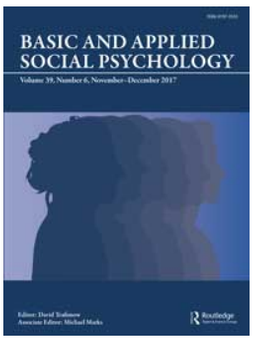

Basic and Applied Social Psychology

ISSN: 0197-3533 (Print) 1532-4834 (Online) Journal homepage: https://www.tandfonline.com/loi/hbas20

Nudge: Concept, Effectiveness, and Ethics

Yiling Lin, Magda Osman & Richard Ashcroft

To cite this article: Yiling Lin, Magda Osman & Richard Ashcroft (2017) Nudge: Concept, Effectiveness, and Ethics, Basic and Applied Social Psychology, 39:6, 293-306, DOI: 10.1080/01973533.2017.1356304

To link to this article: https://doi.org/10.1080/01973533.2017.1356304

View supplementary material

Published online: 09 Nov 2017.

Submit your article to this journal

Article views: 2897

View related articles

View Crossmark data

Citing articles: 4 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=hbas20

BASIC AND APPLIED SOCIAL PSYCHOLOGY
2017, VOL. 39, NO. 6, 293–306
https://doi.org/10.1080/01973533.2017.1356304

Routledge
Taylor & Francis Group

Check for updates

Nudge: Concept, Effectiveness, and Ethics

Yiling Lin, Magda Osman, and Richard Ashcroft

Queen Mary University of London

ABSTRACT

Nudges are psychologically informed tools designed to promote behavioral change in order to improve health and well-being. In this review, we focus on three areas of concern: theory, evidence base, ethics. We begin by discussing the problems arising from the theoretical framework that nudges are based on and propose an alternative framework that helps to classify nudges into two types (Type 1 and Type 2). We then evaluate the evidence for nudges in the health domain, drawing attention to critical empirical issues (internal and external reliability) that explain the limited evidence base for their effectiveness. The review ends with an examination of the implications of the theoretical and empirical issues we discussed with respect to current debates regarding the ethics of nudge.

Noncommunicable diseases (NCDs)—principally, heart diseases, stroke, cancer, diabetes, and chronic lung diseases—are responsible for almost 70% of global deaths (World Health Organization [WHO], 2017a). However, most NCDs can be reduced by targeting four main risk factors: tobacco use, physical inactivity, harmful use of alcohol, and an unhealthy diet. These factors are speculated, by some, to have a common cause, which is poor health choices resulting from our psychology (Thaler & Sunstein, 2008). That is, we make choices in our day-to-day lives based on heuristics (such as anchoring, availability, representativeness), and biases (optimism, overconfidence, status quo) that drive poor lifestyle choices. Thus, based on this speculation, a potential way of targeting NCDs is to identify the psychological factors that contribute to poor health choices and use behavioral interventions to exploit social scientific research on human behavior: "nudges."

A nudge is "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives" (Thaler & Sunstein, 2008, p. 6). Nudges influence choice behavior in a variety of ways that include (a) provision of information (e.g., leaflets about the benefits of climbing stairs), (b) correcting misapprehensions about social norms (e.g., informing individuals of peer group behavior such as statistics of average alcohol intake), (c) altering the profiles of different choices (e.g., making healthy food appear more prominent in the canteen), (d) implementing default options (e.g., changing an organ donation legislative system from opt-in to opt-out; Bonell, McKee, Fletcher, Wilkinson, & Haines, 2011). The guiding principle behind these examples is to make the "better" option more convenient or salient for the decision maker to select; this option is better because it maximizes future health, wealth, and well-being.

The views on the use of nudges in the health domain range from those praising their potential benefits (Mills, 2013; Saghai, 2013; Sunstein, 2015) to those that raise doubts as to how they are used (Goodwin, 2012; Mols, Haslam, Jetten, & Steffens, 2015; Osman, 2016; Selinger & Whyte, 2012) and whether they are effective (Bonell et al., 2011; Marteau, Ogilvie, Ronald, Suhrcke, & Kelly, 2011; Rayner & Lang, 2011). If these doubts are warranted, then one area that needs inspection is the theoretical framework on which nudges are built on, given that this forms the rationale for how they are supposed to operate (Baldwin, 2014). We use this as a point of departure in our review by first examining the proposed psychological mechanisms that underpin nudges and the problems associated with the theoretical framework adopted. We use this opportunity to propose an alternative theoretical account of nudges as a way to rethink the evidence base of nudge interventions in the health domain. We end the review by discussing how the issues we raise also have implications for ethical debates (Saghai, 2013) and their impact on applied social policy issues (Hansen & Jespersen, 2013).

CONTACT Magda Osman m.osman@qmul.ac.uk School of Biological and Chemical Science, Queen Mary University of London, Mile End Road, London E1 4NS, United Kingdom.

Supplemental data for this article can be accessed on the publisher's website.

© 2017 Taylor & Francis

Dual-system theoretical underpinnings of nudges

The current theoretical framework used to support the foundations of nudge is based on a dual system framework (Kahneman & Frederick, 2002; Stanovich, 1999). The idea is that our judgment, decision-making, and reasoning processes are underpinned by two distinct cognitive systems: System 1 and System 2. The received view on dual-system theories (DST) is that generally there are some family resemblances in the way the two systems operate. In general, System 1 processes are heuristic-based, intuitive, biased, associative, automatic, and System 2 processes are rule based, analytical, flexible, and slow (Evans & Stanovich, 2013; Kahneman, 2011). Nudge theorists (Sunstein, 2015, 2016a; Thaler & Sunstein, 2008) claim that the basis on which we make poor lifestyle choices is commonly through the activation of System 1 type processes. Because of this, a practical way of generating positive behavioral change is to target System 1 process by reorienting features of the choice contexts on which heuristics and biases are invoked; this is typically achieved covertly without the decision maker's awareness. This in turn relies on a distinction between implicit and explicit processes whereby implicit processing occurs without awareness, whereas explicit processing is deliberate and is accompanied by awareness (Evans & Over, 1996; Stanovich & West, 2000). We focus on two core issues in the evaluation of System 1/System 2 distinctions: (a) the nature of the interaction between the two systems and (b) the lack of precision around the details of key distinctions between the two systems (Osman, 2016). Our evaluation is designed to better understand how to target behavior via nudges, because having a better idea of the actual underlying mechanism that guides decision making should reveal where and where not nudges are likely to be effective (Grüne-Yanoff, 2016; Grüne-Yanoff & Hertwig, 2015).

Critical issues with DST

Issue 1

The nature of the relationship between system 1 and 2. DST do not all make the same claims regarding the relationship between systems (Evans & Stanovich, 2013). Some dual-system theorists claim that the systems are interactive (Sloman, 1996), some claim that they operate in parallel (Epstein, 1973, 1994; Evans & Over, 1996), and some claim that they operate serially (Gilbert, 1989); while others propose that System 1 is the default system, and only later does System 2 kick in to monitor the outputs of System 1 (Evans & Stanovich, 2013; Kahneman, 2003)—known as the default-interventionist approach. If, as nudge proponents claim, Systems 1 and 2 are dissociated and they do not interact, then it makes better sense to try to directly tap into System 1 processes to generate behavioral change; this is the default system that gives rise to many of the core decision-making and reasoning processes that drive suboptimal lifestyle choices. Though, by doing so, nudge theorists and practitioners need to identify which common suboptimal behaviors are exclusively driven by System 1 processes, which thus far the program of nudges has yet to do. However, if it is the case that the two systems do interact, then does it still make sense to introduce an intervention solely on System 1? If not, then nudges may not operate in the way that they are intended. In other words, if the grounds on which they are based is theoretically problematic, this may also explain why they are not as effective as they are purported to be.

