.18$, $p < .01$, but no main effect for condition; $\chi^2(1, N = 580) = 0.77$, $p = .38$.

[9] We checked whether the decision pattern differed between the three food products. It was found that the model with the factor product did not add significant value (χ^2 [26] = 35.61, $p = .10$) and was therefore omitted.

[10] The slopes for the choose condition; $F(1, 6) = 117.9$, $p < .001$, $R^2 = 0.97$, and the reject condition; $F(1, 6) = 58.92$, $p < .001$, R^2=0.91, differed; $t(7) = -4.016$, $p < .01$.

choose task. Hence, the consumers in a rejection frame have fewer resources left for the decision itself and are therefore less discriminative between the options (i.e., the mean preferences for one of the options are less strong).

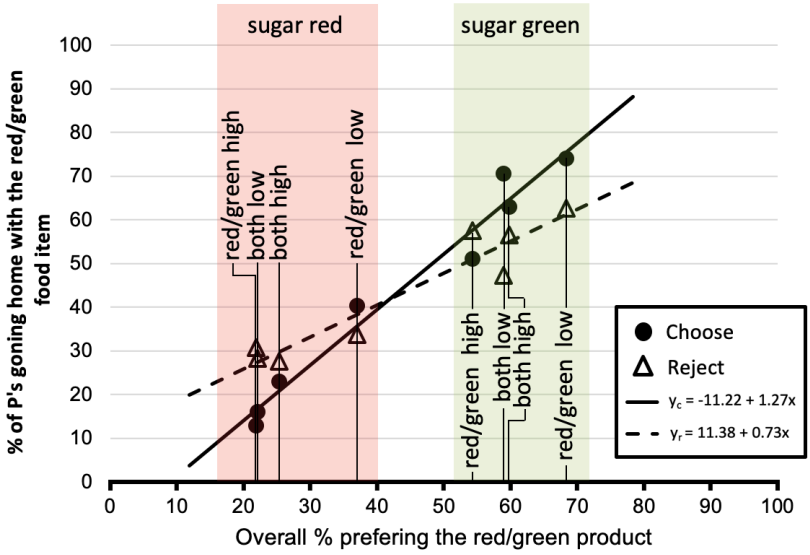


Figure 3. Results of the two linear regression slopes indicating a match with the accentuation hypothesis. Vertical text in the figure labels the calorie content of the corresponding dot and triangle.

To sum up, the results of the studies indicate that the decision frame (choose or reject) has an influence on choice and that the accentuation hypothesis can explain the difference accurately. These findings could therefore help to understand the framing effect and influence what we eat.

Manuscript 2: Nudged to a Menu Position: The Role of "I'm Loving It"!

What we eat, that is, what we choose from a supermarket shelf, the restaurant buffet or menu, appears to be influenced by the position of the items. And although research indicates that those positions influence our choices, the literature describes a potpourri of different effects. For example, a center-stage effect (preference of the middle option) was found by Christenfeld (1995), Keller et al. (2015) and Rozin et al. (2011), but this effect could not be found by others (Dayan & Bar-Hillel, 2011; Rozin et al., 2011; Wansink & Hanks, 2013) who instead found a top-choice – also called primacy – or a recency effect (preference for the first and last option respectively). So, it would appear that consumers sometimes prefer the middle but also sometimes the first or last option. Therefore, to develop a theoretical framework one has to find the factors that can explain the different position effects found in the literature. We identified two variables that we assume influence position effects: identity and preference. Identity refers to the perception of the options in a set as either identical (e.g., a set of canned tomatoes) or non-

identical (e.g., a set of cereal bars with different ingredients) and preference refers to a situation where you "love" an item (or not). For example, studies (Christenfeld, 1995; Valenzuela & Raghubir, 2009) that used identical items reveal the tendency to choose items placed in the middle. This is probably due to the belief of consumers that the most popular item is usually placed in the middle. Studies outside the food context, for example Rodway et al. (2016) found that when people had to choose their preferred item from a set of similar items the ones in the middle were picked more often. Therefore, we assume that preference mediates the center-stage effect, that is, that when preference plays a role (e.g., entrée of your preferred cuisine style – the reason why you built a preference and where everything looks tasty) a center-stage effect should occur. But, when preference plays no role (e.g., appetizers, to which you are more indifferent), a top-choice effect should be found.

To test our hypothesis, we tested 45 participants using a computer-based restaurant menu study. Seven menus with completely different dishes for each of five different cuisine styles were designed. First, participants had to indicate which of the five cuisine styles was their most and least preferred one. Then, participants saw menus of their most and least preferred cuisine style on the screen and had to choose the option they would most likely order in a restaurant. Each menu contained the meal types of appetizers, entrées and dessert with each course being shown separately.[11] It was subsequently found that participants chose entrées of their most preferred cuisine style 24% more often when placed in the middle of the choice set compared to the top or the bottom area.[12] This center-stage effect occurred only in the preferred cuisine style, we assume because of the perception of the items as identical (e.g., "all Italian entrées look so yummy!") and thus the choice is based on the popularity belief of the middle option (see Valenzuela & Raghubir, 2009). Appetizers at the top of the menu were chosen about 21% more often independent of the preference for the cuisine style.[13] The top-choice effect is therefore probably based on the belief that items on the top are of a higher price or quality (see Valenzuela & Raghubir, 2015). Regarding desserts no position effect could be found. We assume that consumers care less about the individual items when it comes to dessert and therefore choose more randomly. This assumption is underpinned by the faster reaction times for desserts.[14] To

[11] During the experiment, the participants saw each menu three times but with switched positions, for example first the "spaghetti alla bolognese" was on top, second in the middle and then in the bottom position.
[12] Compared to the top area, χ^2 (1, N = 535) = 5.250, $p < .05$, or the bottom area; χ^2 (1, N = 535) = 7.308, $p < .01$.
[13] Participants chose more appetizers from the top area than the bottom area (most preferred: χ^2 [1, N = 496] = 5.879, $p < .05$; least preferred: χ^2 [1, N = 522] = 5.180, $p < .05$) and in the least preferred condition more from the top than the middle area; χ^2 (1, N = 517) = 6.284, $p < .05$.
[14] Reaction times for appetizer (M = 12,289 ms, SD = 7,909 ms) and desserts (M = 9,293 ms, SD = 5,836 ms) differed significantly; t(2984) = 11.782, $p < .001$.

sum up the results, a framework as a "guideline for choice architects" was created to help choice architects to decide where to put healthy items (see Figure 4).[15]

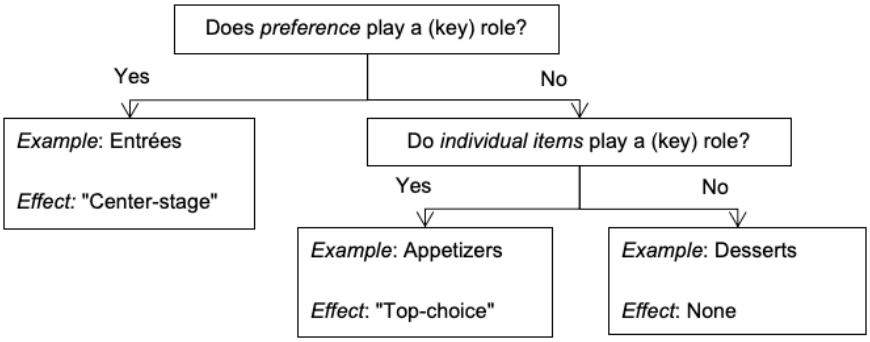


Figure 4. Theoretical framework for menu choices as a guideline for choice architects[16]

One further aspect that could be shown by this study is that some consumers have preferences for one specific dish (e.g., spaghetti alla bolognese) and chose that item whenever it showed up independently from its position (on average in 16% of the menus such repeated choices occurred). Parker (2017) describes similar results from a study where one-sixth of the workers reported having eaten the same lunch (e.g., ham sandwich) every day for six years. Thus, the lunch seems to become a default that is seldom questioned. Developing interventions that switch the default to a healthy meal could thus be shown to be effective.

Manuscript 3: When Too Few Is Bad for the Environment. Choice Set Size and Default Effects for Electricity Products

Like the people who consume the same menu every day (see Parker, 2017), we consume electricity every day without thinking too much about our consumption (i.e., the origin of our electricity). Although the electricity market has been liberalized in many countries, most people stay with their energy provider and their default electricity product (Pichert & Katsikopoulos, 2008). This is important especially in regard to the ongoing climate crisis: people do not actively switch to renewable electricity products. Therefore, the tendency of people to stay with the default has been reported in many fields, from organ donation (i.e., donate organs as default, Johnson & Goldstein, 2003) to the printing option of an institution (i.e., printing double-sided as

[15] Note, a limitation of the study is that we did not manipulate the healthiness of the dishes. So, future studies have to be conducted to check whether our framework actually fosters healthy food choices.
[16] Note, our research indicates that those effects in general can be found, but also that the factor time (i.e., how often one picks from the menu or – more applied – if someone is a regular customer) plays a role. For more details see the attached full paper.

default, see Rutgers, 2020), and has proven to be an effective nudging technique. In the energy sector, Pichert and Katsikopoulos (2008; similarly, Ebeling & Lotz, 2015) found, depending on the electricity product set as a default – that is, a more environmentally friendly "green" or less environmentally friendly "grey" product – people chose the respective option more or less often. Thereby, compared to a neutral condition (without a default), setting a default could nudge people to stay more often with the grey product. On the other hand, setting the default to the green product could not nudge more people to stay in the green product than if they had to actively choose it. Therefore, both studies (Ebeling & Lotz, 2015; Pichert & Katsikopolus, 2008) used binary choice sets. However, research outside the electricity context (Simonson & Tversky, 1992) has shown that adding a third option (e.g., an eco product; high in price and environmental friendliness) to a binary choice set (e.g., a grey electricity product; low in price and environmental friendliness, and a green electricity product; medium in price and environmental friendliness) can change the overall choice pattern crucially and either lead to a *compromise effect* or a *polarization effect* (e.g., Simonson & Tversky, 1992). According to the compromise effect, by adding this third option the middle option (e.g., green product) becomes relatively more popular. Therefore, consumers have to consider both attributes of the options (e.g., price and environmental friendliness) as equally important. When consumers consider the advantages of one attribute (e.g., environmental friendliness) but perceive the other's advantages as negligible (e.g., price), the addition of a third option leads instead to a polarization effect. Accordingly, the middle option becomes relatively more popular by adding the option high in the considered attribute (e.g., eco product) to the choice set (e.g., with the other two products) but not when adding the option with the neglected attribute (e.g., grey electricity) to the choice set (e.g., to a green and eco electricity choice set). Nevertheless, so far it is unclear whether adding an eco electricity product would indeed increase or decrease the relative share of the middle product.

In an online experiment, we therefore tested 241 participants, confronting them with three electricity products: grey, green and eco. The products differed in their environmental friendliness (low to high) and their price (low to high). Participants were randomly assigned to one of four conditions: grey default, green default, eco default or no default. In the no-default condition participants had no default set and had to actively choose their product. In the default conditions, participants were logged into the respective product and were asked if they would like to switch to either of the other products and how much they would be willing to pay for their default (or actively chosen product) and which switch costs/gains they would expect for each of the other products. Thereby, our results show that in the no-default condition, most participants chose the green product (56%) followed by the eco (29%) and the grey option (15%). Under the

default conditions, we found a default effect for all products,[17] that is, participants stayed in their defaults more often (grey default; 48%, green default; 76%, eco default; 61%). This means, unlike previous studies (e.g., Pichert & Katsikopoulos, 2008), we found positive default effects, that is, defaults did not work by keeping participants in the grey electricity product but instead nudged them into green (additionally 20%) or eco energy (additionally 33%) products (see Figure 5).

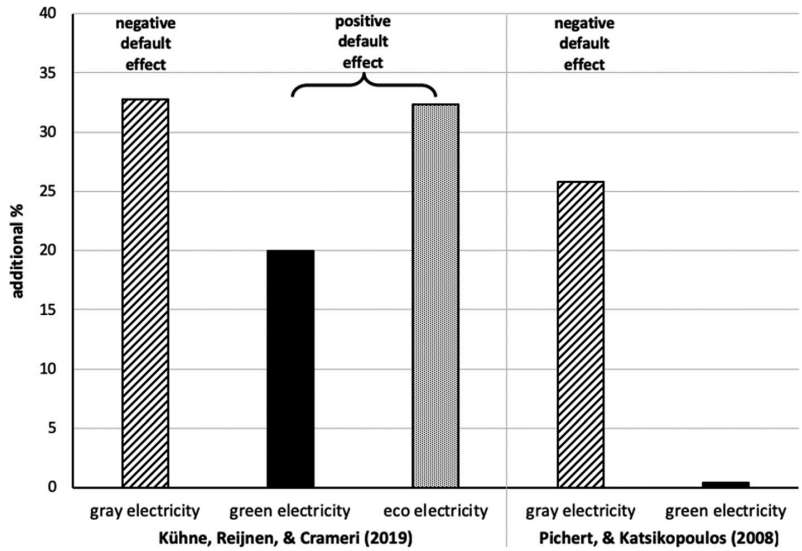


Figure 5. The bars indicate the additional percentage of choices made in favor of the default option relative to its control, the no-default condition.

We furthermore compared our three-option choice set with the binary choice set of Pichert and Katsikopoulos (2008) and found that adding an eco product tends to increase environmentally friendly choices.[18] Accordingly, the pattern of the results suggests that the tendency for an increased popularity of the middle option can be ascribed to the polarization effect. This assumption is underpinned by the stated switch costs/gains, which indicate that participants weighted the dimension of environmental friendliness higher than the dimension of price.[19] Participants expected prices for the electricity products (reference and switching prices)

[17] To test for default effects, we used generalized linear models for Poisson distributions and compared each default position with its equivalent in the no-default condition. For all three comparisons (see left side of figure 5) we found default effects (grey: $b = 1.65$, $SE = 0.34$, $p < .001$, $V = 0.18$; green: b = 0.91, $SE = 0.31$, $p = .004$, $V = 0.11$; eco: $b = 1.36$, $SE = 0.31$, $p < .001$, $V = 0.17$).

[18] Note, this is just a tendency and is not significant. The popularity of the green option in the green default condition increased from 54% to 62% ($z = 1.14$, $p = .255$) and in the no-default condition from 67% to 79% ($z = 1.67$, $p = .096$).

[19] Regarding the prices for the electricity products, we found that the expected gains for switching to a less environmental friendly product were higher than the expected costs for a switch to a more environmental friendly product in the no-default condition ($t(56) = 4.63$; $p < .001$, $g = 1.01$).

to be quite unrealistic (a finding also observed by Tabi et al., 2014). Despite those high reference and switch prices, however, participants did choose environmentally friendly options. We conclude that enlarging the choice set with an eco electricity product and placing the default judiciously could be an efficient way to foster sustainable electricity consumption and to do more for the environment.

General Discussion and Conclusion

In the three manuscripts the effectiveness of the nudging approach could be demonstrated in different fields. The first manuscript showed that when consumers make a decision between two food options (one option with all average attributes and one with bad and good attributes), the frame of the decision is a predictor of the decision but is moderated by the consumers' overall preference (where sugar seems to play a major role) for one of the options. Thus, the accentuation hypothesis could predict the observed decision pattern. The second manuscript showed the importance of preference and identity when it comes to position effects in restaurant menus. A framework was designed to guide choice architects to the sweet spots on the restaurant menu. Both manuscripts thereby reveal insights into how one could nudge people into healthier eating habits. The third manuscript shows that judiciously set defaults and enlarged choice sets with a more environmentally friendly electricity product can help to engage people in the energy transition. Therefore the manuscripts show that nudging can be used in different fields (food choices and sustainable electricity choices) and with different mechanisms and effects (frames, positions and defaults).

Overall, nudging seems to be a powerful tool. But the manuscripts also illustrate the complexity of the nudging approach and behavioral change generally: identifying the relevant factors for behavior change and understanding how they interact in a specific situation is difficult. My co-authors and I therefore tried to find the patterns described in different studies and contexts, establishing a possible explanation and testing it. The established framework and theoretical contributions should help to plan future studies and interventions, as well as to assist in building up a theory on the nudging approach. In contrast, today, many of the effects reported in the nudging literature are the so-called one-word explanations (e.g., center-stage, top-choice effect etc.) and many lack a theoretical background. Note, not all one-word explanations have a lack of theoretical foundation (e.g., the concept of loss-aversion has a theoretical foundation), but if there is a lack, they can be used to explain almost every outcome (see Gigerenzer, 1998). Thus Gigerenzer (1998) notes that those explanations were loosely characterized in the 1970s as ad hoc

explanations. Even now, 50 years later, there is still no compelling theory to explain many of these effects. Those one-word explanations, without a theoretical background, are thus still used, especially in applied research to explain outputs (as, for example, in the food domain; Wansink, 2007). This lack of theory has its consequences: when just focusing on the use of effects, there is little scientific progress regarding the understanding (i.e., explanation and integration) of a phenomenon (as defined by Feller & Stern, 2007). And, connected to this, the outcome of an intervention (the application of a phenomenon) is less predictable. That this is the case regarding nudging is clearly underlined by David Halpern (2019b), chief executive of the BIT, who states that only one in four interventions or even fewer are successful (actually, we do not know the real numbers because institutions seldom make all their attempted interventions public). This indicates that many ideas and interventions really do not work. And that is exactly what Halpern (2019a) states: "Lots of good ideas won't work and that's ok." Of course, behavioral economics teams such as the BIT have shown that they are highly cost-efficient (see Halpern, 2015). However, one could try to reduce the number of interventions required to find an effective one by establishing a more reliable theoretical background to the nudge approach. Furthermore, we would better understand what we are doing. Studies in use-inspired basic research can help to categorize different effects, such as when does a top-choice and center-stage affect occur, before launching it in an applied setting.[20] Choice architects could then plan, for example, a restaurant menu intervention based on a framework or theory and check whether a top-choice effect occurs for entrées and whether healthy items would be picked more often from that position. Thereby, one would not need to test for all kinds of ideas based on the one-word explanations but arrive at an effective intervention more quickly. Furthermore, today one-word explanations can always be attached to explain the outcome post hoc. Without a theoretical background those explanations are not falsifiable and thus do not foster scientific progress. Here this dissertation tries to differ; by establishing clear, falsifiable theoretical frameworks. Accordingly, our studies are of course limited: we used convenience samples (mostly students from different faculties) and choices with no consequences (e.g., financially) and the choices are made in a fairly abstract manner (e.g., food choices based on labels). So, if future studies can verify the established theoretical frameworks, they may help to understand the nature of some of the one-word explanations and assist in building a more appropriate theory in the nudge domain.

The nudge domain – and thus indirectly the nudges used in our studies – has been criticized for reducing the freedom of choice. For example, critics state that nudges are

[20] Of course, even interventions based on a solid theoretical background cannot replace testing interventions first in smaller control trials before scaling up, but less failed trials could be expected.

paternalistic, because people often do not know (transparency) that their decision is influenced and thus their freedom of choice is restricted (see, for example, Rebonato, 2014)[21]. On the other hand, proponents (e.g., Thaler & Sunstein, 2009) of the nudge approach point out that the freedom of choice is still contained – people can always choose what they want, no option is forbidden or highly incentivized. What on first sight might not be clear; those two camps differ fundamentally in their understanding of freedom of choice per se and this has barely been addressed in the discussion so far. Therefore, looking at the literature on different definitions of freedom of choice[22] may help: for example, Pattanaik and Xu (1990) established axioms which lead to a (in their own words) rather naïve rule for judging freedom of choice. In short, they state that the higher the number of options the more freedom of choice.