To explore Issue 1 further we consider the bat and ball task (Kahneman & Frederick, 2002) that Thaler and Sunstein (2008) used to illustrate the relationship between Systems 1 and 2. The task involves presenting participants with this simple description and a question: "A bat and ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?" Typically the majority of participants will intuitively answer 10 cents (Kahneman & Frederick, 2002) when the correct answer is 5 cents. Just based on this task alone, DST vary significantly in their explanations for this common error. One interpretation is that System 1 is invoked automatically and is the sole driver of the error (Kahneman & Frederick, 2002, 2005), which is the favored interpretation of Thaler and Sunstein (2008). An alternative explanation is that System 2 is also in operation but fails to detect the error generated by System 1 (Kahneman & Tversky, 1982), and a third interpretation is that because System 2 is slower than System 1, System 2 detects the error but cannot intervene quickly enough to prevent the error being made (Gilbert, 1989; Stanovich & West, 2000). The role that System 2 plays in this task in turn has implications for how researchers develop methods to ameliorate the error effectively. Thus, the same problem extends to nudges, in which the reasons behind a poor choice being made, in turn will impact what appropriate intervention is needed to reduce the chances of it continually being made.

Issue 2

Lack of precision regarding the critical distinctions between the two systems. There are essentially three different proposals regarding the core qualitative difference between System 1 and System 2. One view is

that they can be differentiated based on their demands on working memory (De Neys, 2006; Evans & Stanovich, 2013; Oppenheimer, 2008); working memory is a system for the temporary holding and manipulation of information during the performance of a range of cognitive tasks. A second claim is that they differ depending on the extent to which metacognitive processes are invoked (Thompson, 2009); metacognition broadly refers to explicit knowledge or beliefs of factors that affect the outcome of a cognitive operation. The third is that they vary to the extent to which representations are accessible (Kahneman, 2003); the ease (or effort) with which particular mental contents come explicitly to mind. Essentially, the greater the dependency on working memory, or metacognitive processes, or difficulty in accessing representations suggests System 2 is in operation, and the opposite applies for System 1.

There are three main problems with the apparent qualitative distinctions between Systems 1 and 2. The first is, and as has been highlighted by DST, that in actual fact the three qualities essentially reduce to one single factor, namely, dependency on working memory (Evans & Stanovich, 2013). This in and of itself is not necessarily a concern, as it suggests strong compatibility between the theoretical claims. What is a concern is that dependency on working memory is not all or nothing, and modeling and empirical demonstrations of the way in which high-order cognition relies on working memory is based on quantitative differences, not qualitative ones (e.g., Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007). Second, the purported core qualitative distinction between Systems 1 and 2—be that working memory, metacognition, or accessibility of representations—is used to explain why one system is implicit/automatic and the other is explicit. As with working memory, measures identifying automatic and explicit processes often rely on speed of response, which is a continuous measure, meaning that there are a variety of judgments, choices, and inferences that are made, some of which are faster than others (Osman, 2004, 2007, 2013, 2014; Osman & Stavy, 2006). The relative nature by which automatic and explicit responses are identified means that it makes better sense to claim that some responses are faster than others, rather than some are automatic and others are explicit. Thus far no dual-system theorist has presented the necessary and sufficient conditions by which to identify automatic responses independently of explicit responses (Osman, 2004). Moreover, not all behaviors associated with System 1 are fast, and not all behaviors associated with System 2 are slow, and to accommodate this, some theorists have proposed a four-system framework (Klaczynski, 2001; Sun, Slusarz, & Terry, 2005). Third, the same initial observations that led to the formation of DST, and were used to theorize about the key qualitative property that determines the difference between the two systems, are redescribed to form predictions of the same observations that identify differences between the two systems. Put simply, it would be akin to detecting that sometimes people make erroneous choices quickly, and correct choices slowly, and then theorize that this is because of two underlying systems— one of which is fast, and the other is slow—each of which differs based on ease of access of information. From this, a prediction is formed that outlines that when the slow system is being used, people will make correct choices, but not when the fast system is invoked. This is referred to as a particular type of circular argument (self-dependent justification), which has been commonly found in the area of DST of decision making, judgment, and reasoning (Hahn, 2011). Given the serious concerns discussed around Issue 1 and 2, we propose an alternative.

A single system account of nudge

Many dual-system theorists (Evans & Over, 1996; Evans & Stanovich, 2013; Kahneman, 2003), as well as critics of them (Osman, 2004, 2013, 2014), agree that a fully dissociationist view of System 1 and System 2 is not adequate for capturing the complexities in which decision-making processes operate. In light of this, and other serious concerns with dual-system frameworks, there are several single-system frameworks (e.g., Kruglanski, 2013; Osman, 2004, 2007, 2013; Osman & Stavy, 2006; Simon, Snow, & Read, 2004). Building on these unitary system frameworks, we extend their proposals by suggesting that they reduce to a parallel constraint satisfaction (PCS) model (Simon et al., 2004). This is essentially a connectionist approach in which the spread of activation among nodes in the network is fully sufficient for the processing of an outcome (e.g., a choice) and the basis on which a decision is made (knowledge, evidence, beliefs) is coded in the network through the pattern of weights among the nodes (Read, Vanman, & Miller, 1997). PCS processing is guided by the goal of maximizing consistency, which means the need to reconfigure/reevaluate/update knowledge/evidence/beliefs from multiple (potentially conflicting) sources that bring about an outcome (i.e., judgment, choice, inference) so that both (representations and outcome) are in alignment (i.e., coherent). Where DST identify distinct types of processes that can be classified as System 1 or System 2, in a PCS model, variations in processing of information is

predicted and modeled according to the degree of restructuring that needs to occur for coherence (between knowledge and behavior) to be achieved (Simon & Holyoak, 2002).

How does a single system account apply to nudges?

We extend the PCS single-system framework to help classify nudges into two types that differ according to the degree to which processing efforts are needed to maintain psychological coherence (see Table 1 for examples). Type 1 nudges target decision-making contexts that generate responses that are not typically accompanied by critical inspection to prompt reconsideration of the choices made. For instance, familiar consumer-based contexts such as supermarkets involve highly practiced patterns of behavior leading to repetitive choices being made. So, in response, Type 1 nudges involve simple interventions such as rearranging the presentations of consumer items in food isles to highlight options that would have ordinarily been ignored. Type 1 nudges minimally disrupt the choice context to prompt some adjustment in the way information within it is processed at the point of decision, but not enough that the decision maker detects any dissonance between the nudged choice and their general value-system. Type 2 nudges aim to promote a sustained reevaluation of the evidence base on which people make their choices, and the choices themselves, by disrupting the coherence between the two. For example, long-term educational campaigns promoting exercise present the benefits of regular exercise, as well as the harmful effects of continuing to be sedentary. Repeated exposure to information of this kind is designed to create dissonance resulting from the costs of maintaining poor habits and the benefits of changing them. Given the cognitive system's need for coherence, a reevaluation and restructuring of knowledge representations is needed to bring in alignment the evidence base (knowledge/beliefs) and choice behavior. It is worth reiterating here that the two types of nudges differ according to the amount of reevaluation of information on which people's choices are made and actions taken based on it, it is not predicated on a difference between qualitative differences in systems of thought. Now that we have considered a simple way to conceptualizing nudges into types, we turn to the evidence base in order to critically consider which type of nudge is shown to be effective.