Another definition has been proposed by MacCallum (1967) who states that one may define a person as free (in choice) as long as he or she lacks constraints to perform certain actions. Regarding those first two definitions, nudges (including non-transparent nudges) cannot be considered as reducing the freedom of choice as they do not forbid any option (i.e., do not reduce the number of options in a choice set) or make some of them significantly more or less costly with regard to money and time and so on (i.e., do not use constrains). The first two definitions of freedom of choice thus look at the options that a person *has,* but one may also look at the choices a person actually *makes* (see Carter, 2004). Therefore, freedom of choice may be defined by the degree that the choice one actually makes is influenced by environmental factors. Following this third definition, nudges would limit the freedom of choice, as they influence the choice one makes (e.g., make people choose the healthy product more often). This is, by definition, the goal of a nudge (but that is also the goal of every political intervention). Nevertheless, with regard to nudging and the third definition one could also argue that every choice set has an order and thus people are always influenced. For example, the top-choice effect (as an example of a non-transparent nudge) influences our choice even when not intentionally used. When not using, for example, the top-choice nudge *intentionally,* the choices are more randomly distributed (e.g., sometimes a healthy and sometimes an unhealthy dish at the top of the

[21] An alternative of the nudge approach is thereby the use of boosts (see Hertwig & Grüne-Yanoff, 2017). Boosts are educative programs which should foster competence in a specific field. More about the discussion between boosts and nudges is stated in the discussion part of the full-length manuscript 3.

[22] One basic question is therefore, what do we define as a choice? (see Hansen & Jespersen, 2013). Do I choose to let the plate fall when it is too hot (reflex)? Do I choose every strike with my pedal when driving a bike (habit)? Do I choose when I do not change my mobile phone contract even though there are better options available (default)? Often choices are defined as a decision between two options involving reflective thinking, but then habitual choices could hardly be defined as choices at all, as they do not involve reflective thinking (see Pfister et al., 2017). Therefore, unconscious nudges cannot be considered as limiting freedom of choice, as they do not involve a choice in this sense. I use the word "choice" here in a broader sense as any choice between options (reflective or not).

menu and thus chosen more often) but that does not mean that one is less influenced. Nudging uses those effects *intentionally* but that does not change freedom of choice per se. In short, there is no neutral design.

It may seem that some choice architectural designs are more neutral than others, for example when making people choose an option actively instead of setting a default (as an example of a rather transparent nudge). But, as Sunstein (2014) points out, demanding an active choice also reduces freedom of choice. This is in regard to the choice a person *has* (first and second definition of freedom of choice), as it reduces the choice set, meaning one cannot choose not to choose anymore. A default thus increases the freedom of choice a person *has* but decreases it regarding the choice a person *makes* (third definition). So, it becomes clearer that the definitions and attributes of freedom of choice may require a more complex model than just a dichotomous distinction as free or not free. One could instead think of freedom of choice as a construct with two poles on a dimension.[23] Accordingly, people seem to have a rather restricted freedom of choice per se (similar to the free-will debate, see Walter, 2016). The intuition that we normally make free choices and only when "a nudge comes along" our choice is abruptly limited is wrong. Nudges either use existing factors (e.g., top-choice effect) or add environmental cues (e.g., setting a default)[24] in an already biased setting and thereby merely adjust the leverage (or probability) to act in a certain way, whereby at least some people with clear preferences will not act accordingly (e.g., people with a nut allergy will probably not eat the peanut bar however it is framed). Nudging changes frequencies and is thereby *probabilistic* rather than *deterministic* (see Hansen & Jespersen, 2013). However, every effective forms of policy, such as taxation and regulation, change probabilities. They thereby also often limit the choice set by forbidding options and seem quite intrusive in terms of limiting freedom of choice regarding the choices people *make* (e.g., incentives), but also regarding the options people *have* (e.g., forbidding options). Freedom of choice cannot simply be defined by the transparency of an intervention. Therefore, the argument that nudging is not ethical per se because it reduces freedom of choice is not compelling, as (a) only some nudges actually change freedom of choice and (b) classical interventions also change freedom of choice (and often in a more intrusive way). So, whether nudging is ethical regarding freedom of choice can only be decided by comparing the nudge (e.g.,

[23] Or even a two-dimensional model regarding freedom of choice one has and the choices one makes.

[24] Note, only when new environmental cues are installed that did not exist before, for example by highlighting a healthy dish on a menu in such a way that it is chosen more often, could be considered as a nudge that reduces freedom of choice regarding the choice that people *make*. Such an intervention would also not increase the choice set and thus would only be a reduction of freedom of choice overall. On the other hand, most nudges cannot be clearly defined as limiting or enhancing freedom. For example, a decoy would increase the choice set (freedom of choice one *has*) but reduces the freedom of choice regarding the choice one *makes*.

setting green defaults) to the alternative method one would choose to reach that goal (e.g., forbid grey electricity).

Whether or not nudges are ethical cannot, of course, be decided only by their degree of freedom of choice. One other important aspect is whether they are accepted by the population. Thus, the acceptance of nudges compared with classical interventions in society is high (see in the food domain Hagmann et al., 2018; and acceptance of nudges overall, Sunstein et al., 2019). Therefore, the acceptance of transparent and conscious nudges (e.g., adding a health label to a product) is higher than for less transparent and unconscious nudges (e.g., reducing plate sizes at the buffet). So, one could assume, choice architects should just stay with the transparent and conscious nudges. But, as Cadario and Chandon (2019a) could show, people have high acceptance rates for the nudges they think are effective and they usually think those are the transparent and conscious nudges. However, in contrast to people's estimations, the unconscious and non-transparent nudges have been shown to be the more effective ones (see in the food domain, Cadario & Chandon, 2019b). As soon as people are informed about the effectiveness of the less transparent nudges, people's acceptance of those nudges rises and gets equal acceptance rates (Davidai & Shafir, 2018). In other words, choice architects that implement non-transparent and unconscious nudges could improve acceptance rates by communicating their effectiveness.

One aspect in the debate about effectivity and ethics of different types of nudges (and interventions in general) is often left out, however: the hedonic costs of interventions. For example, Liem et al. (2012) found that adding health information (e.g., reduced salt content) reduced expected and perceived liking of a soup, which means reduced hedonistic value of the product. Furthermore, Thunström (2019) showed that adding a calorie label (a transparent and conscious nudge) to food caused negative emotions in people with lower levels of self-control. They thereby had a harder time regulating their food choices, that is, the nudge did often not induce a behavior change (e.g., low self-control participants chose the healthier food item less often). On the other hand, people with high self-control have been shown to be more likely to regulate their behavior due to the calorie label and thereby experience positive emotions. However, so far very little research has been conducted about the hedonistic consequences of nudges and it is not clear how one could add this factor to a conclusive picture of the effectivity and ethics of nudges. So, when designing nudging interventions, a variety of factors have to be addressed, such as the direct effect of the nudge (e.g., number of people switching to the healthier food), freedom of choice (regarding the options one has and the decision one makes), acceptance (by the population), hedonistic costs (such as emotional consequences) and if or how the nudge is

communicated to the people. Thus, most of those factors depend on the transparency and the degree of conscious process of the nudge. This should make clear that whether a nudge is effective and ethical has to be concluded for each specific situation.

To sum up, our studies could show the effectiveness of the nudging approach in two different fields, healthy food choices and sustainable consumption. Therefore, different nudges with different degrees of transparency and conscious processing were used. We furthermore tried to foster a theoretical foundation of the nudging approach, designing a framework, and discussing key variables for the specific nudges. Overall, in the last ten years the nudging approach has been shown to change behavior effectively in many fields. Accordingly, major challenges for the next years seem to be the establishing of a theoretical background, as well as further investigating the possibilities and limits of the approach by addressing different aspects of its effectiveness and ethics.

References

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*, Article 676. https://doi.org/10.1186/s12889-016-3272-x

Behavioral Insight Team. (2014). *EAST: Four simple ways to apply behavioural insights*. https://www.bi.team/wp-content/uploads/2015/07/BIT-Publication-EAST_FA_WEB.pdf

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J., & Galing, S. (2017). Should governments invest more in nudging? *Psychological Science*, *28*(8), 1041-1055. https://doi.org/10.1177/0956797617702501

Bindoff, N. L., Willebrand, J., Artale, V., Cazenave, A., Gregory, J., Gulev, S., Hanawa, K., Le Quéré, C., Levitus, S., Nojiri, Y., Shum, C. K., Talley, L. D., & Unnikrishnan, A. (2007). Observations: Oceanic climate change and sea level. In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. J. Tignor, & H. L. Miller (Eds.), *Climate Change 2007: The physical science basis* (pp. 385–432). Cambridge University Press.

Bray, G. A. (2004). Medical consequences of obesity. *The Journal of Clinical Endocrinology & Metabolism, 89*(6), 2583–2589. https://doi.org/10.1210/jc.2004-0535

Boles, M., Adams, A., Gredler, A., & Manhas, S. (2014). Ability of a mass media campaign to influence knowledge, attitudes, and behaviors about sugary drinks and obesity. *Preventive Medicine*, *67*, 40–45. https://doi.org/10.1016/j.ypmed.2014.07.023

Cadario, R., & Chandon, P. (2019a). Effectiveness or consumer acceptance? Tradeoffs in selecting healthy eating nudges. *Food Policy*, 85, 1–6. https://doi.org/10.1016/j.foodpol.2019.04.002

Cadario, R., & Chandon, P. (2019b). Which healthy eating nudges work best? A meta-analysis of field experiments. *Marketing Science*. https://doi.org/10.1287/mksc.2018.1128

Carter, I. (2004). Choice, freedom, and freedom of choice. *Social Choice and Welfare*, *22*(1), 61–81. https://doi.org/10.1007/s00355-003-0277-z

Cartwright, E. (2018). *Behavioral Economics*. Routledge.

Cecchini, M., & Warin, L. (2016). Impact of food labelling systems on food choices and eating behaviours: A systematic review and meta-analysis of randomized studies. *Obesity Reviews*, *17*(3), 201–210. https://doi.org/10.1111/obr.12364

Chance, Z., Gorlin, M., & Dhar, R. (2014). Why choosing healthy foods is hard, and how to help: presenting the 4Ps framework for behavior change. *Customer Needs and Solutions*, *1*(4), 253–262. https://doi.org/10.1007/s40547-014-0025-9

Chen, J., & Proctor, R. W. (2017). Role of accentuation in the selection/rejection task framing effect. *Journal of Experimental Psychology: General*, *146*(4), 543–568. https://doi.org/10.1037/xge0000277

Christenfeld, N. (1995). Choices from identical options. *Psychological Science*, *6*(1), 50–55. https://doi.org/10.1111/j.1467-9280.1995.tb00304.x

Colombo, L., Nicotra, E., & Marino, B. (2002). Preference reversal in decision making: The attraction effect in choice and rejection. *Swiss Journal of Psychology, 61*(1), 21–33. https://doi.org/10.1024//1421-0185.61.1.21

Cialdini, R. B., Demaine, L. J., Sagarin, B. J., Barrett, D. W., Rhoads, K., & Winter, P. L. (2006). Managing social norms for persuasive impact. *Social Influence, 1*(1), 3–15. https://doi.org/10.1080/15534510500181459

Darke, P. R., & Chung, C. M. (2005). Effects of pricing and promotion on consumer perceptions: It depends on how you frame it. *Journal of Retailing*, *81*(1), 35–47. https://doi.org/10.1016/j.jretai.2005.01.002

Davidai, S., & Shafir, E. (2018). Are ‘nudges’ getting a fair shot? Joint versus separate evaluation. *Behavioural Public Policy*, 1–19. https://doi.org/10.1017/bpp.2018.9

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

Ebeling, F., & Lotz, S. (2015). Domestic uptake of green energy promoted by opt-out tariffs. *Nature Climate Change, 5*(9), 868–871. https://doi.org/10.1038/nclimate2681

Feller, I., & Stern, P. C., (2007). *A strategy for assessing science: Behavioral and social research on aging*. National Academies Press. https://doi.org/10.17226/11788

Food Standards Agency. (2016). *Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets*. https://www.food.gov.uk/sites/default/files/media/document/fop-guidance_0.pdf

Franckle, R. L., Levy, D. E., Macias-Navarro, L., Rimm, E. B., & Thorndike, A. N. (2018). Traffic-light labels and financial incentives to reduce sugar-sweetened beverage purchases by low-income Latino families: A randomized controlled trial. *Public Health Nutrition*, *21*(8), 1426–1434. https://doi.org/10.1017/S1368980018000319

Gigerenzer, G. (1998). Surrogates for theories. *Theory & Psychology, 8*(2), 195–204. https://doi.org/10.1177/0959354398082006

Gigerenzer, G. (2015). *Bauchentscheidungen: die Intelligenz des Unbewussten und die Macht der Intuition*. Bertelsmann Verlag.

Gilbert, D. T. (1999). What the mind's not. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 3–11). Guilford. http://www.danielgilbert.com/Gilbert%20%28What%20the%20minds%20not%29.PDF

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment.* Cambridge University Press. https://doi.org/10.1017/CBO9780511808098

Hagmann, D., Siegrist, M., & Hartmann, C. (2018). Taxes, labels, or nudges? Public acceptance of various interventions designed to reduce sugar intake. *Food Policy, 79*, 156–165. https://doi.org/10.1016/j.foodpol.2018.06.008

Halpern, D. (2015). *Inside the nudge unit: How small changes can make a big difference*. Random House.

Halpern, D. (2019a, October 3). *David Halpern discusses the frontiers and challenges for behavioural science at BX2019* [Video]. YouTube. https://www.youtube.com/watch?v=Li7TFa692os

Halpern, D. (2019b). Finding the 'unicorns': Behavioural science in government grows up. *Civil Service Quarterly, 21*, 16–21. https://quarterly.blog.gov.uk/wp-content/uploads/sites/5/2019/10/6.5989_CO_CSQ_21-Oct2019_web_v141.pdf

Hansen, P. G. (2016). The definition of nudge and libertarian paternalism: Does the hand fit the glove? *European Journal of Risk Regulation, 7*(1), 155–174. https://doi.org/10.1017/S1867299X00005468

Hansen, P. G., & Jespersen, A. M. (2013). Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy. *European Journal of Risk Regulation, 4*(1), 3–28. https://doi.org/10.1017/S1867299X00002762

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

Hawkes, C. (2013). *Promoting healthy diets through nutrition education and changes in the food environment: An international review of actions and their effectiveness*. Nutrition Education and Consumer Awareness Group. http://www.fao.org/3/i3235e/i3235e.pdf

Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, *12*(6), 973–986. https://doi.org/10.1177/1745691617702496

Hieke, S., & Wilczynski, P. (2012). Colour me in: An empirical study on consumer responses to the traffic light signposting system in nutrition labelling. *Public Health Nutrition, 15*(5), 773–782. https://doi.org/10.1177/1745691617702496

Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives?. *Science, 202*(5649), 1338–1339. https://doi.org/10.1126/science.1091721

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

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

Liem, D. G., Miremadi, F., Zandstra, E. H., & Keast, R. S. (2012). Health labelling can influence taste perception and use of table salt for reduced-sodium products. *Public Health Nutrition*, *15*(12), 2340–2347. https://doi.org/10.1017/S136898001200064X

Ly, K., Mazar, N., Zhao, M., & Soman, D. (2013). *A practitioner's guide to nudging*. Rotman School of Management. http://dx.doi.org/10.2139/ssrn.2609347

MacCallum, G. C., Jr. (1967). Negative and positive freedom. *The Philosophical Review, 76*(3), 312–334. https://doi.org/10.2307/2183622

Maule, J., & Villejoubert, G. (2007). What lies beneath: Reframing framing effects. *Thinking & Reasoning*, *13*(1), 25–44. https://doi.org/10.1080/13546780600872585

National Academies of Sciences, Engineering, and Medicine. (2016). *Attribution of extreme weather events in the context of climate change*. National Academies Press. https://doi.org/10.17226/21852

Newman, G. E., & Shen, Y. J. (2012). The counterintuitive effects of thank-you gifts on charitable giving. *Journal of Economic Psychology*, *33*(5), 973–983. https://doi.org/10.1016/j.joep.2012.05.002

Norman, E. (2010). "The unconscious" in current psychology. *European Psychologist*, *15*(3), 193–201. https://doi.org/10.1027/1016-9040/a000017

OECD. (2017a). *Behavioural insights and public policy: Lessons from around the world*. OECD Publishing. https://doi.org/10.1787/9789264270480-en

OECD. (2017b). *Obesity update 2017*.

Parker, R. (2017, February 13). Got a ham sandwich for lunch today? One in six workers admit eating the same thing every day for two years. *Daily Mail Online*. https://www.dailymail.co.uk/news/article-4218246/Would-eat-ham-butty-day-20-years.html

Pattanaik, P. K., & Xu, Y. (1990). On ranking opportunity sets in terms of freedom of choice. *Recherches Économiques de Louvain/Louvain Economic Review*, *56*(3), 383–390. https://doi.org/10.1017/S0770451800043955

Pfister, H. R., Jungermann, H., & Fischer, K. (2017). *Die Psychologie der Entscheidung*. Springer. https://doi.org/10.1007/978-3-662-53038-2

Pichert, D., & Katsikopoulos, K. V. (2008). Green defaults: Information presentation and pro-environmental behaviour. *Journal of Environmental Psychology, 28*(1), 63–73. https://doi.org/10.1016/j.jenvp.2007.09.004

Rebonato, R. (2014). A critical assessment of libertarian paternalism. *Journal of Consumer Policy, 37*(3), 357–396. https://doi.org/10.1007/s10603-014-9265-1

Rodway, P., Schepman, A., & Thoma, V. (2016). Reachability does not explain the middle preference: A comment on Bar-Hillel (2015). *i-Perception, 7*(2), 1–5. https://doi.org/10.1177/2041669516639959

Rogers, W. A., Pak, R., & Fisk, A. D. (2007). Applied cognitive psychology in the context of everyday living. In F. T. Durso, R. S. Nickerson, S. T. Dumais, S. Lewandowsky, & T. J. Perfect (Eds.), *Handbook of applied cognition* (pp. 1–27). Wiley. https://doi.org/10.1002/9780470713181

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.

Rutgers (2020). Print green. https://oit-nb.rutgers.edu/service/printing/printgreen

Schneider, H., & Venetz, W. (2014). *Cost of obesity in Switzerland in 2012*. Bundesamt für Gesundheit. https://www.bag.admin.ch/dam/bag/de/dokumente/npp/forschungsberichte/forschungsberichte-e-und-b/cost-of-obesity.pdf.download.pdf/cost-of-obesity.pdf

Schellenhuber, H. J., (2006). *Avoiding dangerous climate change*. Cambridge University Press. http://gyohe.faculty.wesleyan.edu/files/2018/05/69.pdf

Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & Cognition, 21*(4), 546–556. https://doi.org/10.3758/bf03197186

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

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29*(3), 281–295. https://doi.org/10.1177/002224379202900301

Stanovich, K. E., & Toplak, M. E. (2012). Defining features versus incidental correlates of type 1 and type 2 processing. *Mind & Society*, *11*(1), 3–13. https://doi.org/10.1177/1745691612460685

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Brookings Institution Press.

Sunstein, C. R. (2014). Choosing not to choose. *Duke Law Journal*, *64*, 1–52. https://scholarship.law.duke.edu/dlj/vol64/iss1/1

Sunstein, C. R., Reisch, L. A., & Kaiser, M. (2019). Trusting nudges? Lessons from an international survey. *Journal of European Public Policy*, *26*(10), 1417–1443. https://doi.org/10.1080/13501763.2018.1531912

Tabi, A., Hille, S. L., & Wüstenhagen, R. (2014). What makes people seal the green power deal? Customer segmentation based on choice experiment in Germany. *Ecological Economics, 107*, 206–215. https://doi.org/10.1016/j.ecolecon.2014.09.004

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Thunström, L. (2019). Welfare effects of nudges: The emotional tax of calorie menu labelling. *Judgment and Decision Making, 14*(1), 11–25.

Tripp, S. (2017, November 13). Oh, BEhave! – Even experimentation experts get surprised. Brooks Bell. https://www.brooksbell.com/resource/blog/oh-behave-even-experimentation-experts-get-surprised/

Valenzuela, A., & Raghubir, P. (2009). Position-based beliefs: The centre-stage effect. *Journal of Consumer Psychology, 19*(2), 185–196. https://doi.org/10.1016/j.jcps.2009.02.011

Valenzuela, A., & Raghubir, P. (2015). Are consumers aware of top–bottom but not of left–right inferences? Implications for shelf space positions. *Journal of Experimental Psychology: Applied, 21*(3), 224–241. https://doi.org/10.1037/xap0000055

Vlaev, I., King, D., Dolan, P., & Darzi, A. (2016). The theory and practice of "nudging": Changing health behaviors. *Public Administration Review, 76*(4), 550–561. https://doi.org/10.1111/puar.12564

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Walter, S. (2016). *Illusion Freier Wille?* JB Metzler. https://doi.org/10.1007/978-3-476-05445-6

Wansink, B. (2007). *Mindless eating: Why we eat more than we think*. Bantam.

Wansink, B., & Hanks, A. S. (2013). Slim by design: Serving healthy foods first in buffet lines improves overall meal selection. *PLoS One, 8*(10), e77055. https://doi.org/10.1371/journal.pone.0077055

Wedell, D. H. (1997). Another look at reasons for choosing and rejecting. *Memory & Cognition, 25*(6), 873–887. https://doi.org/10.3758/BF03211332

World Health Organization (WHO). (2020, April 1). *Obesity and overweight.* https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

Williams, E.P., Mesidor, M., Winters, K., Dubbert, P. M., & Wyatt, S. B. (2015). Overweight and obesity: Prevalence, consequences, and causes of a growing public health problem. *Current Obesity Reports, 4*(3), 363–370. https://doi.org/10.1007/s13679-015-0169-4

Acknowledgements

I would like to thank all the people who have supported me in this endeavor. In particular, I would like to thank Ester Reijnen for introducing me to the world of cognitive research and for all her patience, valuable advice and encouragement during these years and especially during the process of writing this thesis. Klaus Opwis for his support and the opportunity to write a thesis. My colleges and co-authors for their dedication and assistance but also for the moments of laughter inside and outside the office. To my family and friends for their love and the occasional reminder to see the other things in life. I am also grateful to have grown up in an environment that fosters education and scientific progress. And I would like to thank the lecturer who proofread this thesis.

Curriculum Vitae

Personal Details

Name	Swen Jonas **Kühne**
Date of birth	7th February 1986

Education

2016 – 2020	Doctoral program in Psychology University of Basel, Faculty of Psychology
2018 – 2020	Academic Teaching Certificate University of Basel
2013 – 2015	Master of Science in Applied Psychology Zurich University of Applied Science, School of Applied Psychology
2010 – 2013	Bachelor of Science in Applied Psychology Zurich University of Applied Science, School of Applied Psychology

Professional experience

2015 – 2020	Research Assistant, Applied Cognitive Psychology Zurich University of Applied Science, School of Applied Psychology
2016 – 2020	Lecturing, *Introduction and History of Psychology* and *Ring Lectures* Zurich University of Applied Science, School of Applied Psychology

Appetite 143 (2019) 104410

Contents lists available at ScienceDirect

Appetite

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

Choosing or rejecting a food item, does framing matter? And what has sugar to do with it!

Ester Reijnen[a,*], Swen J. Kühne[a], Markus Stöcklin[b], Jeremy M. Wolfe[c,d]

[a] Zurich University of Applied Sciences, Switzerland
[b] University of Basel, Switzerland
[c] Brigham & Women's Hospital, Boston, USA
[d] Harvard Medical School, Boston, USA

ARTICLE INFO

Keywords:
Preference order
Choose/reject
Sugar
Traffic light label
Accentuation

ABSTRACT

The color code of "Traffic Light Labels" (TLL) on food items indicates the amount (e.g., green = low) of fat, saturates, sugar and salt it contains. Consider two ways to select among food items (e.g., two cereal bars) based on their TLLs. You might *choose* between the two items or you might *reject* one of the two. Furthermore, differences between choose and reject might be driven more strongly by one factor (e.g., sugar) than by others. In Study 1 our participants made choose or reject decisions between food items with an *all-orange* TLL (all moderate) and a *2 red/2 green* TLL (2 negative/2 positive). Both items had equal energy/caloric content. We found that, independent of the condition (Choose/Reject), participants went home more often with the 2 red/2 green item if sugar was green. This effect was stronger in the Choose than in the Reject condition. In Study 2, we additionally manipulated the energy content (low, high) of the items. In the case where both food items had a low energy content, similar results as in Study 1 were observed. If either or both items had high energy content, the choose/reject interaction with sugar disappeared. Only differences in energy content played a role in the reject condition. Overall, our results can be better explained by an "accentuation hypothesis" than by a "compatibility hypothesis". These findings could be used by choice architects to fight the current obesity crisis.

If we had a better understanding of the underlying mechanisms of our daily decisions about food, it would be easier to develop intervention strategies that could address issues like the global obesity crisis (see NCD Risk Factor Collaboration, 2016, for future trends). For example, consider a decision between consuming a hazelnut cereal bar or a crunchy chocolate bar. That decision could be framed as a choice of one of the bars or a rejection of one of them. Does that framing matter? Might you end up with the hazelnut cereal bar in one case and with the crunchy chocolate bar in the other? Regardless of how it is framed, the decision could be informed by the increasingly common "front-of-package" labels, that display information about the food's nutrients such as fat, saturates, sugar, etc. Does that information (especially the sugar content information) interact with the framing of the decision in terms of choosing or rejecting a food item? We focus on the sugar content information since recent research has shown that obesity might be tightly connected to the growing consumption of *sugar-sweetened* foods and beverages (see Johnson et al., 2007). Previously, sugar had not been considered as a key contributor to obesity. In the era of the heart attack of US president Dwight D. Eisenhower in 1955, researchers like Ancel Keys (Page et al., 1961) proposed that heart attacks are caused by *fatty* diets (see his seven-country study). As a consequence, fat (though A. Keys was actually talking about saturates, not fat per se) was considered as the chief culprit in coronary diseases and obesity. Hence, it was suggested that US citizens should lower their fat intake from 45 to 30%. This, in turn, led food producers to reduce the fat content of food items and to introduce so-called "light" or "diet" food items to the market. Those food items found approval though, unbeknownst to many, the reduced fat content went along with an increase in sugar. The Keys study might have been flawed by his omission of data points from some countries (justified or not). Even at that time scientists, such as John Yudkin (1964), argued that one should actually focus on sugar as the cause for obesity rather than fat. Only more recently was his claim heard, as more recent research supports Yudkin's claim that consumption of sugar-sweetened food and beverages (see Hu, 2013 or Luger et al., 2017, for a review paper) is directly related to obesity. However, the findings about the role of sugar are not entirely

* Corresponding author. ZHAW Zurich University of Applied Sciences, School of Applied Psychology, Pfingstweidstrasse 96, Postfach 707, CH-8037, Zürich, Switzerland.

E-mail address: ester.reijnen@zhaw.ch (E. Reijnen).

https://doi.org/10.1016/j.appet.2019.104410
Received 22 February 2019; Received in revised form 15 August 2019; Accepted 18 August 2019
Available online 19 August 2019
0195-6663/

straightforward. Some countries, such as Germany, have relatively low obesity rates, despite their high sugar consumption (Ferdman, 2015). Hence, high sugar consumption should be considered as just one of the contributors to obesity (Siervo et al., 2014). Nevertheless, while our daily sugar consumption (e.g., in 2016/2017: Australia: 40 sugar cubes à 4 g/day; EU: 25 cubes/day, Statista, 2017) has been considered as too high (see World Health Organization, 2015), intervention strategies that could help us to reduce our daily sugar consumption are rare, especially if one excludes paternalistic approaches (e.g., sugar taxes). One promising strategy is the use of so-called *front-of-package labels,* especially the traffic light label (TLL) developed by the British Food Standards Agency (2016) and others. The TLL uses a simple color-coding scheme to inform consumers whether a food item is low (green color), medium (orange) or high (red) in fat, saturates, sugar and salt. The TLL uses *graphic and symbolic* elements in order to improve the understanding of nutrient information by low-literate consumers (Viswanathan, Torelli, Xia, & Gau, 2009) and to meet the consumers' time constraints. The labels do have some effect. For example, Franckle, Levy, Macias-Navarro, Rimm, and Thorndike (2018), showed that the use of a TLL did reduce the monthly consumption of unhealthy "red" beverages by 9%-points compared to a control group. These results are consistent with the findings from Cecchini and Warin's (2016) meta-analysis showing that the TLL is an effective way of steering consumers toward choosing healthier food items. The TLL seems to have an advantage over labels such as the guideline of daily amount (GDA), which usually displays only the energy content and its relation (in percentage) to the recommended daily consumption. However, many issues remain unknown concerning the use of TLL information. For example, how do consumers weigh and integrate the multiple pieces of information displayed (fat, saturates, sugar, etc.) into a single value or utility of the food's overall healthiness in order to form the basis for binary decisions. Are people operating according to a rational theory of choice (see WADD[1]) or do they use some kind of simple heuristic (e.g., LEX[2])? Support for the use of simple heuristics comes from, for example, Hieke and Wilczynski's (2012) study showing that consumers seem to preferentially consider *sugar* and *fat* information in their evaluation of the food item's overall healthiness. On the other hand, in Scarborough et al.'s (2015) study, consumers seemed to focus on *saturates* and *salt*. The difference in findings has been attributed to the fact that different food items were used (yoghurt vs. ready meals) and/or to cultural differences (Germany vs. UK).

We are focused on a somewhat different question. Does framing the decision as a choice of one item lead to a different preference order and hence to a different choice pattern compared to framing the decision as a rejection of an item.

From the perspective of a rational theory of choice, a consumer's well-defined preference orders (e.g., prefer the hazelnut cereal bar over the crunchy chocolate bar) should be independent of, for example, the context in which decisions are made, such as the methods of elicitation (procedure invariance). However, research has shown that because our preferences are being "constructed" at the time of elucidation, they are open to being manipulated by different factors such as, for example, *scale compatibility* (see, for example, Slovic, Griffin, & Tversky, 1990).

In addition to scale compatibility, Shafir (1993) showed that there also exists an effect of the compatibility of the *reasons available* for choosing or rejecting an option and the task at hand. More precisely, he stated that (p. 547) an "option's advantages provide compelling reasons for choosing, and thus make choices easier to determine and justify to oneself and to others. An option's disadvantages provide natural reasons for rejecting, and thus make rejection easier to determine and justify". Hence, according to Shafir, positive features should be *weighted more* when *choosing* than when rejecting, whereas the reverse should be true for negative features. Hence, when faced with a decision between a healthy, but expensive dish (say Sushi) and a less healthy, but cheap dish (Hot dog), people will, on the one hand, tend to choose Sushi because it is healthier. At the same time, if the decision is framed as a rejection, people will tend to reject Sushi (i.e., end up with the Hot dog) because it is more expensive. This pattern has been observed by Shafir (1993) in his experiment where he presented his participants with the task (problem 1) to either award ("choose situation") or deny ("reject situation") sole custody of the child to parent A or B. In this study, parent A just had "average" features (e.g., average health; thereby called the *impoverished* option) parent B had positive features (e.g., above-average income) and negative features (e.g., minor health problems; the *enriched* option). Again, according to the rational theory of choice, choosing and rejecting should be complementary and percentages for choosing (Pc) and rejecting (Pr) should add up to 100%. If you choose parent A 70% of the time in the choose condition, you should reject parent A 30% of the time in the reject condition. However, according to Shafir (1993) Pc + Pr should be greater than 100 percent for the enriched option, since that option should not only be chosen more often in the choose condition but also be rejected more often than the impoverished option in the reject condition. That is exactly what Shafir found: Depending on whether participants had to choose or reject, parent B got more (in 64% of the cases) or less (45%) often sole custody of the child (64% + [100%–45%] = 119%; see the discussion section for contradictory results to Shafir's paper as well as alternative explanations for the observed effect).

In the present paper, the question is whether we find a similar compatibility principle pattern of results in the choosing or the rejecting of food items that incorporate the TLL. A useful feature of TLLs is that they can be manipulated in the way that they can represent, as in Shafir's (1993) study, the so-called enriched option (labels with 2 red and 2 green traffic lights) or the "average/*impoverished*" option (labels with all-orange traffic lights). However, a key difference from Shafir's study is that in our study the same feature (e.g., sugar content) can be either negative (e.g., high in sugar = red) or positive (e.g., low in sugar = green) in the enriched option. In Shafir's study, the same feature was consistently positive or negative (e.g. in his problem 1 income, if mentioned, was positive: above-average). Our TLL version allows us to investigate the interaction of specific features with the choose/reject framing. To sum up, in our study we not only want to replicate Shafir's basic findings in a new environment - the context of food choice[3] - but also to extend his research by varying the feature's value.

1. Study 1

To answer our research questions, we presented two cereal bars with their respective TLLs (specifications see Method part) to our participants and asked them either to choose or reject one of them. We chose cereal bars because they have a large variability in their nutrient profiles (e.g., low in fat but high in sugar or vice versa).

2. Method

2.1. Participants

580 participants (M_{age} = 26.0; SD_{age} = 5.8; 69.7% female) from the ZHAW Zurich University of Applied Sciences and the greater area of

[1] WADD = weighted additive rule (see Gigerenzer & Selten, 2002, for an overview).

[2] LEX = lexicographic heuristic (see Gigerenzer & Selten, 2002, for an overview).

[3] Nagpal, Lei, and Khare (2015) conducted a study which investigates – though the title suggest otherwise – the effect of a default in the context of food; that is, whether a pre-selection (= reject condition) of all healthy and unhealthy toppings on a list has a different effect on choice than when none of the toppings were pre-selected (= choose condition).

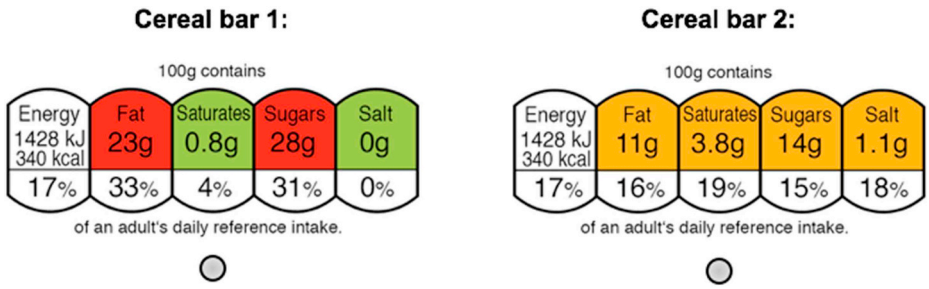


Fig. 1. Participant's choice situation (left: red/green label; right: all-orange label). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Zurich took part in this web-based study. Participants from the ZHAW School of Applied Psychology could receive course credit for their participation (which 9.5% did). All participants gave informed consent.

2.2. Stimulus material

Participants were simultaneously presented with *two cereal bars* respectively their traffic light labels (TLL; see Fig. 1 for an illustration). One of the two cereal bars always had four *orange traffic lights*, while the other one had *two red* and *two green* lights (e.g., fat and sugars red, saturates and salt green). We had 6 versions of the "red/green" cereal bar: representing all possible combinations of two red and two green traffic lights.[4] The color-coding followed the "guide to creating a front of pack nutrition label" developed by the UK Department of Health (2013). Labels were created in such a way that the overall energy content of the all-orange and the red/green cereal bars was equal, as well as the mean percentage of the nutrition profile (e.g., cereal bar 1: [33% + 4% + 31% + 0%]/4 = 17%; cereal bar 2: [16% + 19% + 15% + 18%]/4 = 17%). Minuscule jitterings were added so that, for example, not all nutrients had the same percentage in the all-orange bar. Thus, rather than all being 17%, they could vary between 15% and 19%.

2.3. Procedure

Participants were first asked to imagine that they were hungry and wanted to buy a cereal bar at a kiosk. Participants were randomly assigned to the Choose or Reject conditions in which they were asked either to *choose* one of the two cereal bars or to *reject* one of them (the words *choose* and *reject* were highlighted). Participants were further randomized to receive one of six versions of the red/green label that was presented next to the all-orange one. For our purposes, the six versions were divided into two groups of interest: sugar-red or sugar-green. Thus, this was a 2 (condition: choose, reject) x 2 (sugar: red, green) factorial design. After participants had made their decision they proceeded to a second task ("healthiness perception rating"). Here they were presented with a series of 15 differently color-coded traffic light *label combinations* in a randomized order. These were: the two labels already presented in their 2 AFC task (the all-orange one and one of the 6 possible 2 red/2 green combinations), as well as the five other 2 red/2 green combinations (the ones not presented in the 2 AFC task) and all possible label combinations with either one green and three red or three green and one red traffic lights (8 in total). For each label, participants were asked to state how often they thought that they could eat food with such a label on it while still following a healthy diet. Participants had to choose from the following three answers: *often*, *occasionally* and *rarely* (adapted from the Dutch Health report, 2006). Finally, we assessed participants' demographic data.

3. Results and discussion

Participants excluded. Participants who needed more than 30 or less than 3 min to complete the whole experiment and/or took more than 300 or less than 5 s to complete the 2AFC task were excluded from the analysis (altogether 2.2% of participants).

Healthiness perception rating. We first looked at participants' perceptions (see Fig. 2) about the *overall healthiness* of the 15 differently color-coded cereal bars (e.g., 1 red/3 green, all-orange, etc.). The focus of our analysis is thereby on the *all-orange* and the *2 red/2 green* bar respectively participants' answers regarding "rarely", "occasionally", and "often". A roughly equal proportion of participants stated that you can eat food labeled with either *four orange* or *2 red/2 green* traffic lights "occasionally" (proportions test; $z = 1.922$, $p = .055$).[5] However, participants thought that they can eat food labeled with the all-orange bar more "often" than food labeled with the 2 red/2 green bar ($z = 11.750$, $p < .001$); whereas the opposite is true for "rarely" ($z = 9.244$, $p < .001$; for more results see footnote[6]).

Thus, participants perceived the all-orange bar overall as healthier (less "rarely" votes) than the 2 red/2 green bars. The question is, do these ratings have an effect on participants choice behavior in the "2 AFC task"?

2 AFC task. Fig. 3 illustrates the number of participants (in %) that "went home" with the red/green cereal bar (note, that when a participant rejected the orange bar, he/she would be going home with the red/green bar). A binary logistic regression was calculated using the R statistics software. The results show a significant main effect[7] for sugar (red vs. green), $\chi^2(1, N = 580) = 45.10, p < .001$. Though there was no main effect for condition (choose vs reject): $\chi^2[1, N = 580] = 0.77$, $p = .38$), there was a significant Sugar x Condition interaction; $\chi^2(1, N = 580) = 8.18$, $p < .01$. Clearly, participants were more likely to go home with a bar having a green light for sugar than with a bar having a red light; indicating that participants did pay attention to the nutrient's (here: sugar) color scheme. Furthermore, the significant interaction indicates that the color-coding of the "feature" sugar was more important when

[4] Note, we presented a red/green combination that actually never appears on realistic labels: saturates "red" and fat "green" - saturates are part of fat and thereby the amount of saturates can never be higher than the total amount of fat. We nevertheless tested those combinations. We, however, asked the participants at the end of the experiment, whether they had noticed it. 62 participants which were confronted with such a combination in the 2 AFC task did notice it. However, the same pattern of results was found when we removed those participants from the analysis.

[5] The same results were found, when we excluded the participants that noticed the impossible combination (see footnote 4). Additionally, choices were considered as independent from each other.

[6] Within the all-orange bar an equal amount of participants thought that they can eat that bar "rarely" and "often" ($z = 1.747$, $p = .081$). The same was true regarding the 2 red/2 green bar for "often" and "occasionally" ($z = 1.164$, $p = .245$).

[7] When not otherwise stated, main effects are reported without higher order factors (similar to a classical ANOVA).