Nudges: Evidence base in the health domain

In this section we examine a range of nudges (Table 1) implemented in four health domains: (a) poor diet, (b) physical inactivity, (c) alcohol overconsumption and (d) tobacco use. This is not a comprehensive review of the evidence in the literature but instead is a focused evaluation of core findings that provide a representative impression of the pattern of evidence in this area (see the appendix for further information).

Poor diet

The size of packaging and the portions of food products has dramatically increased over the past 30 years (Young & Nestle, 2002), which in turn has affected our food consumption. A common Type 1 nudge approach to address overconsumption has been to change visual cues in a food environment which may consist of the availability of certain foods; the variety of food assortments; size of food packages and portions; or shape/size of plates, glasses, and bowls. These cues are often used to imply a consumption norm that helps regulate how much we eat or drink in a food environment (Wansink, 2004). In addition, many people adhere to the norm "plate clean," which means that in a food establishment if the plate is large, and the portions match the size, people consume to the size of the portion on the plate, and not to the point of being sated (Schwarz, 1998). For example, depending on the plate size, up to 45% more food is consumed in a Chinese buffet setting (Wansink & van Ittersum, 2013). A comparison of 58 studies (6,603 participants) in a recent Cochrane review found that people consistently ate more food when offered larger portions, packages, or items of tableware than when offered smaller versions (Hollands et al., 2015). Introducing Type 1 nudges that reduce plate size in food establishments was shown to have a reasonable effect in reducing intake but was dependent on participants being unaware of the

Table 1. A summary of the evidence on the effectiveness of Type 1 and Type 2 Nudges.

Health domain	Type 1 Nudges		Type 2 Nudges	
Unhealthy Diet	smaller plate sizes	*Limited effect*	Calorie labelling	*Ineffective*
			Traffic light labelling	*Effective*
Physical Inactivity	Footprints	*Ineffective*	Motivational posters	*Mixed evidence*
Harmful Use of Alcohol	Adopting straight glassware	*Insufficient evidence*	Correct social norm misapprehension	*Mixed evidence*
Tobacco Use	Shorter cigarettes	*Ineffective*	Health warnings on branded packs	*Mixed evidence*
			Health Warnings on plain packs	*Effective*

manipulation (Holden, Zlatevska, & Dubelaar, 2016). However, it isn't clear that the intervention successfully generalizes to contexts beyond the one in which the nudge was implemented. Moreover, there are many factors that significantly influence our eating habits. Eating is often a social activity, and we take our cues as to how much to consume from our dining partners, and distractions encourage us to overconsume (e.g., watching television, watching movies at the cinema), which suggests that there are multiple countervailing factors that limit the scope of Type 1 nudges (Wansink, 2004).

In this context, as an alternative to Type 1 nudges, Type 2 nudges focus on improving the presentation of information on which people make their food choices, such as providing calorie counts on food menus to draw attention to both healthy and unhealthy options (Downs, Loewenstein, & Wisdom, 2009). However, evidence suggests that this has had limited effects in increasing healthy food choices (Loewenstein et al., 2012). One reason for this is that calorie labels do not provide an obvious reference point as to which specific options are best for people. In and of themselves, they do not reliably motivate people to systematically monitor and translate calorie counts to shift their choices over a sustained period. Indeed, a recent systematic review of the impact of reading calorie labels at the point of purchase or consumption has had little to no effect on positively changing people's choice behavior (Kiszko, Martinez, Abrams, & Elbel, 2014; Sinclair, Cooper, & Mansfield, 2014). The review found that regardless of the length of the intervention, the Type 2 nudge was generally ineffective at changing behaviour over a 4-week period (Elbel, Gyamfi, & Kersh, 2011; Elbel, Kersh, Brescoll, & Dixon, 2009; K. L. Webb, Solomon, Sanders, Akiyama, & Crawford, 2011); 2-month period (Holmes, Serrano, Machin, Duetsch, & Davis, 2013); and 13-month period (410 Finkelstein, Strombotne, Chan, & Krieger, 2011). To deal with these concerns, traffic light systems have been used as a way to make calorie information more salient and intuitively simpler to interpret (Sacks, Rayner, & Swinburn, 2009; Sonnenberg et al., 2013). By using simple visual cues (e.g., red = highly fat food, amber = moderately healthy foods, green = very healthy foods), the signals directly connect calorie counts to their impact on health (House of Lords, 2011) and provide a relevant reference point (Liu, Wisdom, Roberto, Liu, & Ubel, 2014). An independent field study conducted by Ipsos Mori showed that 35% of customers of a major UK supermarket actively look at traffic light labels when they shop, and 92% of those find these labels easy to understand. Also, over the 12-week period, sales of food items with mostly green traffic lights grew to 46.1%, whereas those with mostly red traffic lights decreased by 24% (House of Lords, 2011). Thus, in the United Kingdom alone, several organizations (The National Institute for Care and Health Excellence; The UK Food Standard Agency) have strongly encouraged food manufacturers and food establishments to use text and traffic lights on food labels/menus because they support better understanding of which food are healthy and which are not.

Physical inactivity

As with poor diet, another major global issue is the significant decreases in regular exercise (Hallal et al., 2012). A simple method of increasing physical activity through a Type 1 nudge involves a point-of-decision prompt that uses visual cues in relevant contexts to encourage people to take the more active of two options (e.g., a choice between taking the stairs rather than the escalator). For instance, painted footprints on stairwells have been used to guide people to take the stairs over elevators. The evidence base for this is not encouraging. Findings show that, perversely, the method increased the selection of the less physical option (Åvitsland, Solbraa, & Riiser, 2017). Although similar to the Type 1 nudge, a favored Type 2 nudge, is to also use a point-of-decision prompt but instead of covert visual cues, educational information is presented that highlights the benefits of regular exercise. This often involves placing posters at start of stairwells or by elevators/escalators that inform people about the calories they would burn or the net positive effects on their health (i.e., increased heart rate), (Andersen, Franckowiak, Snyder, Bartlett, & Fontaine, 1998; Blamey, Mutrie, & Aitchison, 1995; Brownell, Stunkard, & Albaum, 1980; A. Lewis & Eves, 2012; Marshall, Bauman, Patch, Wilson, & Chen, 2002; Nomura, Yoshimoto, Akezaki, & Sato, 2009; O. J. Webb & Eves, 2007). A recent review of the evidence reported that across 11 studies the success rate was at chance levels (Soler et al., 2010), but a second review reported an overall positive effect ranging between 0.3% and 10.6% (Nocon, Müller-Riemenschneider, Nitzschke, & Willich, 2010). A speculation in the difference between these two reviews is that there is variability in the length of the intervention. The length of the intervention is often between 4 and 12 weeks, but with some notable exceptions [24 weeks (Kerr, Eves, & Carroll, 2001b; 9 months: Lee et al., 2012)]. But further work has shown that there is no association between effectiveness of the Type 2 nudge and length of intervention. Another potential explanation for the mixed findings is that the locations in which the nudge was implemented varied, and so it is hard to compare like for like, for instance, comparing

returning populations taking the stairs at train station versus those at a shopping mall. To further uncover the precise reasons for why Type 2 nudges vary in their effectiveness in this domain, the types of informational prompts have been evaluated. A specific message such as "7 minutes of stair climbing protects your heart" was shown to be more effective than a general message such as "Stay healthy, use the stairs" (Puig-Ribera & Eves, 2010).