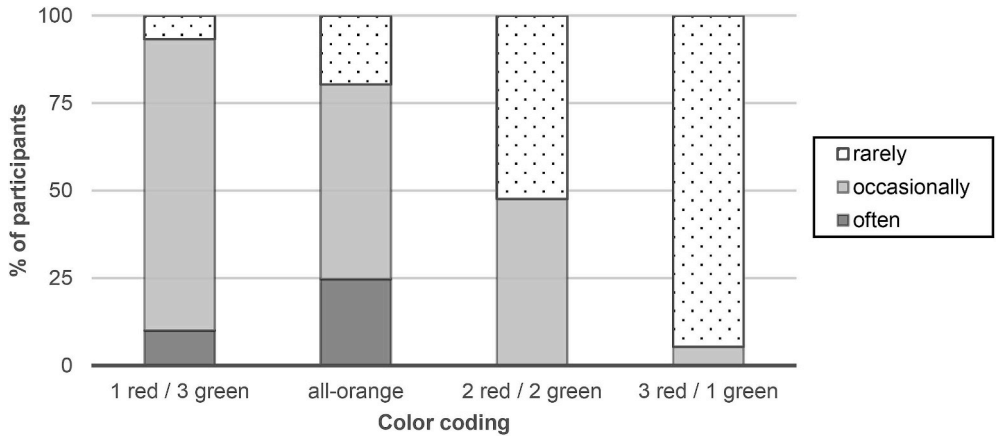


Fig. 2. The percentage of assignments to the 3 answer categories depending on the coloring of the label.

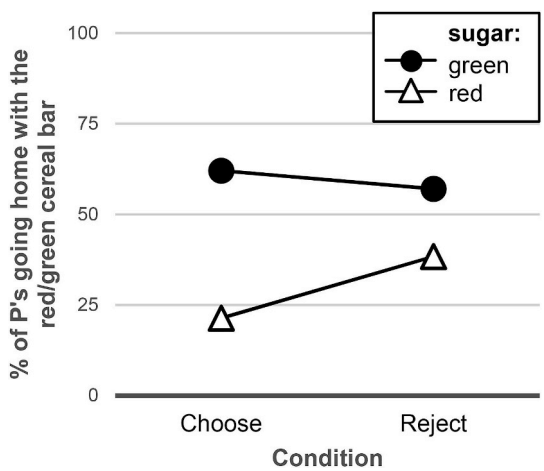


Fig. 3. Number of participants (in %) going home with the red/green cereal bar.

participants had to choose than when they had to reject.[8]

Considering participants "healthiness perception rating" results, we see the odd result that although participants perceived the all-orange bar in that "healthiness perception rating" (less "rarely" votes) as healthier than the 2 red/2 green bars, they "went home" in the 2 AFC task as equally often with the all-orange (51.7%) as with 2 red/2 green (48.3%) bar ($z = 0.789$, $p = .430$). That is, the explicit healthiness perception ratings did not correspond with the actual choice behavior in the 2 AFC task.

How do the results of our Study 1 relate to the results of Shafir (1993)? To recapitulate, as in Shafir's study our participants had to choose or reject one of two options. Additionally, in our study a given feature's (sugar) value could either be negative (red = high in sugar) or positive (green = low in sugar). Now, our results on the one hand replicate Shafir's (1993) basic finding (i.e., a difference between the choose and reject conditions). However, our results go beyond the original finding, showing that this choose/reject effect is modulated by the specific value of the feature. That is, while our participants in the sugar "green" condition behaved in line with Shafir's (1993) expectations, they did not in the sugar "red" condition. A version of our "red" condition finding has also been observed by Ganzach (1995) in a problem where participants had to choose or to reject a job candidate. Among others, Chen and Proctor (2017; see also Wedell, 1997; Colombo, Nicotra, & Marino, 2002) trace that difference in findings back to the fact that the choose/reject effect is not caused by the compatibility principle, as stated by Shafir, but by the *accentuation hypothesis*. That hypothesis states that "when the enriched option is liked overall more than the impoverished option, given that people will tend to be more discriminating in choice, and will therefore accord more importance to positive features, the choices of the enriched option (where positive features are extremely so) will be greater than the rejections (of the alternative). If, on the other hand, the enriched option is not liked much overall (which means that its positive features have a smaller weight in its evaluation), it is the rejections of the impoverished option, rather than the choices of the enriched option, that will be dominant (i.e., the proportions of the former will be greater as compared to the latter)" (Colombo et al., 2002, p. 21). Hence, participants should go home more often with the red/green cereal bar in the Choose condition than the Reject condition if overall attraction is high, respectively vice versa if overall attraction is low.[9] If the preference is the same, the differences between choose/reject should disappear. Since we have just four data points in Study 1 we cannot run the analysis that is needed to back-up either principle/theory. We will come back to this point in Study 2.

Furthermore, in our Study 1, both labels indicated the same energy content. Stubbs, Ferres, and Horgan (2000) found a positive correlation between participants' food density (KJ/g) and energy intake (number of calories consumed in a day): The higher the food density, the more calories were consumed in a day. Hence, the energy content might also be a factor that should be considered in fighting the obesity crisis. Yet, to date it is unclear whether, and to what extent, people consider energy content in their food choices. Studies with restaurant menus have shown that the number of calories displayed, using caloric labels or caloric flags, have little or no effect on the number of calories purchased (Swartz, Braxton, & Viera, 2011; see also; Kiszko, Martinez, Abrams, & Elbel, 2014). With the exception of Cioffi, Levitsky, Pacanowski, and Bertz (2015), all these studies have used short periods of observation. Using longer time periods of observation, Cioffi et al. (2015) found a significant reduction of high calorie food items purchased and a corresponding increase in purchase of low caloric food items. Other than these restaurant menu studies, a few studies have investigated the effects of caloric labels on purchase decisions. In this respect, Kiesel and

[8] Note, since we found the same pattern of results regarding the red sugar line independent of whether it was paired with fat, saturates or salt (pairing: $\chi^2[2, N = 292] = 1.21$, $p = .546$; Pairing x Condition interaction: $\chi^2[2, N = 292] = 0.01$, $p = .933$) we did run our analysis on aggregated data. The same was true for the pairing of "green" sugar with other nutrients (pairing: $\chi^2[2, N = 288] = 1.52$, $p = .468$; Pairing x Condition interaction: $\chi^2[2, N = 288] = 3.,37$, $p = .186$).

[9] Under the compatibility principle, more participants should - independent of the overall attractiveness – go home with the red/green cereal bar in the choose than in the reject condition.

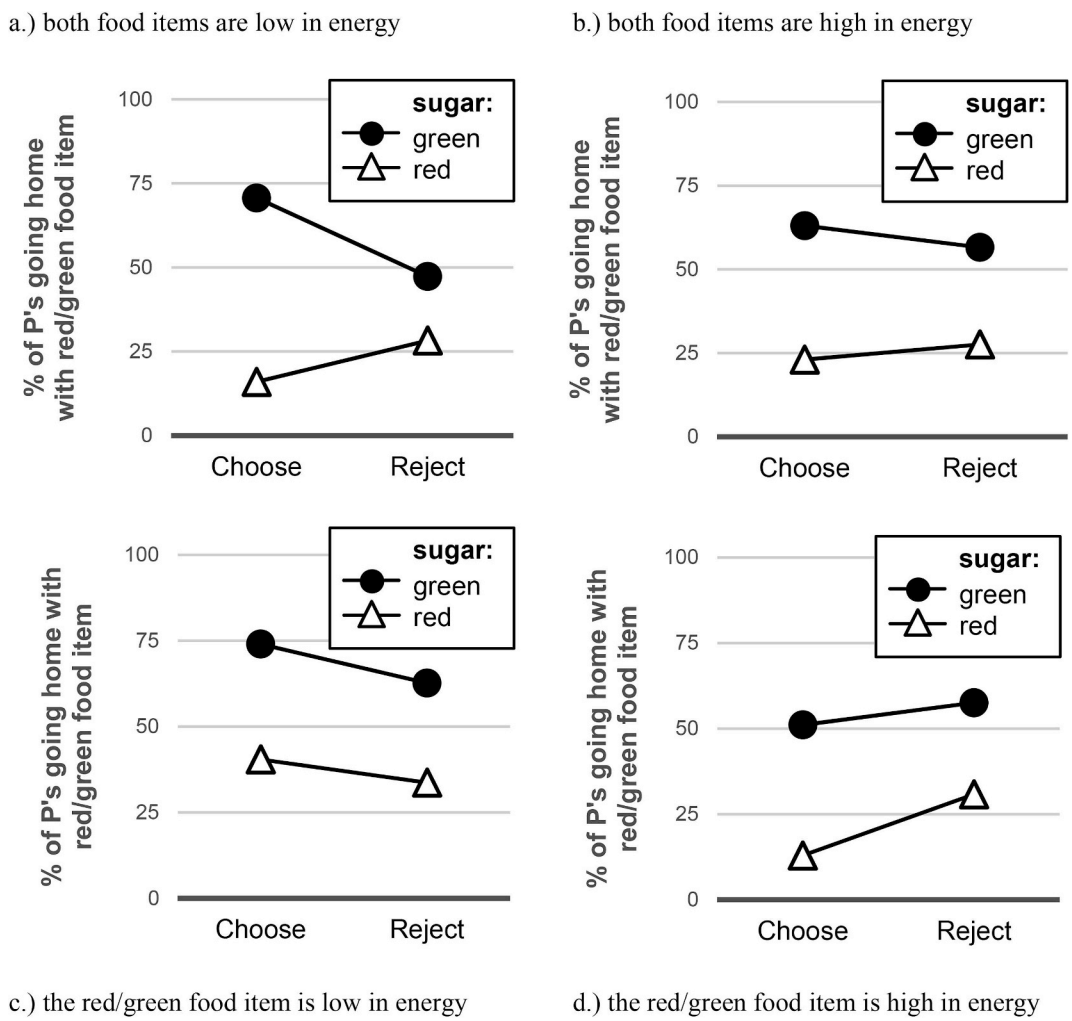


Fig. 4. a.) both food items have a low energy content (340 kcal; see Study 1), b.) both have a high energy content (470 kcal), c.) the red/green food item has the lower energy content, d.) the red/green food item has the higher energy content.

Villas-Boas (2013) have found that when microwave popcorn in supermarkets were in addition labeled with a "low calorie" label, popcorn sales increased; however, this is only so when that label did not contain other information such as "low fat". In Study 2, we investigate if caloric content information also influences people's choices and, if it does, is that influence different if participants are asked to choose or reject.

4. Study 2

In Study 2 we manipulated the energy content of the 2 bars, since the total amount of energy in a food item, next to sugar, might be important information in fighting the obesity crisis. Additionally, we expanded our materials from the cereal bars of our first study to include chips and cereals. We added these food item categories in order to diversify the nutrient profiles. Note that all of those items can be made from different ingredients and vary broadly in their overall healthiness. Finally, our Study 2 should help us to solve the conundrum of our finding in Study 1 that the participants in the sugar "green", but not in the sugar "red" conditions behaved in accord with Shafir's (1993) expectations. By varying the energy content of the food items, we end up with enough data points to run the appropriate analysis to explain whether the compatibility or the accentuation hypothesis can explain our data.

5. Method

5.1. Participants

520 participants (M_{age} = 25.7; SD_{age} = 5.8; 72.3% female) from the ZHAW Zurich University of Applied Sciences and the greater area of Zurich took part in this web-based study. Participants could take part in a raffle for one of 10 cinema vouchers, each with a value of 17.- swiss francs (equals 17 $). Students from the School of Applied Psychology could opt for course credit instead of the raffle (12% did this). All participants gave informed consent.

5.2. Stimulus material and procedure

The same stimulus materials and procedures as in Study 1 were used with the following exceptions. First, the energy content of the two items could vary in the following way: either both items could have the same energy content (both 340 kcal as in Study 1 or both 470 kcal) or a different energy content. If different, either the red/green item or that all-orange item had a high energy content (470 kcal) while the other was low (340 kcal) or vice versa. This produces four "*energy content combinations*". Second, we just used two variations of the "red/green" bar; either fat and sugar were high (red) or fat and salt were high (red), the other two ingredients were green, which produces two *sugar* conditions. Third, in Study 2, each participant had to make three decisions, one each for cereal bars (as in Study 1), chips and cereals. Finally, the

healthiness perception rating was replaced by the "*food category healthiness perception rating*". That is, participants had to state (on a 7-point Likert scale), regarding each of the three food items, how much they agreed with the statement "cereal bars [chip packages or cereals] are usually healthy (food item category health perception).

6. Results and discussion

Participants excluded. Participants who took more than 30 or less than 3 min to complete the whole experiment and/or took more than 300 or less than 4 s to complete the 2AFC task were excluded from the analysis (7% of total participants; 3% were removed by the 2AFC filter). Additionally, we excluded another two participants because they took part in the study twice.

2AFC task. Participants' choice behavior was not statistically different between the 3 food items (cereal bar, chips packages and cereals; χ^2 [2, N = 1560] = 4.40, p = .11) and because, respectively, the model fit was not better with the factor food-type included (χ^2 [26] = 35.61, p = .10), we combined the 3 food items and hence ran our analysis on the aggregated data.

Data were first analyzed separately, for each of the four energy content combinations, by running four binary logistic regressions. Fig. 4 a – d displays the percentage of participants who "went home" with the red/green food item. The caloric content of the foods modulates the effects of *Sugar* and *Choose/Reject* variables[10]. Turning first to Fig. 4a, when both food items are low in energy content (340 kcal), the result is a close replication of Study 1. In that condition, there is a significant main effect *for sugar*, $\chi^2(1, N = 384) = 50.74$, $p < .001$, as well as a significant Sugar x Choose/Reject Condition interaction; $\chi^2(1, N = 384) = 13.37$, $p < .001$. However, unlike Study 1, we also found a significant main effect for Choose/Reject Condition: $\chi^2(1, N = 384) = 4.26$, $p < .05$ (see Fig. 4a). In Fig. 4b, when both food items were *high* in energy (470 kcal), the Sugar x Condition interaction disappeared: $\chi^2(1, N = 393) = 1.33$, $p = .249$. The same is true for the two combinations where each item had a different energy content (Fig. 4c & d, both χ^2's < 3.07, p's > 0.079). The main effect of sugar is significant in all cases (all χ^2's > 21.91, p's < 0.001). When sugar is red, the item is always less desirable. The main effect of Choose/Reject Condition is significant when the red/green item is high in energy and the all-orange item is low, (Fig. 4c, $\chi^2[1, N = 384] = 8.20$, $p < .01$). Under the conditions shown in Fig. 4b and d, the effects of Choose/Reject Condition are not significant (both χ^2's < 0.97, p's > 0.324).

Fig. 4 may seem to show a disappointing diversity of results. However, combining conditions on the same graph reveals a fairly clear-cut story (see Fig. 5a & b). When participants are asked to <u>reject</u> one item, all of the conditions produce essentially the same result. Energy content has no effect (both χ^2's < 0.18, p's > 0.668) and participants go home with green sugar items more often than red sugar items (both χ^2's > 13.79, p's < 0.001). When participants had to <u>choose</u> an item, energy content did not make a difference in the case where both food items were equal in their energy content (χ^2 [1, N = 388] = 1.60, p = .205), whereas it did when they were different in the energy content (low vs. high; χ^2 [1, N = 386] = 16.81, $p < .001$). This is clear in Fig. 5b. Relatively high caloric content reduces the appeal of an item with green sugar (χ^2 [1, N = 192] = 16.81, $p < .001$) while relatively low caloric content increases the appeal of an item with red sugar (χ^2 [1, N = 194] = 10.65, $p < .01$). Overall, when participants are asked to Reject an item, it appears that their decision process is impoverished relative to the Choose condition. Fig. 5a shows that the preference for green-light sugar is less pronounced (as in Study 1). Fig. 5b shows that the effect of energy is greatly reduced. It is as if participants do not think about reject decisions as robustly as they think about choose decisions.

Compatibility principle vs. accentuation hypothesis. Response compatibility proposes that observers will focus on the positive feature when choosing and will focus on the negative features when rejecting. To check whether the choose/reject effect is based on response compatibility or rather on the overall attractiveness of the enriched option, we calculated linear regression models over the choose/reject data. For the data plotted in Fig. 6, Shafir (1993) compatibility account predicts two *parallel* regression slope lines with the choose line above the rejection line. In contrast, the accentuation account predicts two intersecting lines with the choose line steeper than the rejection line. If the lines lie on top of each other there is no framing effect at all.

As shown on the x-axis in Fig. 6, the relative attractiveness of the enriched option explains a significant amount of variance in the proportion of P's going home with the red/green food item in the choose condition, $F(1, 6) = 117.9$, $p < .001$, $R^2 = 0.97$, as well as in the reject condition, $F(1, 6) = 58.92$, $p < .001$, $R^2 = 0.91$. The coefficient of the regression function was significant in the choose condition, $\beta = 1.27$, $t(6) = 13.34$, $p < .001$, and in the reject condition, $\beta = 0.73$, $t(6) = 7.68$, $p < .001$.[11] More importantly, the difference between the two coefficients was significant, $t(7) = -4.016$, $p < .01$, with a steeper choose slope supporting a pattern predicted by the accentuation hypothesis. That is, when sugar is green, the enriched option is more preferred in the choose than the reject condition; whereas the opposite is true when sugar is red (a pattern observed also in our Study 1).

Food category healthiness perception rating. The three food categories were rated differently regarding their healthiness: $F(2, 1557) = 297.8$, $p < .001$. Post-hoc tests showed significant differences between all pairings (all p's < 0.001). Chips were rated as the least healthy ($M = 1.73$, $SD = 0.97$), followed by the cereal bars ($M = 2.85$, $SD = 1.18$). Cereals were considered to be the most healthful items ($M = 3.46$, $SD = 1.28$). Note, however, though participants differentiated between the categories in terms of the foods' perceived healthiness, this did not translate into a significant effect on their choice behavior in the 2AFC task.

7. General discussion

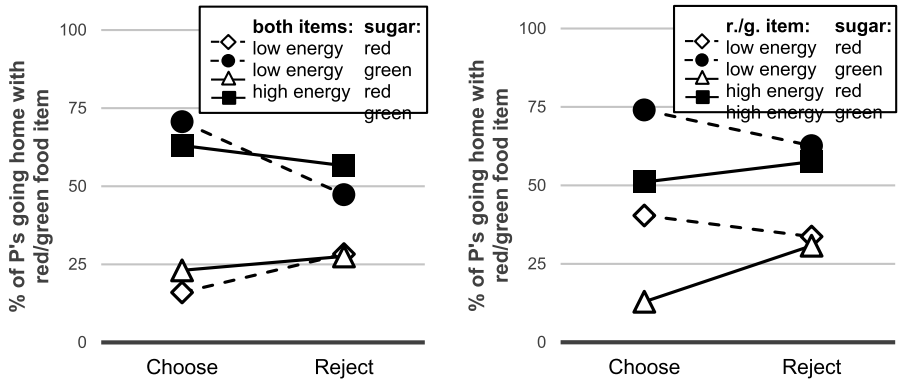
In general, front-of-package food labels, such as the TLL, seem to positively impact people's food item choices (Cecchini & Warin, 2016). However, it has been unclear how specific colors (green or red) in the TLL interact with the form of the question, the frame. Does it matter if the decision is framed as choosing one food item or rejecting the other?

In our Study 1, we found that participants based their decisions more on the sugar content than on any other variable. They went home more often with a food item that was low in sugar (green) than with one that was high in sugar. Moreover, the sugar content interacted with the framing of the decision as choice or rejection. The sugar effect was much stronger when participants had to choose a food item, than when they had to reject one. Study 2 (where we additionally manipulated the energy content) replicated the results of Study 1 in the case where both food items had a low energy content. However, when both items had a high energy content or when each had a different energy content (low vs. high or vice versa) the interaction of the decision frame with sugar was no longer significant. Furthermore, the energy content differentially affected decisions only in the choose frame and when the food items had a different energy content.

Overall, results from both studies can be better explained by the *accentuation theory* (e.g., see Wedell, 1997; Chen & Proctor, 2017) than by Shafir (1993) *compatibility hypothesis*. Recall that the accentuation theory states that decisions are made based on the "overall attractiveness" of the food item, and not based on the features that are more compatible with the task (e.g. the negative one in the reject task). In

[10] overall main effect of energy: $\chi^2(3, N = 1560) = 23.29$, $p < .001$.

[11] Similar values have been found by, for example, Wedell (1997): $b_{choose} = 1.20$, $b_{reject} = 0.82$.



a.) both items have the same energy content b.) food items have a different energy content

Fig. 5. Shows the comparison when both food items had the same (a.) or a different energy content (b.).

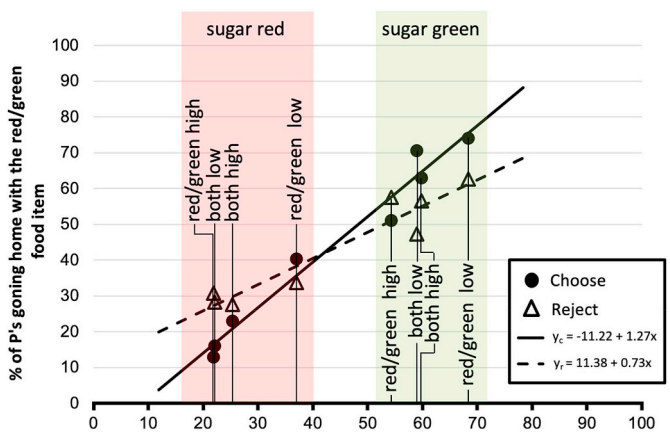


Fig. 6. Results of the two linear regression slopes indicating the match with the accentuation hypothesis (y-axis: for the Choose condition: % of P's going home with the red/green food item in the Choose condition (p_s); for the Reject condition: % of P's going home with the red/green food item in the Reject condition (p_F); x-axis: $p_o = (p_s + p_F)/2$. p_o serves as a measure of the relative attractiveness of the enriched option. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

line with that assumption, we found different slopes (between overall and for each condition's percentage of P's going home with the red/green item) for the choose and the rejection condition. In other words, in the case where more than 50% of the participants went home with a red/green food item, more of them were in a choose frame than in a rejection frame; the opposite was true in conditions where less than 50% of the participants went home with the red/green food item.

Chen and Proctor (2017) explain this type of effect with the *cognitive effort hypothesis,* which claims that cognitive effort is higher in a rejection task than in a choosing task. Assuming that a rejection task is more effortful, participants should be left with fewer cognitive resources for the decision (all orange vs. red/green food item) itself, since understanding the task already consumes some of the resources. Hence, participants should become less discriminating in the rejection task (i.e., percentage for choosing the all-orange vs. the red/green food items should become more similar) than in the choosing task. Their assumption was supported by data coming from various methods (i.e., verbal protocols, eye-movement tracking, decision-making response times, etc.; see also Shiv and Fedorikhin's (1999) study for supporting results). There is some support for the cognitive effort hypothesis in the data from our Study 2. Specifically, our participants spent more time in the rejection condition (M = 31.6 s, SE = 1.0) than in the choose condition (M = 28.5 s, SE = 0.8; $t[1558] = 2.34$, $p < .05$).

7.1. Real-world situations: is it about choosing or rejecting?

The decisions that people make are influenced by the framing of that decision as a choice or a rejection (see also, Tversky & Kahneman, 1981 regarding gains or losses). Is this relevant in the case of daily food choices? One could propose that daily food decisions are always *choices* and not *rejections.* Perhaps under conditions more natural than those of our studies, individuals ask themselves to choose between, say, spaghetti Bolognese (an impoverished option; all nutrients orange) and chicken nuggets with salad (enriched option; red in fat and salt, green in saturates and sugar). Perhaps, they do not ask themselves, "which of these two dishes will I reject?"

Nagpal, Lei and Kahre (2015) argue that both frames can be at work in their study which examines the impact of the choose/reject frame on the total number of healthy/unhealthy food items purchased, and on the healthy/unhealthy ratio of food items purchased. As a real-world example, they note that "Subway allows customers to build a sandwich by *choosing* from a range of healthy (e.g., cucumber, tomato) and unhealthy (e.g., cheese, ham) items. Alternatively, Dominos offers pre-prepared pizzas consisting of healthy (e.g., mushroom, onion) and unhealthy (e.g., pepperoni, bacon) toppings but allows customers to customize it by *rejecting* the items they do not like (…)." (Nagpal, Lei, & Kahre 2015, p. 422; see also Levin, Prosansky, Heller, and Brunick (2001) for examples outside the food domain).

Furthermore, Colombo et al. (2002) state that whenever customers have to state their preference by deciding what they do not want, they are in the reject frame. For example, when they know that they do not want to eat spaghetti with meat.

Therefore, the question should be whether the power of framing can be used by choice architects (i.e., a health coach) to get people to make healthier food choices, especially with regard to the obesity crisis. We illustrate our affirmative answer with a hypothetical scenario (for simplicity's reasons, we move from the nutrient level to a more general level). Imagine, a health coach guiding his client in the choice between spaghetti Bolognese, a dish of average health and medium price and either a healthy, but expensive sushi or an unhealthy, but cheap hamburger. In this example, the price (expensive, medium, cheap) might be analogous to the sugar content (red, orange, green), while the feature healthiness might be analogous to the less effective salt content (red, orange, green) value of our enriched (red/green) food items in Study 1 and 2. How could the health coach steer the choice? Let us first consider the situation where the client can choose between spaghetti Bolognese and sushi, but does not prefer sushi. Here, the coach might be most successful promoting a sushi choice by invoking a reject frame. However, if the choice is between spaghetti Bolognese and a hamburger and the client prefers hamburgers, the coach can reduce the likelihood of that hamburger choice by invoking a rejection framework. If, however, the key feature is not price, but health, then choose frames should

be used instead of reject frames.

While our preliminary results may point to such real-life application, further experiments are needed. One unanswered question is, for example, whether framing influences not only the choice between an impoverished and an enriched food item, but also between two positive (i.e. all nutrients = green) or two negative (i.e. all nutrients = red) ones, respectively how this could look like? For example, in Perfecto, Galak, Simmons, and Nelson (2017) study, participants had to choose/reject one of two words, both of which were either positive (joy vs. kiss) or negative (murderer vs. tumor) valence. Although the authors did not focus on the choice pattern itself, they found that a match in features – that is, either a choice between two positive words or a rejection between two negative words - had an impact on confidence (i.e., participants' confidence in their choice) and perceived consensus (i.e., the percentage of people who would make the same choice). They found that in case of a "match", confidence and perceived consensus in their choice increased. This is of importance, among other things, since, for example low confidence can lead to people actively avoiding making a decision at all (see Botti & Iyengar, 2006).

Another unanswered question is, whether the choose/reject effect found for food items can be mitigated/enhanced by combining it with other factors. For example, in Colombo, Nicotra and Marino's (2002) study participants had to choose/reject between an impoverished and enriched item (i.e., cars). In some conditions they added a *decoy* (a slightly worse enriched or a slightly worse impoverished option). They assumed that in the case of a low overall preference, the selection/rejection effect should disappear respectively in the case of a high overall preference, be strengthened. Their hypothesis was, however, only confirmed in the case of high overall preference.

Nevertheless, based on their study, we would expect that by supplementing the choice between "spaghetti Bolognese" (here the overall preferred one) and "chicken nuggets with salad" with a slightly worse option (e.g. spaghetti Bolognese with pork instead of veal), the following would happen: Presumably, more people would eat spaghetti Bolognese with veal in a choice frame (compared to the case without the decoy). And, presumably, fewer people would end up eating the Bolognese in a reject frame. Thus, future studies in the food context should look closer into the combination of decoys and decision.

In sum, with Colombo, Nicotra, and Marino's (2002) words, people's food item choices are not "simply a matter of following one's clear preferences on the basis of the characteristics of a specific option" (p. 21), an assumption of the *rational theory*.

Why? First, when calculating the value or utility of the decision options, people do not take into account all the features listed (e.g., sugar, salt, etc. content). Rather, as can be seen in our studies, they seem to focus on one or two features, such as the sugar and the energy content. Secondly, the features are evaluated/weighted differently depending on the context. In our experiments, it is the choose/reject framing that changes the weights. Last, but not least, in our work, the effect of frame seems to be more in accord with the accentuation hypothesis as opposed to the compatibility principle. Taken together these findings might help to understand the framing effect, and hopefully leads to future experiments' aim to develop interventions to fight the current obesity crisis.

Declaration of interest

None.

Author note

Ester Reijnen, School of Applied Psychology, ZHAW Zurich University of Applied Sciences; Swen J. Kühne, School of Applied Psychology, ZHAW Zurich University of Applied Sciences; Markus Stöcklin, Faculty of Psychology, University of Basel, Jeremy M. Wolfe, Department of Surgery, Brigham and Women's Hospital & Department of Ophthalmology, Harvard Medical School.

References

Botti, S., & Iyengar, S. S. (2006). The dark side of choice: When choice impairs social welfare. *Journal of Public Policy and Marketing, 25*, 24–38.

Cecchini, M., & Warin, L. (2016). Impact of food labelling systems on food choices and eating behaviours: A systematic review and meta-analysis of randomized studies. *Obesity Reviews, 17*(3), 201–210.

Chen, J., & Proctor, R. W. (2017). Role of accentuation in the selection/rejection task framing effect. *Journal of Experimental Psychology: General, 146*(4), 543.

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.

Colombo, L., Nicotra, E., & Marino, B. (2002). Preference reversal in decision making: The attraction effect in choice and rejection. *Swiss Journal of Psychology, 61*(1), 21.

Department of Health (2013). Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets. Retrieved from https://www.gov.uk/.

Dutch Health report (2006). Richtlijnen goede voeding 2006 (guidelines healthy nutrition 2006). Retrieved from http://www.gezondheidsraad.nl/.

Ferdman, R. A. (2015). *Where people around the world eat the most sugar and fat.* February 5 The Washington Post. Retrieved from https://www.washingtonpost.com.

Food Standards Agency. Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets. (2016). Retrieved from https://www.gov.uk/government/publications November 8.

Franckle, R. L., Levy, D. E., Macias-Navarro, L., Rimm, E. B., & Thorndike, A. N. (2018). Traffic-light labels and financial incentives to reduce sugar-sweetened beverage purchases by low-income Latino families: A randomized controlled trial. *Public Health Nutrition, 21*(8), 1426–1434.

Ganzach, Y. (1995). Attribute scatter and decision outcome: Judgment versus choice. *Organizational Behavior and Human Decision Processes, 62* 113-113.

Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptive toolbox.* MIT press.

Hieke, S., & Wilczynski, P. (2012). Colour Me in – an empirical study on consumer responses to the traffic light signposting system in nutrition labelling. *Public Health Nutrition, 15*(5), 773–782.

Hu, F. B. (2013). Resolved: There is sufficient scientific evidence that decreasing sugar-sweetened beverage consumption will reduce the prevalence of obesity and obesity-related diseases. *Obesity Reviews, 14*(8), 606–619.

Johnson, R. J., Segal, M. S., Sautin, Y., Nakagawa, T., Feig, D. I., Kang, D. H., et al. (2007). Potential role of sugar (fructose) in the epidemic of hypertension, obesity and the metabolic syndrome, diabetes, kidney disease, and cardiovascular disease. *American Journal of Clinical Nutrition, 86*(4), 899–906.

Kiesel, K., & Villas-Boas, S. B. (2013). Can information costs affect consumer choice? Nutritional labels in a supermarket experiment. *International Journal of Industrial Organization, 31*(2), 153–163.

Kiszko, K. M., Martinez, O. D., Abrams, C., & Elbel, B. (2014). The influence of calorie labeling on food orders and consumption: A review of the literature. *Journal of Community Health, 39*(6), 1248–1269.

Levin, I. P., Prosansky, C. M., Heller, D., & Brunick, B. M. (2001). Prescreening of choice options in 'positive'and 'negative'decision-making tasks. *Journal of Behavioral Decision Making, 14*, 279–293.

Luger, M., Lafontan, M., Bes-Rastrollo, M., Winzer, E., Yumuk, V., & Farpour-Lambert, N. (2017). Sugar-sweetened beverages and weight gain in children and adults: A systematic review from 2013 to 2015 and a comparison with previous studies. *Obesity Facts, 10*(6), 674–693.

Nagpal, A., Lei, J., & Khare, A. (2015). To choose or to reject: The effect of decision frame on food customization decisions. *Journal of Retailing, 91*(3), 422–435.

NCD Risk Factor Collaboration (2016). Trends in adult body-mass index in 200 countries from 1975 to 2014: A pooled analysis of 1698 population-based measurement studies with 19·2 million participants. *The Lancet, 387*(10026), 1377–1396.

Page, I. H., Allen, E. V., Chamberlain, F. L., Keys, A., Stamler, J., & Stare, F. J. (1961). Dietary fat and its relation to heart attacks and strokes. *Circulation, 23*(1), 133–136.

Perfecto, H., Galak, J., Simmons, J. P., & Nelson, L. D. (2017). Rejecting a bad option feels like choosing a good one. *Journal of Personality and Social Psychology, 113*, 659.

Scarborough, P., Matthews, A., Eyles, H., Kaur, A., Hodgkins, C., Raats, M. M., et al. (2015). Reds are more important than greens: How UK supermarket shoppers use the different information on a traffic light nutrition label in a choice experiment. *International Journal of Behavioral Nutrition and Physical Activity, 12*(1), 151.

Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & Cognition, 21*, 546–556.

Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research, 26*(3), 278–292.

Siervo, M., Montagnese, C., Mathers, J. C., Soroka, K. R., Stephan, B. C., & Wells, J. C. (2014). Sugar consumption and global prevalence of obesity and hypertension: An ecological analysis. *Public Health Nutrition, 17*(3), 587–596.

Slovic, P., Griffin, D., & Tversky, A. (1990). Compatibility effects in judgment and choice. In R. M. Hogarth (Ed.). *Insights in decision making: A tribute to Hillel J. Einhorn* (pp. 5–27). Chicago, US: University of Chicago Press.

Statista (2017). *Pro-Kopf-Konsum von Zucker in ausgewählten Ländern weltweit von 2013/2014 bis 2016/2017 (in Kilogramm).* Retrieved from https://de.statista.com/.

Stubbs, J., Ferres, S., & Horgan, G. (2000). Energy density of foods: Effects on energy intake. *Critical Reviews in Food Science and Nutrition, 40*(6), 481–515.

Swartz, J. J., Braxton, D., & Viera, A. J. (2011). Calorie menu labeling on quick-service

restaurant menus: An updated systematic review of the literature. *International Journal of Behavioral Nutrition and Physical Activity, 8*(1), 135.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453–458.
Viswanathan, M., Torelli, C. J., Xia, L., & Gau, R. (2009). Understanding the influence of literacy on consumer memory: The role of pictorial elements. *Journal of Consumer Psychology, 19*(3), 389–402.
Wedell, D. H. (1997). Another look at reasons for choosing and rejecting. *Memory & Cognition, 25*(6), 873–887.
World Health Organization (2015). Guideline: sugars intake for adults and children. Retrieved from https://www.who.int.
Yudkin, J. (1964). Dietary fat and dietary sugar in relation to ischaemic heart-disease and diabetes. *Lancet, 2*, 4–5.

Journal of Consumer Policy
https://doi.org/10.1007/s10603-019-09413-4

SHORT COMMUNICATION

Nudged to a Menu Position: The Role of "I'm Loving It"!

E. Reijnen[1] · S. J. Kühne[1] · H. M. von Gugelberg[1] · A. Crameri[1]

Received: 3 September 2018 / Accepted: 19 March 2019/Published online: 25 May 2019

Abstract

The *position* of food items on supermarket shelves or dishes on restaurant menus seems to influence consumers' choices. However, it is still unclear which position is the most favourable, respectively which factor can explain the variety of different position effects observed (e.g., centre-stage effect). We assume that this factor is based on whether or not "your love" (or your preference) for the cuisine of the restaurant (e.g., Italian) where you have dinner plays a role regarding your dish choice or not. Hence, in a computer-based study, participants had to choose dishes, appetizers, entrées and desserts from menus whose cuisine they most (e.g., Italian) or least preferred (e.g., Japanese). We found that regarding the meal type entrées, preference indeed played a role. Regarding menus from their most preferred cuisine, participants chose significantly more often entrées positioned in the centre of the menu. No such effect could be found regarding menus from their least preferred cuisine. Regarding the meal type appetizer, preference did not seem to play a role; hence, participants did, regarding both preferences, choose more appetizers positioned at the top of the menu. Regarding desserts, no effects could be found. A developed theoretical framework tries to illustrate how preference comes into play, by changing the way the dishes within a meal type are perceived, and hence modulates the different position effects observed. The framework should provide choice architects with guidelines about where they could place healthier dishes on a menu to fight the current overweight and obesity crisis.

Keywords Position effects · Preference · Cuisine · Choice · Menu

"Eternally lures the pizza" (Six 2015) – though formulated in a catchy language – headlines like these want us to draw attention to our *unhealthy eating behaviour*, leading to the current overweight and obesity crisis. A crisis that not only costs human lives (due to secondary diseases such as cardiovascular disease, diabetes, cancer, etc.) but also generates enormous economic costs (alone 7,990 mio. Swiss francs per year in Switzerland; Schneider and Venetz 2014).

✉ E. Reijnen
ester.reijnen@zhaw.ch

[1] School of Applied Psychology, ZHAW Zurich University of Applied Sciences, Pfingstweidstrasse 96, Postfach, 8037 Zurich, Switzerland

Since eating out is increasing (Rudolph and Bassett 2015), restaurants might be a target to counter the current *overweight and obesity crisis*. Especially since eating out seems to be associated with "lower diet quality, greater energy intake, and higher body weight" (Long et al. 2015, p. e11). The question is which strategies could be used by restaurants to make their customers eat healthier? They could simply *ban* unhealthy dishes from their menu or make them *more expensive*. However, those strategies could probably only be implemented by the government, since restaurant owners have stated their concern that they might lose customers by, for example, banning unhealthy dishes from their menus (Estrade et al. 2014; Lombardini and Lankoski 2013). Another strategy that restaurants could follow is to complement dishes with helpful health information such as nutrition facts (Long et al. 2015). This would be in line with people's desire to be informed or educated rather than subjected to governmental strategies such as taxation (gfs.bern 2014). Alas, that strategy assumes that customers make *rational* dish choices. That requires not only that they are aware of all the relevant information but also that they integrate these information properly in order to choose the dish with the highest (subjective or objective) value (e.g., the healthiest one). Research, however, has repeatedly shown that this not how people choose, mainly due to human processing capacity limits. Hence, that strategy might not be the most successful one, and a conclusion also reached in the review and meta-analysis paper by Long et al. (2015). Therefore, over the last years, another strategy called "nudging" has become popular (see Sunstein 2014 for a brief summary). Since people decide *intuitively* instead of *rationally* (Thaler and Sunstein 2009), it is assumed that by changing people's *choice architecture*, their choices can be improved not only to their own benefit but also to the benefit of the whole society. Such a change could be, for example, to simplify the aforementioned nutrition facts by using labels or health claims (Bauer and Reisch 2019; Kozup et al. 2003). Or, as in the context of this study, it could be a change of the *position of the dishes* or items in general (Bucher et al. 2016). Importantly, nudging – as opposed to the banning or taxing strategy – does not forbid any options (the unhealthy hamburger is still on the menu), nor does it significantly change them through economic incentives (the hamburger is not made more expensive; Thaler and Sunstein 2009). Hence, customer freedom is not at risk (Sunstein 2015).

The idea that changing the position of items is an effective strategy for behaviour change has been found in different areas of our lives, from *medical treatment decisions* and *leisure activities* (i.e., gambling), to daily habits such as choosing *food items* from supermarket shelves to buffets (e.g., breakfast) and food trays.