It is also worth noting that studies examining Types 1 and 2 nudges typically involve point-of-decision prompts placed at stairwell with only one to two flights of stairs. In order to experience any significant impact on cardiorespiratory fitness women need to climb at least six flights of stairs daily (Boreham, Wallace, & Nevill, 2000), and men need to climb 25 flights to result in any significant improvement in aerobic power (Fardy & Ilmarinen, 1975). Even if some of the studies have shown positive effects of nudges, they would fall short of any meaningful impact on actual physical fitness levels.

Alcohol overconsumption

The evidence base for nudges designed to reduce alcohol overconsumption accounts for only 7.3% of all behavioral intervention studies in the health domain (Hollands et al., 2013), despite the severity of the problem (Magnusson, 2009). Akin to the Type 1 nudges used to reduce food consumption via altering the size of food containers, a similar rationale has been adopted in the context of alcohol consumption. This typically involves offering alcohol in tall, narrow glasses as opposed to short, wide glasses in drinking establishments (bars and public houses; Wansink & van Ittersum, 2005). This is motivated by work showing that the rate of alcohol consumption is related to the shape of glassware, which is slower in a straight glass compared to a curved glass (Attwood, Scott-Samuel, Stothart, Munafò, & Campanella, 2012). A recent systematic review of studies examining the use of this nudge reported that there was not enough evidence to estimate the effect on reducing consumption (Hollands et al., 2015).

In line with other Type 2 nudges discussed so far, a preferred method is to provide explicit information as a means of generating behavioral change, for example, providing a more accurate idea of the safe quantities to consume through the use of social norm cues. As social creatures, people are sensitive to majority influences, and this can have a strong pervasive influence on behavior (Bullers, Cooper, & Russell, 2001; Ennett et al., 2006; Pearson & West, 2003). That is, a wealth of evidence from social psychology shows that people behave in accordance with their peers (J. V. Wood, 1989), often as part of a group mentality; there is a strong drive to belong and be accepted by a group. Given that consuming alcohol is typically a social activity, the claim is that using social norm cues (i.e., the typical amount of alcohol a particular social group consumes) is a more efficient way of helping people regulate their alcohol consumption by evaluating it relative to the consumption of their peers (Nishida, Akaoka, & Nishizawa, 1975). For instance, heavy drinkers often judge their alcohol consumption to be equal to or even less than their peers, even though it is substantially greater (Perkins, Meilman, Leichliter, Cashin, & Presley, 1999; they feel as if they can reasonably justify their behavior by rationalizing that it is no different than their peers'). To correct misapprehensions of social norms in a student population, several studies using self-reported survey responses have shown that Type 2 social norm interventions (through educational campaigns) implemented over 1 year (Gomberg, Schneider, & DeJong, 2001) and 5 years (Haines & Spear, 1996) have successfully reduced alcohol consumption. However, a different review of 66 studies analyzed alcohol reduction at 4 months postintervention and found that the effect sizes were small and were unlikely to be of meaningful benefit in practice (Foxcraft, Moreira, Santimano, & Smith, 2015). It is worth noting that when surveyed, students doubted the credibility of the educational campaign messages (Thombs, Dotterer, Olds, Sharp, & Raub, 2004). In addition, some have suggested that the effectiveness of the nudge needs to take into account campus sizes in which norm misperception may be harder to correct if "everybody knows everyone else," and thus students are more confident in their estimates of others' drinking levels (Borsari & Carey, 2003). Moreover, it is possible that the average or typical norm used to compare drinking levels in these Type 2 nudges does not represent the ideal normative reference point (M. A. Lewis & Neighbors, 2006). In other words, feedback that involve best friends' drinking rather than typical student drinking level would be more specific and may have a stronger influence, assuming the peers are actually consuming alcohol within healthy limits (Baer, Stacy, & Larimer, 1991; Borsari & Carey, 2003; M. A. Lewis & Neighbors, 2006).

Tobacco use

Another serious problem is tobacco consumption, which kills around 6 million people each year (WHO, 2016). To target this, Type 1 nudges promoting smoking cessation have focused on increasing the availability of shorter cigarettes; however, a systematic review found that when compared to standard-sized cigarettes, there was no

overall reduction in tobacco consumption (Hollands et al., 2015). Alternatively, a more common route is to adopt educational campaigns that are typical of Type 2 nudges. In many Western countries, health warnings on cigarette packages are among the most common means of increasing smokers' awareness of the risks of smoking (Hammond et al., 2006). It is now mandatory that consumers of tobacco products have a "fundamental right to health information, including accurate information about the harms of tobacco use" (WHO, 2015). A comprehensive review by Hammond et al. (2006) found that smokers' knowledge of toxic constituents in tobacco smoke was very low even among smokers in affluent and educated countries in the world. Although presenting health information is a common practice for many countries across the world, the style of presentation of health information differs between countries, making it difficult to evaluate the effectiveness of these messages on reducing consumption.

More recently, Type 2 nudges such as plain cigarette packaging which standardises the shape, color, and method of opening the package, as well as the health warnings themselves. The aim of is to fulfill several objectives that include reducing the attractiveness of consuming tobacco (Hammond, Daniel, & White, 2013; Hammond & Parkinson, 2009; Moodie & Mackintosh, 2013; Moodie, Mackintosh, Hastings, & Ford, 2011) and restricting use of the pack as a form of advertising and promotion while increasing the size of the health warnings (Maynard, Munafò, & Leonards, 2013; Moodie et al., 2012; Munafò, Roberts, Bauld, & Leonards, 2011). A large review of 37 studies concluded that plain packaging was rated as less attractive and contained poorer quality products than branded packaging (WHO, 2017b), although this did not indicate the impact on tobacco consumption.

Empirical work looking at the impact of health warnings on tobacco consumption is still in its infancy, but the findings are promising and suggest that they indeed reduce acute craving and are often associated with more negative perceptions of smoking (Brose, Chong, Aspinall, Michie, & McEwen, 2014). In the short term, plain packaging has been shown to encourage cessation for up to 2 weeks (Moodie & Mackintosh, 2013). In the medium and long term, there is evidence to suggest that plain packaging decreased tobacco consumption 6 months postintervention (Dunlop, Dobbins, Young, Perez, & Currow, 2014) as well as 12 months postintervention (Wakefield et al., 2015). Meanwhile Australia, as the first WHO member to implement standardize packaging, has also seen a statistically significant decline in smoking prevalence as a result of this Type 2 nudge (Australian Government, Department of Health, 2016).

Empirical issues concerning nudge

Having examined the available evidence of both Type 1 and Type 2 nudges, we evaluate the methodological issues concerning the implementation of nudges designed to promote health behaviors. As with any intervention designed to improve behavior, the most reliable way to confidently make casual inferences about a manipulation and its possible effect is to compare it against a control condition (randomized control trials). However, randomized control trials are hard to introduce in field work, and so this, along with other factors, limits the ability to draw firm conclusions as to the effectiveness of nudges. Beyond this, we next raise two key points that we consider need addressing in future empirical research on nudge.