Christenfeld (1995), for example, found that customers seem to favourably pick, from horizontally arranged *identical* food items (e.g., canned tomatoes) on the supermarket shelf, the *middle* ones. Similarly, Keller et al. (2015) found that when participants had to choose between *different* food items (i.e., cereal bars), shifting the apple bar from the left (out of three *horizontal*) to the middle position increased its choice from 13.3% to 36.7% (Rozin et al. 2011 for a similar finding for *vertically* positioned food items). However, Wansink and Hanks (2013) as well as Rozin et al. (2011) could not replicate this finding of a preference for the middle option or the so-called "centre-stage effect." When having to choose between likewise different food items (e.g., different ingredients at a salad bar such as chicken, egg, tomatoes, etc.), their customers did choose most often items from the first (primacy or *top-choice effect*) or the last position[1] (recency-effect). Finally, some authors did not find a position effect at all (van Kleef et al. 2012).

[1] or positions.

More importantly, only one study has investigated position effects in *restaurant menus* (Dayan and Bar-Hillel 2011). Unlike the food item choices in the studies mentioned above (with the exception of Keller et al. 2015), dish choices from restaurant menus do not require any physical effort on the part of the customer. However, the challenge of the study by Dayan and Bar-Hillel (2011) is that, depending on the meal type (i.e., appetizers, entrées or desserts), different position effects have been found. A significant position effect could only be found regarding *appetizers*; appetizers placed at the beginning or at the end of a menu seemed to be chosen more often than the ones in the middle. However, no such effect could be found for the entrées and the desserts. Considering that entrées are probably the centrepiece when you are dinning out, these results are not very promising regarding the possibility to nudge customers towards certain positions to make them, for example, choose healthier dishes.

Hence, it is a major challenge to develop a theoretical framework of factors that can explain the different position effects found so far in food position studies (Bar-Hillel 2015; see Hollands et al. (2017) for the more general TIPPME concept regarding possibilities of altering physical micro-environments to change health-related behaviour). Especially since, based on the differences in the basic settings of those studies, such as the identity of the items (identical vs. different or non-identical) and the arrangement of the items (vertical vs. horizontal), all of these differences could come into question as factors.

We think that the critical factor is whether the individual items of a set are physically *identical* or *different*/*non-identical* or, as we will show later, are simply perceived that way (e.g., due to the factor preference). If we take a look at the food position studies, from this perspective, we see that choice studies (Christenfeld 1995; Valenzuela and Raghubir 2009) using *identical items* reveal a consistent tendency to choose items presented in the middle. According to Valenzuela and Raghubir (2009), this centre-stage effect is produced by people's *popularity belief*, the belief that marketers place the most popular item in the middle position, i.e., "most preferred by other consumers and has the highest market share in its category" (Valenzuela and Raghubir 2009, p. 186). On the other hand, studies using *non-identical items* seem to produce – only at first glance, though – a less clear picture, that is, different position effects. While Valenzuela and Raghubir's (2009) study, which used different flavoured chewing gums (spearmint, peppermint and winter frost), found a preference for the middle option, ballot voting studies (comparable to choosing from a restaurant menu) found that candidates listed first received a greater portion of votes than when listed at any other position (Koppell and Steen 2004; Miller and Krosnick 1998). According to Miller and Krosnick (1998), this top-choice effect is produced by the "satisficing principle"; this principle assumes that since customers want to conserve resources, they select the most accessible satisfactory item presented (even if it is not optimal). Thereby customers probably rely on their belief that items positioned at the top are the ones with the highest price and quality (Casasanto 2009; Valenzuela and Raghubir 2015). This top-choice effect has also been found in restaurant menu choices or restaurant online deliveries (Dayan and Bar-Hillel 2011; Murphy et al. 2006), however, in combination with a recency-effect (i.e., the last item on the list is chosen most often; Ert and Fleischer 2016).

Hence, while the *centre-stage effect* could be observed for identical and non-identical items, the *top-choice effect* could only be observed for non-identical items. However, we assume that the centre-stage effect can actually only be found with identical items. Then, how do we explain the found centre-stage effect for non-identical items. We assume that by some factor, the items in these studies, although *physically different* from each other, were perceived as identical (though not indifferent from each other). That factor is, from our view, *preference*,

that is, whether your desire regarding a "set of items," for example chewing gums, plays a role or not.

However, can preference affect our choices at all? In that respect, Strauss et al. (2013) found that preference changes the preferred order of items. They primed their participants with images of *positive* green objects (e.g., ripe kiwi) and *negative* red objects (e.g., infected eye), respectively. That is, both "positive" or "negative" could be associated with both "green" or "red." After being primed, participants had to state which colour, green or red, they preferred (to cover the true intention of the experiment, these two colours were embedded in a number of other colours). Strauss et al. (2013) found that participants' preference for the colours green and red changed depending on the priming condition (going from preferring green over red to red over green, respectively). However, that result could only be observed when their participants had to rate the images of the positive and negative objects (e.g., ripe kiwi) regarding whether they preferred them or not, but not when they simply had to click on the centre of these objects. That is, "positive/negative" had to be connected with you "preferring/not preferring the objects," a crucial point, especially for later. Preference, however, seems also to be able to induce position effects, as shown by Li and Epley (2009), yet in a non-food or non-menu study. In one of their experiments, participants had – after being primed with either good ("preferred," e.g., Miss Universe contestants) or bad ("not preferred," e.g., women with craniofacial syndrome) objects – to choose between three non-identical objects (paintings or female faces). They found that participants (in their *simultaneous* condition) favoured the *middle option* when primed with the good stimuli, while they favourably selected the last option when primed with the bad stimuli. Since their main focus was on explaining position effects when objects were presented in a sequential order, they did not run any statistical analysis[2] over the above-mentioned results nor did they explain those results in detail. They just mentioned that the effect depended on whether the overall choice set was preferred or not. Similarly, Rodway et al. (2016) found, in a study where their participants had to either choose the most or least preferred pictures in a row of five similar ones (e.g., butterflies), a centre-stage effect for the most preferred (in vertical and horizontal arrangements) pictures, respectively no effect for the least preferred pictures. Hence, when preference seems to play a role, a centre-stage effect can be observed for the most preferred items, respectively no effect for the least preferred items, whereby preference makes the items of the preferred item set identical towards each other ("all are equally loveable").

We, therefore, hypothesize that when your *preference* or love for a set of items – in our case the dishes of a certain menu type – plays a role, a centre-stage effect (mediated by the popularity belief) for the preferred set of items (e.g., Italian cuisine) should occur, since the dishes per meal type (e.g., entrée) became (evidently) identical, just, for example, nine Italian entrées, and not *Spaghetti alla Bolognese*, *Gnocchi al Forno* and so forth.

That, however, should only be the case regarding the meal type entrée (not for the meal types appetizers and desserts), since entrées are the probable cause for having established a preference for a given cuisine in the first place. Formally:

H1. Regarding the meal type entrées, we expect a centre-stage effect for the most preferred cuisine, respectively no effect for the least preferred cuisine.

[2] Rodway et al. (2016) while running the missing analysis in the Li and Epley (2009) study found indeed a significant centre-stage effect for good, but not for bad items.

On the other hand, when preference plays no role, a top-choice effect (if at all) – based on the satisficing principle – should be observed, since the items of the set are non-identical. This should be the case for appetizers and desserts.

In that vein, Pereira and Villodres (2002) have shown that when the political party – democratic or republican did not play a role, candidates presented at the top of the voting list were preferred, probably based on the belief that the most capable candidates are listed at the top. Hence, we hypothesize that, regarding appetizers and desserts, where preference should not play a role, a top-choice effect for both the most preferred and the least preferred *cuisine* should be observed (at least if the individual items—as in the party analogy – play a role). Formally:

H2. Regarding the meal type appetizers and desserts, we expect a top-choice effect for both the most and the least preferred cuisine.

Method

Participants

Forty-five participants (53% female) of the Zurich University of Applied Sciences, ZHAW, and the Zurich community took part in the study. Participants' ages ranged from 18 to 62 years ($M = 32.8$; SD = 12.06). All participants gave written informed consent.

Apparatus

The experiment was conducted on Apple Mac computers (Apple Inc., USA) on 20-in. CRT monitors (hp p1230, Hewlett-Packard Company L.P, USA) with a screen resolution of 1024 × 768 pixels and 100 Hz. The experiment was programmed with the MATLAB Software and presented by means of the Psychtoolbox. The participants were seated 57 cm in front of the monitor. The experiment took place in a dark and silent room.

Material

For the cuisines, Italian, Asian, Mexican, Japanese and American menus were created whereby each of them contained three meal types: appetizers (six dishes), entrées (nine dishes) and desserts (six dishes). Each meal type (appetizers, entrées, desserts) could be divided into top, middle and bottom positions, or "area" how we call it (we summarized two respectively three positions into one area). For each cuisine, seven completely different menus (regarding the dishes they contain) were created (i.e., only one of them contained the "Spaghetti alla Bolognese"), symbolizing seven different Italian restaurants. Subsequently these seven entire menus (with appetizers, entrées and desserts) were tripled (accompanied by a layout change). From one triplet to the other, dishes were moved from the top to the middle, from the middle to the bottom, and from the bottom to the top area. For example, if the entrée dish Spaghetti alla Bolognese has been positioned in the top area in the first menu of the triplet, it then was moved to the middle area in the second menu of the triplet and to the bottom area in the last menu of the triplet. This resulted in a total of 21 menus per cuisine (see Fig. 1 for examples). Menus did not include prices.

Fig. 1 **a**) Examples for the Mexican cuisine. **b**) Examples for the Japanese cuisine

Procedure and Design

First participants had to choose their *most* and *least* preferred cuisine (Italian, Asian, Mexican, Japanese, American) before they completed five practice and 42 experimental trials.[3] Each trial started with a fixation cross (500 ms), followed by a menu (either from their most or least preferred cuisine), whereby for each menu, the appetizers were presented first, followed by entrées and then desserts. Each meal type was presented on the screen until the participant made his or her choice. This resulted in a *preference* (most/least preferred cuisine) × *area* (top, middle, bottom) design for each of the three meal types (appetizers, entrées and desserts).

Results

Overall, one participant had to be removed from the analysis for dubious behaviour. As can be seen in Table 1, Italian was the most preferred cuisine followed by the Asian cuisine. On the other hand, Japanese was the least preferred cuisine followed by the American cuisine. No clear preference for the Mexican cuisine emerged.

Table 2 shows the percentage of participants that did choose a dish from the upper, middle or bottom area, broken down by hypothesis, meal type and preference. Data were – based on our hypothesis – analysed separately per *meal type* (starting with the entrées, followed by the appetizers and desserts).

Entrées

Trials with choices made in less than 1,000 ms or more than 55 000 ms (1.2%) as well as trials in which a participant ordered the same dish in the appropriate triplet (15.6%) were removed from the analysis.

[3] Forty-two experimental trials/participant: 7 (number of menus per cuisine) × 3 (dish area: top, middle, bottom) × 2 (cuisine: most/least preferred). Since this applies to all three meal types (appetizers, entrées, desserts), we end up with 125 choices per participant.

Table 1 Number (percentage) of participants choosing the most and least preferred cuisine

Preference	Cuisine					Total
	Asian	Italian	Japanese	Mexican	American	
Most	147 (16%)	609 (66%)	42 (5%)	63 (7%)	63 (7%)	924 (100%)
Least	42 (5%)	21 (2%)	609 (66%)	42 (5%)	210 (23%)	924 (100%)
Total	189 (10%)	630 (34%)	651 (35%)	105 (6%)	273 (15%)	1,848 (100%)

We tested the effects of the factors preference and area by means of the chi-square test for one-way tables. Overall, no significant effects were found for preference, χ^2 (1, $N = 1{,}538$) = 0.010, $p = .919$, and area, χ^2 (2, $N = 1{,}538$) = 3.657, $p = .161$. However, a closer look at the data (single comparisons) showed that in the most preferred condition, participants did choose more dishes from the middle than the top, χ^2 (1, $N = 535$) = 5.250, $p < .05$, and bottom area, χ^2 (1, $N = 535$) = 7.308, $p < .01$. No such results were found for the least preferred condition (all χ^2 values < 0.124, all p values > .725).

Since some studies question the stability of preferences (see Mueller and Szolnoki 2010), we analysed the data separately for the first and second half of the trials. While no advantage for the middle area in the most preferred condition could be found for the first half, all χ^2 values < 2.510, all p values > .113, a middle preference (see Fig. 2) could be observed for the second half of the trials (all χ^2 values > 4.942, all p values < .026).[4]

Appetizers and Desserts

Appetizers

Again, trials with choices made in less than 1,000 ms or more than 45 000 ms (0.9%) as well as trials in which a participant ordered the same dish in the appropriate triplet (18.7%) were removed from the analysis.

Overall, we found no significant effect for preference, χ^2 (1, $N = 1{,}489$) = 0.151, $p = .698$, but one for area, χ^2 (1, $N = 1{,}489$) = 13.259, $p < .01$. Single comparisons showed that participants in both preference (most and least) conditions chose more appetizers from the top than the bottom area (most preferred: χ^2 (1, $N = 496$) = 5.879, $p < .05$; least preferred: χ^2 (1, $N = 522$) = 5.180, $p < .05$). In the least preferred condition, we additionally found that more appetizers were chosen from the top area compared to the middle area, χ^2 (1, $N = 517$) = 6.284, $p < .05$.

Again, we compared the first half of the trials with the second half to determine the origin of the effect. That analysis revealed that the observed effects could only be found in the first half of the trials (trials 1–21; see Fig. 3). That is, participants did choose more dishes from the top area compared to middle and bottom area in both preference conditions (all χ^2 values > 5.015, all p values < .05). Interestingly, all effects disappeared in the second half of the trials (22–42).

[4] See the discussion section for a different result regarding the development of the centre-stage effect in the most preferred condition.

Table 2 Number (percentage) of menus selected per meal type and area

Hypothesis	Meal type		Area			
			Top	Middle	Bottom	Total
H1	Entrées					
	Preference	Most	241 (31%)	294 (38%)	232 (30%)	767 (100%)
		Least	255 (33%)	254 (33%)	262 (34%)	771 (100%)
		Total	496 (32%)	548 (36%)	494 (32%)	1,538 (100%)
H2	Appetizers					
	Preference	Most	275 (37%)	241 (33%)	221 (30%)	737 (100%)
		Least	287 (38%)	230 (31%)	235 (31%)	752 (100%)
		Total	562 (38%)	471 (32%)	456 (31%)	1,489 (100%)
	Desserts					
	Preference	Most	246 (33%)	248 (33%)	250 (34%)	744 (100%)
		Least	244 (32%)	245 (33%)	264 (35%)	753 (100%)
		Total	490 (33%)	493 (33%)	514 (34%)	1,497 (100%)

Desserts

Again, trials with choices made in less than 1,000 ms or more than 30 000 ms (1.7%) as well as trials in which a participant ordered the same dish in the appropriate triplet (17.4%) were removed from the analysis.

Overall, no significant effect of preference, χ^2 (2, $N = 1497$) = 0.054, $p = .816$ and area, χ^2 (2, $N = 1497$) = 0.685 $p = .710$, was found. An equivalent number of dishes were ordered from each area whether the cuisine was the preferred or not. Additionally, no single comparisons (between the different areas, separately for the preference conditions) reached significance (all χ^2 values < 0.709, all p values > .400).

The same pattern of results was found for the first (all χ^2 values < 2.420, all p values > .120) and second half of the trials (χ^2 values < 1.536, all p values > .215).

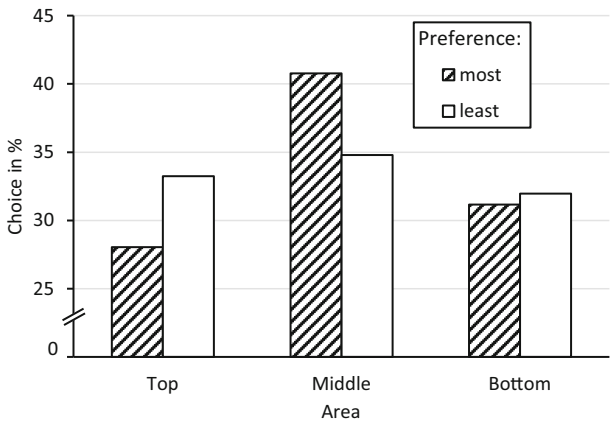


Fig. 2 Entrée choices in percentages for the second half (22–42) of the trials

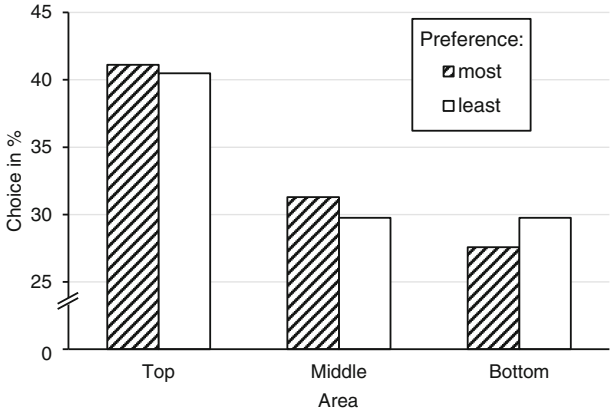


Fig. 3 Appetizer choices in percentages for the first half (1–21) of the trials

Discussion

In summary, regarding the meal type *entrées*, preference seemed – as predicted by our hypothesis H1 – to play a role. Hence, dishes positioned in the middle area were chosen most often (an advantage of about 24%) in the most preferred cuisine, but not in the least preferred cuisine. That "centre-stage" effect, however, could only be found in the second half of the trials. Regarding the *appetizers*, however, preference did not play a role. Hence, dishes from the top area were chosen most often (as predicted by our H2); this independent of whether the cuisine was the most or least preferred one. This top-choice effect could, however, only be found in the first half of the trails. For the meal type *desserts*, where preference also played no role; however, no such top-choice effect could be found.

Last but not least, about one sixth of our participants did, independent of the meal type, choose the same dish over and over again (e.g., the lasagna). This result is in line with a study described by Parker (2017), where one sixth of the participants (workers) reported to have eaten the same lunch every day for the last six years. They did so even though they were bored with their own lunch choice. Future research must reveal whether there could be potential nudging techniques to even change these people's choices in a beneficial way (i.e., healthier).

Based on our results, we propose the following theoretical framework to be considered as a guideline for choice architects (see Fig. 4). In support of the model, different generalized linear models (GLMs) were calculated (the appropriate results are displayed in the footnotes 5, 7 and 8). The GLM results thereby confirm (with one exception) the results presented before. This illustrated the robustness of our results despite running different analyses.

The key factor in the framework is whether your preference towards "a set of items" plays a role or not (i.e., different position effects are produced between the most and least preferred set of items, here cuisine).

Whenever that could potentially be the case, choice architects should position the items to be sold (e.g., the healthiest dish) in the *middle* or centre[5] of a set of items (e.g., entrées). Strictly taken, one could argue that the centre-stage effect – if preference plays a role – is only found regarding the preferred, but not for the unpreferred, set of items. Which is true, but is it not the case that customers

[5] The GLM with Poisson distribution regarding the entrées a significant effect for the middle area in the most preferred condition $z = 1.96$, $p < .05$. However, no other factor reached significance; all z's < 1.60 and all p's $> .10$.

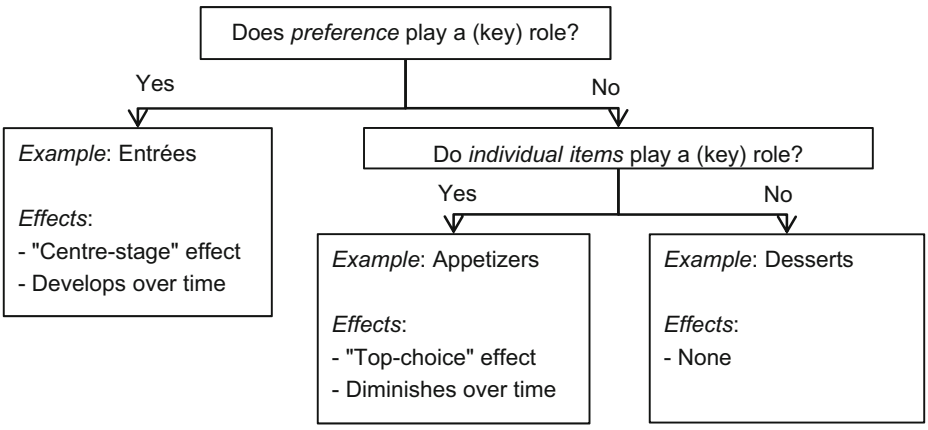


Fig. 4 Theoretical framework for menu choices as a "guideline for choice architects"

usually also buy only from the preferred set of items, but not from the unpreferred one? For example, if you need a laptop and you love Apple products, you are going to choose between MacBooks, but not between HP or ASUS laptops. Remember, by preference respectively when preference plays a role, the items of the preferred set are going to be perceived as *identical*, hence "everything on the menu is so yummy"! Then based on the popularity belief, customers choose the item in the centre most often. Similarly Bar-Hillel (2015) propose that that centre-stage effect requires no individual item processing.

This framework should also hold for other choice sets, for example, *ballot voting*. In that case, we would expect that if you have to vote for a candidate of your preferred party (e.g., democrats), you choose a candidate placed in the middle of the list most often, since all candidates have – by preference – become equal, for example, regarding their competence.

However, as we have seen in restaurant menus, preference for a certain cuisine did not play a role regarding appetizers and desserts. That might also be the case in ballot voting, for example, when you also have to vote for a candidate from the "young" democrats. Based on the belief that the best items are placed at the top, items from those positions should be chosen most often (e.g., top-choice effect).

These examples show also that further research is needed to define the categories where preference might usually play a role and where not. Maybe it is not whether you have to choose within, for example, the "regular" democrats (i.e., preferred Italian entrées) respectively the young democrats (i.e., preferred Italian appetizers), but maybe whether you have to vote for candidates for the National or Local Council.

Further, since we could only observe the top-choice effect regarding appetizers and not desserts, we assume that that effect is mediated by the factor of whether the *individual items* of the set play a role. Remember, in both cases (appetizers and desserts), the items are *non-*

[0] Reaction times (RTs) were significantly slower for the meal type appetizer (M = 12 289 ms, SD = 7909 ms) than for the meal type desserts (M = 9293 ms, SD = 5836 ms), $t(2984) = 11.782$, $p < .001$. This could indicate that participants cared about the individual dishes in the meal type appetizers, but not in the meal type desserts.
[0] The GLM with Poisson regarding appetizers showed a significant effect for the top area $z = 3.97$, $p < .001$, and a significant decrease in choices in the top area for the second half of the trials; $z = -2.315$, $p < .05$. No other factor reached significance; all z's < 1.22 and all p's > .22.
[0] Regarding desserts no factor reached significance in the GLM; all z's < 0.89 and all p's > .37.

identical. Hence, the question is do customers care whether they choose a "good" (here good does not mean popular) item or not. If they do (as regarding the appetizers[6]), based on people's belief that items placed at the top are associated with a higher price or quality than the ones placed at the bottom (Valenzuela and Raghubir 2015), a top-choice effect should be observed. As we have seen, this effect should diminish over time.[7] If, on the other hand, they do not, no effect (as with the desserts[8]) should be observed. Going back to the ballot voting example, if preference for a political party does not play a role, yet voters want to choose the best candidate available on the list, they should choose the one on the top of the list. If not, no position should be favoured.

To wrap it up, we assume that Dayan and Bar-Hillel (2011) as well as, for example, Kim et al. (2019) did not find a centre-stage effect since they did not take *preference* into account.

Limitations and the Non-replicability Crisis of This Type of Research

Limitations We did not manipulate our dishes regarding their healthiness. Hence, further research must be conducted in order to test whether, for example, the observed centre-stage effect for entrées still arises when those vary in healthiness (e.g., the healthy entrées is placed in the middle). Furthermore, it might be interesting to see whether adding prices to the dishes would change the observed choice patterns. Since one does not usually eat for free, and the willingness to pay can vary greatly among people, price might be a factor that needs to be added in future restaurant menu studies.

Non-replicability Crisis of This Type of Research The different position effects found in the food domain might add to the expressed concern about the non-replicability of psychological study results (Pashler and Wagenmakers 2012). However, what does non-replicability mean after all? Stroebe and Strack (2014) thereby distinguish between *exact* and *conceptual* replications. In exact replications, the question is whether the same study results can be observed when the *independent and dependent variables* of the study to be replicated are exactly reproduced (Maxwell et al. 2015; Stroebe and Strack 2014). To our knowledge, there exist no such studies in the area of food item positioning. Anyhow, of more scientific value are – according to the authors – *conceptual replications*. They try to answer the question whether a different operationalization of the theoretical constructs that underlie the independent and dependent variables of the study to be replicated leads to the same study results. Since the studies mentioned so far all differ, especially in their independent but also in their depended variables, we can hardly consider these studies as conceptual replications. Only if another study, for example, operationalizes our preference concept differently and cannot replicate our findings, we can talk about a failed conceptual replication. And even then, it is premature to speak of a replication crisis without considering, for example, possible mediating variables.

As described, although at first glance at the literature, there seems to be a broad variety of results, they can be explained by the theoretical framework stated above.

Overall, we hope that our research did provide clarification of why different position effects can be found in the food domain. Furthermore, we hope that our theoretical frameworks provide a valuable guideline for choice architects.

Acknowledgments We would like to thank Daniel Furter, Susan Keller and Michelle Petrig for helping to design the menus as well as for helping run the experiments.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

Bar-Hillel, M. (2015). Position effects in choice from simultaneous displays: A conundrum solved. *Perspectives on Psychological Science, 10*, 419–433.

Bauer, J. M., & Reisch, L. A. (2019). Behavioural insights and (un)healthy dietary choices: A review of current evidence. *Journal of Consumer Policy, 42*, 3–45.

Bucher, T., Collins, C., Rollo, M. E., McCaffrey, T. A., De Vlieger, N., Van der Bend, D., et al. (2016). Nudging consumers towards healthier choices: A systematic review of positional influences on food choice. *British Journal of Nutrition, 115*, 2252–2263.

Casasanto, D. (2009). Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General, 138*, 351–367.

Christenfeld, N. (1995). Choices from identical options. *Psychological Science, 6*, 50–55.

Dayan, E., & Bar-Hillel, M. (2011). Nudge to nobesity II: Menu positions influence food orders. *Judgment and Decision making, 6*, 333–342.

Ert, E., & Fleischer, A. (2016). Mere position effect in booking hotels online. *Journal of Travel Research, 55*, 311–321.

Estrade, M., Dick, S., Crawford, F., Jepson, R., Ellaway, A., & McNeill, G. (2014). A qualitative study of independent fast food vendors near secondary schools in disadvantaged Scottish neighborhoods. *BMC Public Health, 14*, 793.

gfs.bern. (2014). Monitor Ernährung und Bewegung 2014. Retrieved from http://www.gfsbern.ch. Accessed 24 Apr 2019.

Hollands, G. J., Bignardi, G., Johnston, M., Kelly, M. P., Ogilvie, D., Petticrew, M., et al. (2017). The TIPPME intervention typology for changing environments to change behaviour. *Nature Human Behaviour, 1*, 0140.

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.

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*, 116–124.

Koppell, J. G., & Steen, J. A. (2004). The effects of ballot position on election outcomes. *Journal of Politics, 66*, 267–281.

Kozup, J. C., Creyer, E. H., & Burton, S. (2003). Making healthful food choices: The influence of health claims and nutrition information on consumers' evaluations of packaged food products and restaurant menu items. *Journal of Marketing, 67*(2), 19–34.

Li, Y., & Epley, N. (2009). When the best appears to be saved for last: Serial position effects on choice. *Journal of Behavioral Decision Making, 22*, 378–389.

Lombardini, C., & Lankoski, L. (2013). Forced choice restriction in promoting sustainable food consumption: Intended and unintended effects of the mandatory vegetarian day in Helsinki schools. *Journal of Consumer Policy, 36*, 159–178.

Long, M. W., Tobias, D. K., Cradock, A. L., Batchelder, H., & Gortmaker, S. L. (2015). Systematic review and meta-analysis of the impact of restaurant menu calorie labeling. *American Journal of Public Health, 105*(5), e11–e24.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist, 70*, 487–498.

Miller, J. M., & Krosnick, J. A. (1998). The impact of candidate name order on election outcomes. *Public Opinion Quarterly, 62*, 291–330.

Mueller, S., & Szolnoki, G. (2010). The relative influence of packaging, labelling, branding and sensory attributes on liking and purchase intent: Consumers differ in their responsiveness. *Food Quality and Preference, 21*, 774–783.

Murphy, J., Hofacker, C., & Mizerski, R. (2006). Primacy and recency effects on clicking behavior. *Journal of Computer-Mediated Communication, 11*, 522–535.

Parker, R. (2017). Got a ham sandwich for lunch today? One in six workers admit eating the same thing every day for two years. *Daily Mail Online*. Retrieved from http://www.dailymail.co.uk. Accessed 24 Apr 2019.

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science, 7*, 528–530.

Pereira, J. M., & Villodres, C. O. (2002). Position effects and party nomination strategies under the limited vote: The 2000 Spanish senate election. *Representation, 38*, 304–316.

Rodway, P., Schepman, A., & Thoma, V. (2016). Reachability does not explain the middle preference: A comment on Bar-Hillel (2015). *i-Perception, 7*(2), 1–5.

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*, 323–332.

Rudolph, T., & Bassett, M. (2015). *Food Consumption 2014: Ess- und Verzehrverhalten in der Schweiz*. St. Gallen: Forschungszentrum für Handelsmanagement.

Schneider, H., & Venetz, W. (2014). *Cost of obesity in Switzerland in 2012*. Retrieved from: http://www.bag.admin.ch. Accessed 24 Apr 2019.

Six, A. (2015). Und ewig lockt die Pizza. *NZZ am Sonntag.* Retrieved from https://nzz.ch. Accessed 24 Apr 2019.

Strauss, E. D., Schloss, K. B., & Palmer, S. E. (2013). Color preferences change after experience with liked/disliked colored objects. *Psychonomic Bulletin & Review, 20*, 935–943.

Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science, 9*, 59–71.

Sunstein, C. R. (2014). Nudging: A very short guide. *Journal of Consumer Policy, 37*, 583–588.

Sunstein, C. R. (2015). Nudges do not undermine human agency. *Journal of Consumer Policy, 38*, 207–210.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth and happiness*. London: Penguin Books.

Valenzuela, A., & Raghubir, P. (2009). Position-based beliefs: The centre-stage effect. *Journal of Consumer Psychology, 19*, 185–196.

Valenzuela, A., & Raghubir, P. (2015). Are consumers aware of top–bottom but not of left–right inferences? Implications for shelf space positions. *Journal of Experimental Psychology: Applied, 21*, 224–241.

van Kleef, E., Otten, K., & van Trijp, H. C. (2012). Healthy snacks at the checkout counter: A lab and field study on the impact of shelf arrangement and assortment structure on consumer choices. *BMC Public Health, 12*, 1072.

Wansink, B., & Hanks, A. S. (2013). Slim by design: Serving healthy foods first in buffet lines improves overall meal selection. *PLoS One, 8*, e77055.

When too Few is Bad for the Environment:

Choice Set Size and Default Effects for Electricity Products.

Swen J. Kühne, Ester Reijnen, & Aureliano Crameri

Zurich University of Applied Sciences

Correspondence concerning this article should be addressed to Swen J. Kühne, ZHAW Zurich University of Applied Sciences, School of Applied Psychology, Pfingstweidstrasse 96, Postfach 707, CH-8037-Zürich, Switzerland. swen.kuehne@zhaw.ch

Abstract

Defaults are an effective tool in shaping consumers' decisions. However, only a few studies have investigated the role of defaults regarding consumers' choices of electricity products. Moreover, each of these studies used binary choice sets (grey vs. green electricity). Notably, decision-making research has shown that consumer choice patterns are considerably

influenced by the size of the choice set (e.g., adding a third option). The question is, does this also hold for defaults, that is, do they function differently depending on the choice set size? In our experimental study, participants could choose between 3 electricity products (grey, green and eco) which varied in their environmental-friendliness and price, with the default randomly being one of the three products. In addition, we had a no default condition. We found – contrary to the other studies – not only a default effect for the least environmentally-friendly grey product, but also for the environmentally-friendlier products green and eco electricity. Moreover, the popularity of the middle option, the green electricity product, was not reduced by adding a third product. The results indicate that increasing the set size by adding an eco-product and by intelligently setting the default, could increase the number of consumers buying environmentally-friendly electricity products.
Keywords: default, electricity products, choice set size, extremeness aversion

When too Few is Bad for the Environment:

Choice Set Size and Default Effects for Electricity Products.

"Consumers are *sometimes* naïve!" This paraphrase of Nobel laureate Richard Thaler's statement (Metzler, 2015) means that we don't always act according to the situation, or in other words environmentally-friendly, such as, for example, buying a fancy SUV with high CO_2 emission. According to Thaler, this "misbehavior" arises because we as consumers are prone to *irrelevant factors* such as *relative position*. Hence, we might have just bought that SUV because it was positioned next to an ugly, small car at the car dealership. According to the standard economic theory (e.g., "expected utility theory" [EUT] and its axiomatic characterizations, see, for example, von Neumann and Morgenstern, 1944) factors such as relative position should be irrelevant, since consumers should always choose the car or option

with the highest (here) *utility*. However, as pointed out in Thaler's statement, this is not always the case. Hence, Thaler introduced his nudging concept, offering us a list of irrelevant factors on which nudges can be developed. Nudges are interventions, that direct consumers into a certain direction (i.e., helping them to act more environmentally-friendly) while preserving their freedom of choice (see Thaler & Sunstein, 2009). One such "nudge" is the use of *default* options (automatic enrollments; Johnson et al., 2012). An example would be making the "double-sided" printing on your computer the default and thereby reducing paper waste. The default option is thus simply what we get if we do nothing (Thaler & Sunstein, 2009). Research, from the health to the financial sector, has shown that defaults can be a quite effective nudge (for a review; Beshears, Choi, Laibson, & Madrian, 2009). Nevertheless, so far in the energy sector only two studies (Ebeling & Lotz, 2015; Pichert & Katsikopoulos, 2008) have, with simple binary choice tasks, *empirically* tested whether defaults also could prompt consumers to buy more environmentally-friendly electricity products. Pichert and Katsikopoulos (2008) found, that depending on whether they made their green or grey electricity product the participant's default, the green product was chosen more or less (green default condition: 68% vs. grey default condition: 41%). However, comparing their "green default" condition to their third neutral (or no default) condition shows that there were not more participants choosing the green product in the former than in the later condition (difference: 0,5%; see the right side of Figure 2). Hence, the default had – as we call it – no "positive" effect. On the other hand, they observed a "negative default" effect in their grey default condition: the grey product, compared to the neutral condition, was chosen more often (and hence, the green product less often). Similar results were found by Ebeling and Lotz (2015), though the (absolute) difference they found in choosing the "green electricity product" between the two default options (opt-in vs. opt-out; they didn't include a neutral control

condition) was enormously high: 61.9% for the real, and 59.6% for the simulated purchase condition. This might be due to the minuscule additional costs of going green (0,3 cents / kWh).

Crucially, both studies investigated default effects only within binary choice sets. However, research has shown that by adding a third option (e.g. C), two things that were initially difficult to compare (e.g. A and B; since one product is better in one attribute, respectively the other better in the other attribute) not only become - relatively speaking - easily comparable (see Ariely, 2009), but also changes the overall choice pattern, either leading to a *compromise* or a *polarization effect* (e.g., Simonson & Tversky, 1992). According to the *compromise* effect, the middle option (e.g., B) should get relatively more popular in the context of the triple (A, B, C) than in either one of the pairs (A – B or B – C). Correspondingly, Simonson (1989) found that by adding a third extreme product C (expensive, high picture quality TV) to the binary choice set A and B (A: low priced, low picture quality TV; B: medium priced, medium picture quality TV) – whereby C was inferior in one attribute (e.g., higher price), but superior in the other attribute (e.g., better picture quality) to A and B – participants did choose the middle product (here B) more often (in 48% instead of 23% of the cases). This pattern could even be observed when participants in the binary choice set were undecided between A and B or even preferred A. According to the "betweenness inequality" property of the standard economic theory (e.g., EUT) the popularity of the middle option should have dropped and that drop should have been larger than that for the extreme option A. The *compromise* effect is attributed to the concept of *loss aversion* (as pointed out by Simonson and Tversky, 1992). However, instead of relating losses / gains to a *neutral reference point*, the products available (i.e., A, B and C) build the reference points. That means the extreme products A or C have, compared to the other extreme product, both a

large disadvantage as well as a large advantage. However, the middle product B has only – compared to either of the two extreme products A or C – small disadvantages / advantages. Since disadvantages loom larger than corresponding advantages, consumers should prefer the middle product (i.e., B) – an assumption confirmed by various studies (Kivetz, Netzer, & Srinivasan, 2004). Importantly, the compromise effect presupposes that the consumer considers both attributes (e.g., quality and price) as *equally* important. If the consumer, however, considers only one attribute in his / her decision, disadvantages and advantages become large for that attribute (e.g., quality), but negligible for the other one (e.g., price). This leads to an asymmetric pattern of extremeness aversion, the *polarization effect*. That is, an increase in the popularity for the middle product B should only be observed by the introduction of one of the two extreme products (A or C) but not both (A and C).

To summarize: So far, the default literature regarding electricity products focused on binary choice sets. For this reason we are interested in the question of whether adding a third, more environmentally-friendly, electricity product impedes default effects. This question is relevant since some countries (e.g., Switzerland) have electricity markets where consumers can choose between 3 (or more) electricity products. It is also relevant for potential search engines that compare electricity products and list, for example, only the three highest rated electricity products. Furthermore, we want to know whether adding a third option changes the popularity of the middle option, the green electricity product: that is, does it lose, stay constant or gain popularity? To answer those questions, we compare the results from Pichert and Katsikopoulos's[1] study with ours. We designed a similar task, with the difference that our

[1] We compared our results to Pichert and Katsikopoulos's (2008) study, since Ebeling and Lotz's (2015) study doesn't have a neutral control condition (details of that study are mentioned in the introduction section).

participants could choose between three different electricity products by adding an even more environmentally-friendly eco electricity product.

Method

Participants

241 participants (M_{age} = 28.2; SD_{age} = 7.2; 53.9% female) from the ZHAW Zurich University of Applied Sciences and the greater area of Zurich took part in this web-based experiment. Participants from the ZHAW School of Applied Psychology (10%) received course credit for their participation.

Stimulus Material and Procedure

The cover story presented at the beginning of the experiment informed the participants that they had just moved to a different location and their new energy company offers them 3 electricity products: grey, green and eco (products were presented simultaneously). Each of the products included a description of the kind of electricity source used (not renewable combined with renewable electricity sources [grey], renewable electricity sources [green], or renewable electricity sources from specially certificated power plants [eco]), as well as the relative price (lower price [grey], middle price [green], higher price [eco])[2]. Electricity source and price were chosen as attributes since they were considered as the most important when it comes to consumers' choice behavior of electricity products (see, for example, Kaenzig, Heinzle, & Wüstenhagen, 2013).

We had four conditions (3 default conditions and 1 no default condition; summarized under the factor "default manipulation"; see Figure 1). The three default conditions differed

[2] No specific price was mentioned.

from each other in whether the grey, the green or the eco electricity product was made the "default" product. In the default conditions participants had to make 2 decisions: whether they wanted to switch from the default product to either of the other two electricity products or not (it was randomized which question was asked first).