Internal reliability of experiments examining nudges

Internal reliability refers to the extent to which a measure is consistent within itself, namely, it generates the same behavior each time it is used within the same context. The Type 1 nudges just reviewed suggest that overall the evidence base is mixed and that the replicability of positive nudge interventions is hard to establish. This raises questions about the reasons for when Type 1 nudges do work, and why the effects are hard to replicate. Loewenstein, Bryce, Hagmann, and Rajpal (2015) speculated that the limited effectiveness of Type 1 nudges results from peoples lack of deep insight into how the nudge is designed to influence their behavior. Ashcroft (2013) proposed that the effectiveness of both Type 1 and 2 nudges in general may depend on the various heterogeneous motivations/value systems people have with regards to changing their behavior in line with a healthier lifestyle. In addition, the fact that nudges are highly context dependent (Kosters & Van der Heijden, 2015) means that some Type 1 nudges are less likely to work in some contexts rather than others, and a clearer understanding of the context in which they are implemented is needed. Thus, a critical step in devising research programs around nudging in health domains is to establish the internal reliability of nudges over time at an epidemiological level but also at the individual level. However, the research practices so far have yet to adopt methodological techniques that tackle any of these issues in depth (i.e., assessing motivation needs, levels of awareness of nudges, characterizing the contexts in which they are implemented) in order to better establish internal reliability.

External reliability of experiments examining nudge

External reliability refers to the extent to which a measure varies from one use to another. With most

nudge field experiments, the difficulty is in reproducing the same conditions in different contexts under which the original intervention was assessed. For instance, consider nudges designed to increase physical activity. The informational prompts have been used in various environmental settings such as libraries (Russell, Dzewaltowski, & Ryan, 1999), underground stations (Blamey et al., 1995), and office buildings (Coleman & Gonzalez, 2001). In meta-analytic reviews of these studies (Andersen et al., 1998; Blamey et al., 1995; Brownell et al., 1980; Kerr, Eves, & Carroll, 2001a; Kerr et al., 2001b; Marshall et al., 2002; Nomura et al., 2009), none suggested that there was a consistent pattern of evidence across the different contents that were studied. The positive impact on behavior as indicated in these studies increased stair use over elevators/escalators varied from around 2% to 12% but not controlling for length of time in which the measure was implemented, that is, 1 month or 3 months. Similar inconsistencies have also been noted for nudges that extend beyond the health domain. For instance, nudges used to increase civic behaviors, such as recycling, volunteering, voting, petitioning, donating, and debating, have shown that the variation in how long and where they are implemented may explain why overall effect size is as low as 9% (John et al., 2013). Thus, with findings such as these there need to be more efforts in standardizing the ways in which nudges are examined in the wild in order to establish external reliability.

Thus, from an empirical perspective the picture appears to be somewhat bleak with respect to establishing good evidence for the effectiveness of nudges in the health domain. The main problem is that it is hard to draw any firm conclusions as to their effectiveness in the long term (i.e., positive change over 1 year or more), which should be the ultimate goal of assessing their effectiveness. Often because the studies are conducted in the field, as with many field-based studies, the problems are that it is hard to run studies on a large sample with proper controls, and it is rare to find field studies that also carry out follow-ups to examine the effects of the nudges in the long term. This does not undermine the program of nudge per se but simply that the evidence to date does not allow researchers to draw strong conclusions about its general effectiveness in generating meaningful positive behavioral change. Moreover, as noted, the limitations in drawing firm conclusions is restricted not only to their effectiveness and reliability over time but also in establishing the generalizability of positive behavioral change beyond the context in which nudges are implemented. Furthermore, the small effect sizes reported in empirical studies means that translating their positive results at a population level may render them less effective than typical social policy methods (i.e., mandates, bans, taxes). Given that there is a growing list of international governments wanting to apply nudge to public policy on important issues such as health and well-being, there is clearly a need to establish further empirical rigor in order to better establish the effectiveness of these behavior interventions (Osman, 2016).

General theoretical reconsiderations of the nudge evidence base

As discussed earlier in this review, a concern for the nudge program is that the theoretical foundations on which it is built are problematic. We have proposed that the types of nudges that have been developed fall into two broad categories, which differ depending on the extent to which they promote a reevaluation of information that informs better decisions (i.e., maximizing long-term gains), so as to bring the new information and choice behavior into greater alignment (greater coherence). This is in contrast to a position of Thaler and Sunstein (2008) and Sunstein (2014, 2016c) that nudges differ according to the underlying differences between System 1 and System 2 thought operations. Building on our proposals and the evidence we reviewed, one reason why Type 1 nudges seem to be ineffective and tend to be short-lived is because they do not engage the decision maker on any substantial level to reexamine the basis on which the decisions are made so as to meaningfully shift their choice behavior. This is consistent with Loewenstein et al.'s (2015) claim regarding the level of insight that people have as to the underlying basis on which Type 1 nudges are designed to influence their behavior. Indeed, without prompting people to think about and acknowledge that they might be eating/drinking less as a result of smaller dinnerware/glassware, any behavioral change is not likely to become sticky (i.e., habitual), or reliably generalize to other contexts outside of where the nudge is present. It has long since been known that habits require sustained and explicit association between situational cues and learned behavioral responses (Hull, 1943), often through repetition of a behaviors in the same context (W. Wood, Quinn, & Kashy, 2002) for the behaviors to generalize beyond them.

The evidence we have reviewed regarding Type 2 nudges typically involve interventions that provide explicit information that is directly connected to the pursuit of a clearly identified goal, which in turn is associated with a specific choice behavior (e.g., reduce unhealthy eating, alcohol consumption, tobacco consumption); this has been in the form of providing

calorie information, a peer group's alcohol consumption, or health warnings on cigarette packages. More to the point, Type 2 nudges seem to be effective in reducing poor health behaviors such as alcohol overconsumption (Haines, Barker, & Rice, 2003; Haines & Spear, 1996) and cigarette smoking (Hancock, Abhold, Gascoigne, & Altekruse, 2002; Hancock & Henry, 2003; Linkenbach, Perkins, & DeJong, 2003) for a period equal to or greater than 12 months. The evidence shows that through repeated intervention over long periods, some Type 2 nudges (particularly those correcting misapprehensions of social norms) can lead to sustainable behavioral change over longer periods (i.e., over 1 year postintervention). Thus, we suggest here that in order to establish reliable methods that promote critical reexamination of one's values, attitudes, and motivations, we advocate that Type 2 nudges should be more frequently used, and over sustained periods (i.e., at least 6–12-months educational campaigns). The rationale for this is that, unlike Type 1 nudges, Type 2 nudges typically encourage a form of reevaluation of behavior through explicit means, and this helps to maintain greater coherence between the information on which new choice behaviors are made coherently. We speculate that it is for this reason that the evidence base suggests that they are relatively more effective in leading to meaningful sustained behavioral change than Type 1 nudges (see Table 1).

Ethical implications

Our review of the theoretical foundations of nudge, as well as the evidence base examining the efficacy of nudges in the health domain, suggests that there is good reason to focus on implementing Type 2 nudges over Type 1 nudges. That is to say, there are theoretical grounds on which Type 2 nudges can be argued to have a more sustainable and deeper impact on generating positive behavioral change, and in line with this, the current evidence base indicates, to some degree, that they are more effective than Type 1 nudges. Not only are there theoretical and empirical grounds for promoting Type 2 nudges, but here we briefly discuss the ethical reasons that corroborate this conclusion.

By definition, nudges are designed to influence choice behavior, but not at the expense of forbidding any options. That is, they preserve people's rights to freely choose whatever option they like, but that the nudge is designed to highlight the option deemed better for them in the long run. This is why the nudge program is liberal paternalistic (Sunstein, 2016c; Thaler & Sunstein, 2008). This approach is designed to spare policymakers any ethical concerns that they might likely face, because a certain choice behavior is being encouraged by the state that promotes a certain value-based lifestyle approach but that preserves the right an individual has to do otherwise (for an in depth discussion, see Osman, 2016). What has come into question is whether nudges easily allow people to do otherwise. Debates have arisen because some have argued that the way in which nudges operate, specifically Type 1 nudges, does not preserve freedom of choice, because choice behavior is predominately influenced without the awareness of the individual (Ashcroft, 2013; Blumenthal-Barby & Burroughs, 2012; Bovens, 2009; Dworkin, 2012; Osman, 2016; Saghai, 2013). That is, it would be hard to do otherwise, if one is choosing an option without any awareness of how and why it is being chosen. From an empirical standpoint this is a moot point because so far there simply isn't enough evidence to suggest that Type 1 nudges work, and if they do, there is no good evidence to suggest that they are influencing behavior on a nonconscious level (Osman, 2014). Rather, the ethical problems that researchers have raised consider the rationale behind the nudge program itself, and the extent to which an endeavor that has global appeal should be a cause for concern. Nudge defenders (Sunstein, 2016b; Thaler & Sunstein, 2008) appeal to comparisons with typical social policy tools such as mandates, taxes, and bans as a way to explicitly and more heavy-handedly steer people to behave in a manner that maximizes their own and society's good (i.e., seat belt laws, fat tax, smoking bans). Although a critical point to highlight here is that typical policy measures are explicit, and often accompanied by educational campaigns, so the public is well aware of the basis on which their choices are being modified, even if they aren't necessarily happy about it (Osman, 2014; Weber, 2017). In this regard, it is worth also considering the fact that public surveys of nudges also suggest that the public show much higher approval ratings for Type 2 over Type 1 nudges (Arad & Rubinstein, 2015; Felsen, Castelo, & Reiner, 2013; Hagman, Andersson, Västfjäll, & Tinghög, 2015; Hedlin & Sunstein, 2016; Jung & Mellers, 2016; Mazzocchi et al., 2015; Reisch & Sunstein, 2016; Reisch, Sunstein, & Gwozdz, 2016; Sunstein, 2016c; Sunstein, Reisch, & Rauber, 2017). Again, this goes to show that, whether or not people are going to modify their behavior in light of nudges, or more typical policy methods of behavioral change, they are supportive of explicit methods that signal what methods are being used and how they change behavior over those that seek to do this covertly, especially without their consent (Osman, 2016). On a macro level, in a democratic society, nudges, like other governance interventions, would be subject to evaluation by the

government official that represents the citizens' interest (Sunstein, 2016a). If there are indeed strong objections against nudge, then public officials would be and should be attentive to public views of their consent to implement them, as well as acquiring evidence to examine their effectiveness. Similarly, the duty of public officials is to promote the welfare of the citizen in the long term, but this cannot be done in a liberal society without discussion and debate.

In conclusion, the motivation behind this review was fairly simple: to better understand how nudges work. With that in mind, once this is achieved, then the better armed we (social scientists, policymakers, practitioner) are designing ways of intervening on behaviors to achieve the best outcome for individuals that need and want it. In this review, we argue that any meaningful change in behavior arises from developing a consistently coherent basis on which people understand the reasons for their decisions and how they enact them. If, through nudges, we want to encourage people to help themselves, particularly in targeting serious problems around NCDs, we need to make the goal of helping oneself making better lifestyle choices a coherent and sustained approach. We argue that for theoretical, empirical, and ethical reasons, this is best achieved through Type 2 rather than Type 1 nudges.

Funding

This work was supported by the Life Sciences Initiative Studentship (LSIPGRS) [Grant Number LSIPGRS].

ORCID

Yiling Lin http://orcid.org/0000-0003-4955-4146
Magda Osman http://orcid.org/0000-0003-1480-6657
Richard Ashcroft http://orcid.org/0000-0001-6065-4717

References

Andersen, R. E., Franckowiak, S. C., Snyder, J., Bartlett, S. J., & Fontaine, K. R. (1998). Can inexpensive signs encourage the use of stairs? Results from a community intervention. *Annals of Internal Medicine*, *129*, 363–369.

Arad, A., & Rubinstein, A. (2015). *The people's perspective on libertarian-paternalistic policies* (Unpublished Manuscript). Tel Aviv University, Tel Aviv, Israel. Retrieved from: http://www.tau.ac.il/~aradayal/LP.pdf

Ashcroft, R. E. (2013). Doing good by stealth: Comments on "Salvaging the concept of nudge". *Journal of Medical Ethics*, *39*(8), 494. doi:10.1136/medethics-2012-101109

Attwood, A. S., Scott-Samuel, N. E., Stothart, G., Munafò, M. R., & Campanella, S. (2012). Glass shape influences consumption rate for alcoholic beverages. *PLoS ONE*, *7*(8), e43007. doi:10.1371/journal.pone.0043007

Australian Government, Department of Health. (2016). *Post-Implementation Review Tobacco Plain Packaging 2016*. Retrieved from http://ris.pmc.gov.au/sites/default/files/posts/2016/02/Tobacco-Plain-Packaging-PIR.pdf

Åvitsland, A., Solbraa, A. K., & Riiser, A. (2017). Promoting workplace stair climbing: sometimes, not interfering is the best. *Archives of Public Health*, *75*, 2. doi:10.1186/s13690-016-0170-8

Baer, J. S., Stacy, A., & Larimer, M. (1991). Biases in the perception of drinking norms among college students. *Journal of Studies on Alcohol*, *52*(6), 580–586.

Baldwin, R. (2014). From regulation to behaviour change: Giving nudge the third degree. *The Modern Law Review*, *77*(6), 831–857. doi:10.1111/1468-2230.12094

Blamey, A., Mutrie, N., & Aitchison, T. (1995). Health promotion by encouraged use of stairs. *BMJ*, *311*(7000), 289–290. doi:10.1136/bmj.311.7000.289

Blumenthal-Barby, J. S., & Burroughs, H. (2012). Seeking better health care outcomes: The ethics of using the "nudge." *The American Journal of Bioethics*, *12*(2), 1–10. doi:10.1080/15265161.2011.634481

Bonell, C., McKee, M., Fletcher, A., Wilkinson, P., & Haines, A. (2011). One nudge forward, two steps back. *BMJ*, *342*, d401. doi:10.1136/bmj.d401

Boreham, C. A. G., Wallace, W. F. M., & Nevill, A. (2000). Training effects of accumulated daily stair-climbing exercise in previously sedentary young women. *Preventive Medicine*, *30*(4), 277–281. doi:10.1006/pmed.2000.0634

Borsari, B., & Carey, K. B. (2003). Descriptive and injunctive norms in college drinking: A meta-analytic integration. *Journal of Studies on Alcohol*, *64*(3), 331–341.

Bovens, L. (2009). The ethics of nudge. In T. Grüne-Yanoff & S. O. Hansson (Eds.), *Preference change: Approaches from philosophy, economics and psychology* (pp. 207–219). Dordrecht: Springer Netherlands.

Brose, L. S., Chong, C. B., Aspinall, E., Michie, S., & McEwen, A. (2014). Effects of standardised cigarette packaging on craving, motivation to stop and perceptions of cigarettes and packs. *Psychology & Health*, *29*(7), 849–860. doi:10.1080/08870446.2014.896915

Brownell, K. D., Stunkard, A. J., & Albaum, J. M. (1980). Evaluation and modification of exercise patterns in the natural environment. *The American Journal of Psychiatry*, *137*(12), 1540–1545.