Figure 1

In the "no default" condition (condition 4), participants could freely choose their preferred electricity product. After participants had made their choice(s) they were asked about: 1.) the price (with a blank field) they would be willing to pay for the assigned default condition, respectively their preferred choice in the no default condition (the so-called *reference price*) and 2.) the costs/gains (*switching costs/gains*) they expected in case they switched to a higher / lower product. These questions allow us the assessment of participants' mindsets according to reference and switching costs/gains over the different conditions. Before concluding with demographic data collection, participants had to rate on a scale (ranging from 0 to 100; 0 = only environment / 100 = only price) how much their decisions have been driven by the attributes environmentally-friendliness or price.

Results

Two participants who needed more than 60 minutes[3] to complete the experiment and one participant with sloppy answer behaviour were excluded from the analysis.

[3] Average time (in minutes) to complete the experiment: $M = 12.37$, $SD = 7.49$.

Default effects

Overall, in both default manipulation (see bold printed diagonal [default] and horizontal [no default] numbers in Table 1) the green electricity seemed to be the most popular and the grey electricity the least popular product. However, more interestingly, we found (analysed with three generalized linear models for Poisson distributed count data and Cramer's V for effect sizes) for all three electricity products *default effects* – as defined by Brown and Krishna (2004): percentage of choosing the default option relative to its control, the no default option (for a graphical representation of our results see the left side of Figure 2).

Table 1

In other words, the grey electricity product was chosen more often (talking about a difference of 33%) in the grey default option than in the corresponding no default option (b = 1.65, SE = 0.34, p < .001, V = 0.18). Analogously, the green and the eco electricity product were chosen more often in the default conditions than the corresponding product in the no default conditions (green: b = 0.91, SE = 0.31, p = .004, V = 0.11; eco: b = 1.36, SE = 0.31, p < .001, V = 0.17).

Hence, similar to Pichert and Katsikopoulos (2008) we observed in the grey condition a "negative"[4] default effect. However, in contrast to their study, we observed a "positive" default effect in the other two conditions, that is, participants did buy more often a green

[4] "negative" in respect to the percentage of products chosen in the no default condition.

respectively an eco product, when they were set as the default than when there was no default. Hence, we found different results.

Figure 2

Extremeness aversion

To define how adding a third option changes the choice pattern, that is, whether a compromise or polarization effect can be observed, we compared the results from Pichert and Katsikopoulos's (2008) study which used a binary choice set with ours (threefold choice set). Note, that – from an environmental perspective – it would be a disadvantage if by adding a third extreme option the middle option, the green electricity, would be chosen less often (at least compared to the grey electricity option). We found that in both default manipulations (default and no default) the popularity of the middle option (green electricity; proportion test) increased (though $p > .05$). Our calculations thereby followed those of Simonson and Tversky (1992). They tested for an increase in popularity, for example, option *b* (equals our green electricity), by comparing the percentage of participants who did choose that option [$P_2(b;a)$] in the binary choice set $\{a,b\}$, with those that did choose that option [$P_3(b;a)$] in the extended choice set $\{a,b,c\}$. The percentage of participants who did choose option *b* in the extended choice set was thereby calculated as follows: $P_3(b;a) = P_3(b;a,c)/[P_3(b;a,c) + P_3(a;b,c)]$; the percentages of *a* and *b* of the extended choice thereby represent the new 100%.

By applying those calculations, we found in the default condition, that the popularity of the middle option (green electricity) increased from 54% of the participants that did choose

the green electricity product in Pichert and Katsikopoulos's (2008) default condition to 62% (.43/(.43+.26), z = 1.14, p = .255; see Table 2 for an overview of the results). Thus, we found an effect in the magnitude of 8%; furthermore, the observed pattern seems to be consistent with the polarization effect[5]. A similar pattern of results could be observed in the no default condition where there was an increase from 67% to 79% (.56/(.56+.15), i.e., 12% increase; z = 1.67, p = .096) for the green electricity option. In sum, adding the eco electricity product as a third option seems to be beneficial. This is all the more the case when considering the participants who have opted for the even more environmentally-friendly eco-product.

Table 2

Reference price and switching costs/gains

Answers from participants who specified a monthly reference price higher than CHF 1000.- (4 of 233 answers) and / or switching costs/gains higher than CHF 500.- (2 of 208 answers) were removed from the analysis.

Do participants – in the different conditions - have different reference prices in mind? We ran a 2 (default manipulation: default vs. no default) x 3 (electricity product: grey, green vs. eco) analysis of variance. Though visual inspection of the mean prices in Table 3 would suggest otherwise, there was neither a significant effect for the default manipulation, $F(1, 223) = 1.17$, $p = .28$, $\eta_p^2 = .01$, the linear trend of the electricity product, $F(1, 223) = 3.56$, $p = .06$, $\eta_p^2 = .02$, nor the default manipulation x linear trend of the electricity product interaction, $F(1,$

[5] Results from the next section and from the rating scale (explicit statements; see page 13) underline this assumption.

223) = 0.56, p = .45, η_p^2 < .01. The absence of a significant linear trend of the electricity products is probably due to the high standard deviations of the stated reference prices (overall *SD* = 102.70).

Table 3

Furthermore, how much money were participants expecting to pay maximum – depending on the condition they have been assigned to – for switching to a more expensive and environmentally-friendlier product, and conversely, what was the minimum discount in CHF they expect to get for switching to a cheaper and less environmentally-friendly product? Switching to an environmentally-friendlier product should – if that dimension primarily matters – be associated with a gain ("I get a qualitatively better product"). Hence, the amount of money people are willing to pay for switching to a higher product (e.g., switching from the default grey to green or eco electricity) should – according to Kahneman and Tversky's (1979) concept of *loss aversion* (i.e., losses loom larger than gains) – be smaller than the amount of money they want to get back when switching to a lower product (e.g., from the default eco to green or grey electricity). However, if the dimension price primarily matters an opposite pattern of results should be observed; respectively, similar switching costs/gains should be observed if both dimensions matter more or less equally.

Overall, all switching costs/gains (with one exception: participants who were logged in into the green default, but switched to the grey product: t [7] = 1.55; p = .164) deviated from 0 (all p < .04, all d > 0.76). However, in the default condition we did not find a difference, t(93)

= 0.202; p = .84, between the switching costs/gains on the loss and gain side of Kahneman and Tversky's (1979) reference function. On the other hand, in the no default condition we did find a difference, $t(56) = 4.63$; $p < .001$, g[6] = 1.01, between the switching costs/gains. These results are in agreement with the polarization effect where participants mainly consider the dimension "environmental-friendliness" in their decisions.

Discussion

In our study, we addressed the question of whether using defaults regarding electricity products might be an effective tool to nudge consumers to make more environmentally-friendly decisions (i.e., buying green electricity products) and thereby meeting the goals set by policymakers (for example defined in the "Energy Strategy 2050" in Switzerland, BFE, 2016). However, contrary to previous studies, we used a choice set with 3 electricity products instead of only 2 products. We thereby found, similar to previous research (i.e., Pichert & Katsikopoulos, 2008), a negative default effect in the grey electricity condition. That is, participants did more often choose the grey electricity product in the default than in the no default condition. However, contrary to previous studies, we also observed – in the green and eco electricity conditions – positive default effects. Hence, setting a default on environmentally-friendly products did indeed result in an increase of participants choosing that product. Furthermore, we were interested in how adding the 3rd product changes the popularity of the middle green product. An increased popularity of the middle option, the green product, was found in the no default condition. The resulting pattern thereby suggests, that it can be ascribed to the polarization, not the compromise effect. This assumption is, on the one hand, supported by our result that participants *stated* switching costs/gains for

[6] corrected Hedges g for unequal sample sizes.

switching to a higher electricity product differed from the ones for switching to a lower product in the no default condition; indicating that participants weighted one dimension (environmental-friendliness) higher than the other (price). On the other hand, participants' explicit statement of which dimension they considered in their choice, by choosing the appropriate number on the rating scale, also supports the polarization effect; 41% of the participants did choose a value on the scale close to 0, signalling the importance of the environmentally-friendly dimension. Finally, our results regarding the reference prices suggest that although on first sight the prices, especially in the default conditions, seem to differ between the electricity products, they do not differ significantly; this is probably due to the large standard deviations. Noteworthy is the fact that the stated average reference prices and the mean switching costs/gains are quite high, indicating that our participants' awareness of their monthly electricity consumption cost and costs/gains for switching to another electricity product are not quite realistic (a finding also observed by Tabi, Hille, & Wüstenhagen, 2014). Interestingly, despite this "wrong" awareness of the reference price, participants did choose the pricier green product. Nonetheless, the overestimated prices might have hindered them to choose the eco electricity product. Further research should thereby focus on ways to tell consumers that switching to a more environmentally-friendly product is not as expensive as they might assume; and that the prices are in the range of costs they are willing to pay for switching to a renewable electricity product.

A study limitation could be that our product prices were not specified such as costs / kWh. Hence, further experiments need to be conducted in order to test how price-sensitive the overall pattern of results found is with real-world incentives. Although, there are reports from real-world examples where people were staying with and paying for a green electricity default (see for example Pichert and Katiskopoulos, 2008).

However, nudges are not the only interventions promoting, for example, environmentally-friendly behaviour. Lately, boosts have become popular. Whereas nudges are theoretically grounded in the *heuristics and biases* (H&B) approach, boosts are grounded in the *simple heuristics* (SH) approach. Although, both approaches assume *bounded rationality* (i.e., assuming that people use heuristics; Simon, 1955), in the H&B approach bounded rationality is considered as the error-prone side to the so-called perfect rational side (i.e., obeying the norm; see Kahneman, Slovic, & Tversky, 1982). On the other hand, the SH approach considers it as a good or even better way to make a choice (Gigerenzer & Gaissmaier, 2011; Gigerenzer & Todd, 1999). Hence, depending on the approach, the human cognitive system is considered as deficient or as efficient (see Jungermann, 1983). Specifically, the Nudge program uses the possibility to re-represent information whereas the boost program tries to "foster existing decision-making competencies and to develop new ones" (p. 146, see Hertwig, 2017) in steering people to a "better" behaviour. So, instead of using a default, one could use boosting to increase knowledge about green electricity products and their importance. Interestingly, some energy companies did put effort to inform people about green electricity, but with little success (see for example Chassot, Wüstenhagen, Fahr, & Graf, 2017). It seems like most people show a lack of motivation and involvement to acquire competency about electricity products and thereby using a nudge seems to be the more efficient intervention (see rule 1 of Hertwig, 2017, on how one can decide between nudging and boosting).

To sum up our findings: Defaults and the addition of an eco electricity product seems to have a positive impact on peoples' decision to buy green electricity products. However, people seem unaware of the real prices for green electricity. We thereby postulate that defaults can't do all the work and need to be complemented, for example, by understandable price

information (e.g. the monthly additional costs for green electricity equals the cost of a cup of coffee). We conclude that placing defaults smartly could be a valuable way for policy makers and energy companies to do more for the environment.

Acknowledgements

We thank Barbara Bleiker and Mirjam Schneeberger for their helpful comments.

Literature

Ariely, D. (2009). Predictably irrational. *The hidden forces that shape our decisions*. London: Harper.

Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2009). The importance of default options for retirement saving outcomes: Evidence from the United States. In J. Brown, J. Liebman, D. A. Wise (Eds.), *Social security policy in a changing environment* (pp. 167-195). Chicago: University Press.

BFE (2016). Energiestrategie 2050 nach der Schlussabstimmung. Retrieved from http://www.bfe.admin.ch/php/modules/publikationen/stream.php?extlang=de&name=de_809960773.pdf

Brown, C. L., & Krishna, A. (2004). The skeptical shopper: A metacognitive account for the effects of default options on choice. *Journal of Consumer Research*, *31*(3), 529-539. doi: 10.1086/425087

Chassot, S., Wüstenhagen, R., Fahr, N., & Graf, P. (2017). Introducing green electricity as the default option. In Herbes, C., & Friege C. (Eds.), *Marketing Renewable Energy* (pp. 109-122). Cham: Springer.

Ebeling, F., & Lotz, S. (2015). Domestic uptake of green energy promoted by opt-out tariffs. *Nature Climate Change, 5*(9), 868-871. doi: 10.1038/nclimate2681

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451-482. doi: 10.1146/annurev-psych-120709-145346

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford: University Press.

Hertwig, R. (2017). When to consider boosting: some rules for policy-makers. *Behavioural Public Policy*, *1*, 143-161. doi: 10.1017/bpp.2016.14

Johnson, E. J., Shu, S. B., Dellaert, B. G., Fox, C., Goldstein, D. G., Häubl, G., & Wansink, B. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, *23*(2), 487-504. doi: 10.1007/s11002-012-9186-1

Jungermann, H. (1983). The two camps on rationality. *Advances in Psychology,* 16, 63-86. doi: 10.1016/S0166-4115(08)62194-9

Kaenzig, J., Heinzle, S. L., & Wüstenhagen, R. (2013). Whatever the customer wants, the customer gets? Exploring the gap between consumer preferences and default electricity products in Germany. *Energy Policy*, *53*, 311–322. doi: 10.1016/j.enpol.2012.10.061

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263-291. doi: 10.2307/1914185

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: University Press.

Kivetz, R., Netzer, O., & Srinivasan, V. (2004). Alternative models for capturing the compromise effect. *Journal of Marketing Research*, *41*(3), 237-257. doi: 10.1509/jmkr.41.3.237.35990

Metzler, M. (2015, September 2015). *Konsumenten sind manchmal etwas naiv*. Neue Zürcher Zeitung am Sonntag. Retrieved from https://nzzas.nzz.ch/

Pichert, D., & Katsikopoulos, K. V. (2008). Green defaults: Information presentation and pro-environmental behaviour. *Journal of Environmental Psychology*, *28*(1), 63-73. doi: 10.1016/j.jenvp.2007.09.004

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*, 99-118. doi: 10.2307/1884852

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*(2), 158-174. doi: 10.1086/209205

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, *29*(3), 281-295. doi: 10.1177/002224379202900301

Tabi, A., Hille, S. L., & Wüstenhagen, R. (2014). What makes people seal the green power deal? Customer segmentation based on choice experiment in Germany. *Ecological Economics*, *107*, 206-215. doi: 10.1016/j.ecolecon.2014.09.004

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth and happiness*. London: Penguin Books.

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: University Press.

Table 1

Association between default/no default and choice. N (%)

	Choice			
Default	Grey	Green	Eco	Total
Grey	**41 (48%)**	30 (35%)	15 (17%)	86 (100%)
Green	8 (9%)	**67 (76%)**	13 (15%)	88 (100%)
Eco	8 (10%)	23 (29%)	**49 (61%)**	80 (100%)
No default	**17 (15%)**	**64 (56%)**	**33 (29%)**	114 (100%)
Total	74 (20%)	184 (50%)	110 (30%)	368 (100%)

Table 2

Effect of the choice set size

	2 vs. (3) options; in %	
Electricity product	Default	No default
Grey electricity	46 (26[7])	33 (15)
Green electricity	**54 (43)**	**67 (56)**
Eco electricity	- (31)	- (29)

Note. In the "Default" column the amount of "stays" per default condition have been summarized and percentages were calculated. The first percentage values were taken from the Pichert and Katsikopoulos study, the one in the brackets is from our study.

[7] The percentage value (26) was calculated as follows: 41/(41+67+49), using the numbers coming from Table 1 (the bold values presented in the diagonal).

Table 3

Reference prices (in bold) and switching costs/gains

	Electricity		
Condition	Grey	Green	Eco
Default			
Grey	**75.62 (*N* = 39)**	+39.67 (*N* = 30)	+46.64 (*N* = 14)
Green	-37.00 (*N* = 8)	**100.77 (*N* = 43)**	+31.92 (*N* = 12)
Eco	-47.50 (*N* = 8)	-34.57 (*N* = 23)	**126.37 (*N* = 38)**
No default			
Grey	**79.29 (*N* = 17)**	+12.44 (*N* = 9)	+16.75 (*N* = 8)
Green	-84.17 (*N* = 18)	**75.22 (*N* = 60)**	+25.04 (*N* = 48)
Eco	-75.00 (*N* = 9)	-62.89 (*N* = 19)	**101.09 (*N* = 32)**

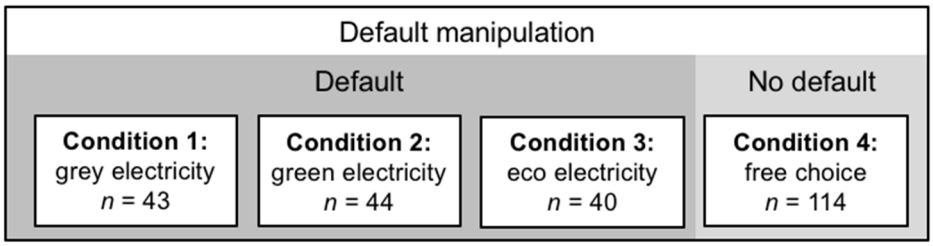


Figure 1. Participants were first randomly assigned to either the default or no default condition and then within the default condition to one of its 3 sub-conditions.

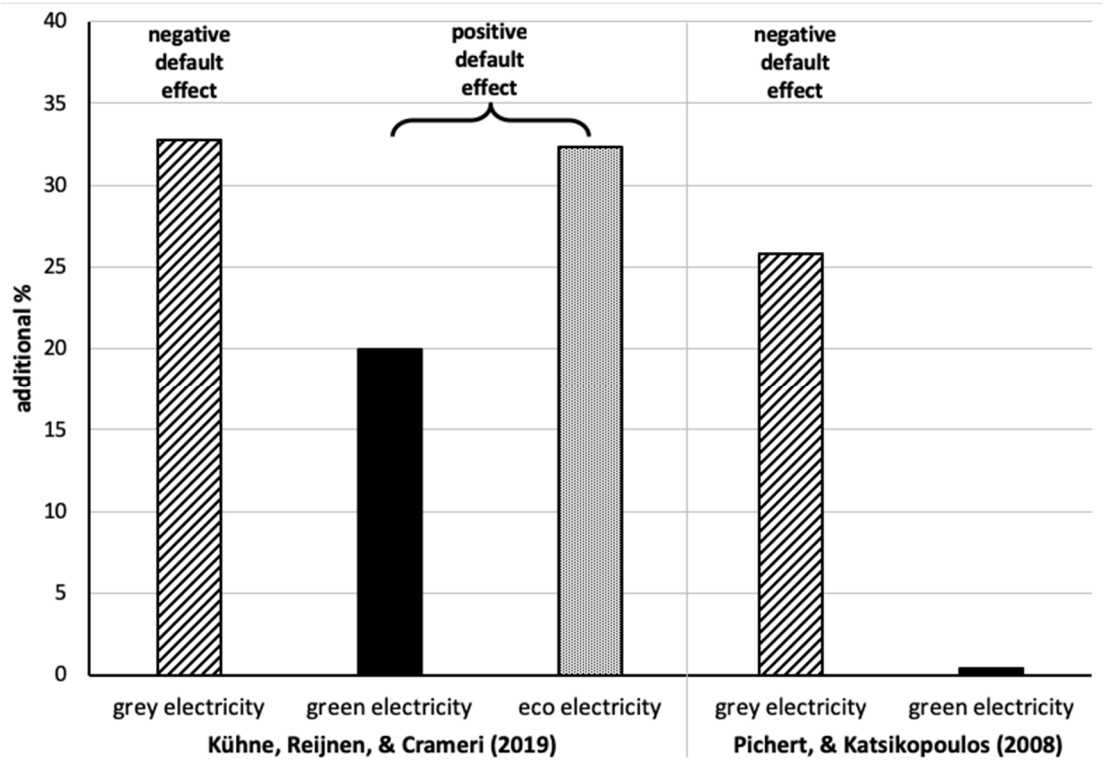


Figure 2. The default effect; bars indicate the additional percentage of choices made in favour of the default option relative to its control, the no default condition.