Bullers, S., Cooper, M. L., & Russell, M. (2001). Social network drinking and adult alcohol involvement. *Addictive Behaviors*, *26*(2), 181–199. doi:10.1016/S0306-4603(00)00099-X

Coleman, K. J., & Gonzalez, E. C. (2001). Promoting stair use in a US–Mexico border community. *American Journal of Public Health*, *91*(12), 2007–2009. doi:10.2105/AJPH.91.12.2007

De Neys, W. D. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, *17*, 428–433. doi:10.1111/j.1467-9280.2006.01723.x

Downs, J. S., Loewenstein, G., & Wisdom, J. (2009). Strategies for promoting healthier food choices. *American Economic Review*, *99*, 159–164. doi:10.1257/aer.99.2.159

Dunlop, S. M., Dobbins, T., Young, J. M., Perez, D., & Currow, D. C. (2014). Impact of Australia's introduction of tobacco plain packs on adult smokers' pack-related perceptions and responses: results from a continuous tracking survey. *BMJ Open*, *4*(12), e005836. doi:10.1136/bmjopen-2014-005836

Dworkin, G. (2012). Lying and nudging. *Journal of Medical Ethics*, *39*(8), 496–497.

Elbel, B., Gyamfi, J., & Kersh, R. (2011). Child and adolescent fast-food choice and the influence of calorie labeling: a natural experiment. *International Journal of Obesity (London)*, *35*, 493–500. doi:10.1038/ijo.2011.4

Elbel, B., Kersh, R., Brescoll, V. L., & Dixon, L. B. (2009). Calorie labeling and food choices: a first look at the effects on low-income people in New York City. *Health Affairs*, *28*(6), w1110–w1121. doi:10.1377/hlthaff.28.6.w1110

Ennett, S. T., Bauman, K. E., Hussong, A., Faris, R., Foshee, V. A., Cai, L., & DuRant, R. H. (2006). The peer context of adolescent substance use: Findings from social network analysis. *Journal of Research on Adolescence*, *16*, 159–186. doi:10.1111/j.1532-7795.2006.00127.x

Epstein, S. (1973). The self-concept revisited: Or a theory of a theory. *American Psychologist*, *28*, 404–416. doi:10.1037/h0034679

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *The American Psychologist*, *49*, 729–724.

Evans, J. S. B. T., & Over, D. E. (1996). *Rationality and reasoning*. East Sussex, UK: Psychology Press.

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241. doi:10.1177/1745691612460685

Fardy, P. S., & Ilmarinen, J. (1975). Evaluating the effects and feasibility of an at work stairclimbing intervention program for men. *Medicine and Science in Sports*, *7*(2), 91–93.

Felsen, G., Castelo, N., & Reiner, P. B. (2013). Decisional enhancement and autonomy: Public attitudes towards overt and covert nudges. *Judgement and Decision Making*, *8*(3), 202–213. doi:10.1177/2332858416674007

Finkelstein, E. A., Strombotne, K. L., Chan, N. L., & Krieger, J. (2011). Mandatory menu labeling in one fast-food chain in King County, Washington. *American Journal of Preventive Medicine*, *40*(2), 122–127. doi:10.1016/j.amepre.2010.10.019

Foxcraft, D. R., Moreira, M. T., Santimano, N. M. L., & Smith, L. A. (2015). Social norms information for alcohol misuse in university and college students (Review). *Cochrane Database of Systematic Reviews*, 12. doi:10.1002/14651858.CD006748.pub4

Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189–211). New York, NY: Guilford Press.

Gomberg, L., Schneider, S. K., & DeJong, W. (2001). Evaluation of a social norms marketing campaign to reduce high-risk drinking at the University of Mississippi*. *The American Journal of Drug and Alcohol Abuse: Encompassing All Addictive Disorders*, *27*(2), 375–389.

Goodwin, T. (2012). Why we should reject "nudge." *Politics*, *32*(2), 85–92.

Grüne-Yanoff, T. (2016). Why behavioural policy needs mechanistic evidence. *Economics and Philosophy*, *32*(3), 463–483. doi:10.1017/S0266267115000425

Grüne-Yanoff, T., & Hertwig, R. (2015). Nudge versus boost: How coherent are policy and theory? *Minds and Machines*, *26*(1–2), 1–52. doi:10.1007/s11023-015-9367-9

Hagman, W., Andersson, D., Västfjäll, D., & Tinghög, G. (2015). Public views on policies involving nudges. *Review of Philosophy and Psychology*, *6*(3), 439–453. doi:10.1007/s13164-015-0263-2

Hahn, U. (2011). The problem of circularity in evidence, argument, and explanation. *Perspectives on Psychological Science*, *6*(2), 172–182.

Haines, M. P., Barker, G., & Rice, R. (2003). The Northern Illinois University experiment: A longitudinal case study of the social norms approach. In H. W. Perkins (Ed.), *The social norms approach to preventing school and college age substance abuse: A handbook for educators, counselors, and clinicians* (pp. 21–34). San Francisco, CA: Jossey-Bass.

Haines, M. P., & Spear, S. F. (1996). Changing the perception of the norm: A strategy to decrease binge drinking among college students. *Journal of American College Health : J of ACH*, *45*(3), 134–40. doi:10.1080/07448481.1996.9936873

Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W. & Ekelund, U., for the Lancet Physical Activity Series Working Group. (2012). Global physical activity levels: Surveillance progress, pitfalls, and prospects. *The Lancet*, *380*(9838), 247–257.

Hammond, D., Daniel, S., & White, C. M. (2013). The effect of cigarette branding and plain packaging on female youth in the United Kingdom. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, *52*, 151–157. doi:10.1016/j.jadohealth.2012.06.003

Hammond, D., Fong, G. T., McNeill, A., Borland, R., Cummings, K. M., & Hammond, D. (2006). Effectiveness of cigarette warning labels in informing smokers about the risks of smoking: Findings from the International Tobacco Control (ITC) Four Country Survey. *Tobacco Control*, *15*, 19–25. doi:10.1136/tc.2005.012294

Hammond, D., & Parkinson, C. (2009). The impact of cigarette package design on perceptions of risk. *Journal of Public Health*, *31*(3), 345–353. doi:10.1093/pubmed/fdp066

Hancock, L., Abhold, J., Gascoigne, J., & Altekruse, M. (2002, August). Applying social norms marketing to tobacco cessation and prevention: Lessons learned from three campaigns (Working Paper No. 6). *The Report on Social Norms.* Available from http://socialnorms.org/wp-content/uploads/2015/05/WP6.pdf

Hancock, L., & Henry, N. (2003). *Perceptions, norms and tobacco use in college residence hall freshmen: Evaluation of a social norms marketing intervention*. San Francisco, CA: Jossey-Bass.

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*, *1*, 3–28.

Hedlin, S., & Sunstein, C. R. (2016). Does active choosing promote green energy use: Experimental evidence. *Ecology*, *43*, 107.

Holden, S. S., Zlatevska, N., & Dubelaar, C. (2016). Whether smaller plates reduce consumption depends on who's serving and who's looking: A meta-analysis. *Journal of the Association for Consumer Research*, *1*(1), 134–146.

Hollands, G. J., Shemilt, I., Marteau, T. M., Jebb, S. A., Kelly, M. P., Nakamura, R., … Ogilvie, D. (2013). Altering micro-environments to change population health behaviour: Towards an evidence base for choice architecture

interventions. *BMC Public Health*, *13*(1), 1218. doi:10.1186/1471-2458-13-1218

Hollands, G. J., Shemilt, I., Marteau, T. M., Jebb, S. A., Lewis, H. B., Wei, Y., … Ogilvie, D. (2015). Portion, package or tableware size for changing selection and consumption of food, alcohol and tobacco. *Cochrane Database of Systematic Reviews*, (9), CD011045. doi:10.1002/14651858.CD011045.pub2

Holmes, A. S., Serrano, E. L., Machin, J. E., Duetsch, T., & Davis, G. C. (2013). Effect of different children's menu labeling designs on family purchases. *Appetite*, *62*, 198–202.

House of Lords. (2011). *Behaviour change*. London, UK: Author. Retrieved from http://www.parliament.uk/business/lords

Hull, C. L. (1943). *Principles of behaviour: An introduction to behaviour theory*. New York, NY: Appleton-Century Crofts.

John, P., Moseley, A., Wales, C., Smith, G., Richardson, L., Stoker, G., & Cotterill, S. (2013). *Nudge, nudge, think, think: Experimenting with ways to change civic behaviour*. London, UK: Bloomsbury Academic.

Jung, J. Y., & Mellers, B. A. M. (2016). American attitudes toward nudges. *Judgement and Decision Making*, *11*(1), 62–74.

Kahneman, D. (2003). Maps of bounded rationality: A perspective on intuitive judgement and choice. In T. Frangsmyr (Ed.), *Lex prix Nobel: The Nobel prizes 2002* (pp. 449–489). Stockholm, Sweden: Almquist & Wiksell International.

Kahneman, D. (2011). *Thinking fast and slow*. London, UK: Allen Lane.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York, NY: Cambridge University Press.

Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. J. Holyoack & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267–293). New York, NY: Cambridge University Press.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 493–508). New York, NY: Cambridge University Press. doi:10.1016/0010-0277(82)90022-1

Kerr, J., Eves, F. F., & Carroll, D. (2001a). Getting more people on the stairs: the impact of a new message format. *Journal of Health Psychology*, *6*(5), 495–500. doi:10.1177/135910530100600503

Kerr, J., Eves, F. F., & Carroll, D. (2001b). Six-month observational study of prompted stair climbing. *Preventive Medicine*, *33*(5), 422–427. doi:10.1006/pmed.2001.0908

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. doi:10.1007/s10900-014-9876-0

Klaczynski, P. A. (2001). Analytic and heuristic processing influences on adolescent reasoning and decision making. *Child Development*, *72*(3), 844–861.

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

Kruglanski, A. W. (2013). Only one? The default interventionist perspective as a unimodel—Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science*, *8*, 242–247. doi:10.1177/1745691613483477

Lee, K. K., Perry, A. S., Wolf, S. A., Agarwal, R., Rosenblum, R., Fischer, S., … Silver, L. D. (2012). Promoting routine stair use evaluating the impact of a stair prompt across buildings. *American Journal of Preventive Medicine*, *42*, 136–141. doi:10.1016/j.amepre.2011.10.005

Lewis, A., & Eves, F. (2012). Prompt before the choice is made: Effects of a stair-climbing intervention in university buildings. *British Journal of Health Psychology*, *17*, 631–643. doi:10.1111/j.2044-8287.2011.02060.x

Lewis, M. A., & Neighbors, C. (2006). Social norms approaches using descriptive drinking norms education: A review of the research on personalized normative feedback. *Journal of American College Health*, *54*, 213–218. doi:10.3200/JACH.54.4.213-218

Linkenbach, J., Perkins, H. W., & DeJong, W. (2003). *Parent's perceptions of parenting norms: using the social norms approach to reinforce effective parenting*. San Francisco, CA: Jossey-Bass.

Liu, P. J., Wisdom, J., Roberto, C. A., Liu, L. J., & Ubel, P. A. (2014). Using behavioral economics to design more effective food policies to address obesity. *Applied Economic Perspectives and Policy*, *36*, 6–24. doi:10.1093/aepp/ppt027

Loewenstein, G., Asch, D. A., Friedman, J. Y., Melichar, L. A., Volpp, K. G., Lowenstein, G., … Volpp, K. G. (2012). Can behavioural economics make us healthier? *BMJ*, *344*, e3482. doi:10.1136/bmj.e3482

Loewenstein, G., Bryce, C., Hagmann, D., & Rajpal, S. (2015). Warning: You are about to be nudged. *Behavioral Science & Policy*, *1*, 35–42.

Magnusson, R. S. (2009). Rethinking global health challenges: Towards a "global compact" for reducing the burden of chronic disease. *Public Health*, *123*, 265–274.

Marshall, A. L., Bauman, A. E., Patch, C., Wilson, J., & Chen, J. (2002). Can motivational signs prompt increases in incidental physical activity in an Australian health-care facility? *Health Education Research*, *17*, 743–749. doi:10.1093/her/17.6.743

Marteau, T. M., Ogilvie, D., Ronald, M., Suhrcke, M., & Kelly, M. P. (2011). Judging nudging: Can nudge improve population health? *BMJ*, *342*, d228.

Maynard, O. M., Munafò, M. R., & Leonards, U. (2013). Visual attention to health warnings on plain tobacco packaging in adolescent smokers and non-smokers. *Addiction*, *108*, 413–419. doi:10.1111/j.1360-0443.2012.04028.x

Mazzocchi, M., Cagnone, S., Bech-Larsen, T., Niedźwiedzka, B., Saba, A., Shankar, B., … Traill, W. B. (2015). What is the public appetite for healthy eating policies? Evidence from a cross-European survey. *Health Economics, Policy and Law*, *10*, 267–292. doi:10.1017/S1744133114000346

Mills, C. (2013). Why nudges matter: A reply to Goodwin. *Politics*, *33*(1), 28–36. doi:10.1111/j.1467-9256.2012.01450.x

Mols, F., Haslam, S. A., Jetten, J., & Steffens, N. K. (2015). Why a nudge is not enough: A social identity critique of governance by stealth. *European Journal of Political Research*, *54*(1), 195. doi:10.1111/1475-6765.12073

Moodie, C., & Mackintosh, A. M. (2013). Young adult women smokers' response to using plain cigarette packaging: a

naturalistic approach. *BMJ Open*, *3*(3), e002402. doi:10.1136/bmjopen-2012-002402

Moodie, C., Mackintosh, A. M., Hastings, G., & Ford, A. (2011). Young adult smokers' perceptions of plain packaging: A pilot naturalistic study. *Tobacco Control*, *20*(5), 367–373. doi:10.1136/tc.2011.042911

Moodie, C., Stead, M., Bauld, L., McNeill, A., Angus, K., Hinds, K., … O'Mara-Eves, A. (2012). *Plain tobacco packaging: A systematic review*. London. Retrieved from http://dspace.stir.ac.uk/bitstream/1893/24420/1/Moodie_et_al_2012_PHRC_PlainPacks.pdf

Munafò, M. R., Roberts, N., Bauld, L., & Leonards, U. (2011). Plain packaging increases visual attention to health warnings on cigarette packs in non-smokers and weekly smokers but not daily smokers. *Addiction*, *106*, 1505–1510. doi:10.1111/j.1360-0443.2011.03430.x

Nishida, Y., Akaoka, I., & Nishizawa, T. (1975). Effect of sex hormones on uric acid metabolism in rats. *Experientia*, *31*, 1134–1135. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1204720

Nocon, M., Müller-Riemenschneider, F., Nitzschke, K., & Willich, S. N. (2010). Review Article: Increasing physical activity with point-of-choice prompts–a systematic review. *Scandinavian Journal of Public Health*, *38*, 633–638. doi:10.1177/1403494810375865

Nomura, T., Yoshimoto, Y., Akezaki, Y., & Sato, A. (2009). Changing behavioral patterns to promote physical activity with motivational signs. *Environmental Health and Preventive Medicine*, *14*, 20–25. doi:10.1007/s12199-008-0053-x

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, *12*(6), 237–241. doi:10.1016/j.tics.2008.02.014

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, *11*, 988–1010. doi:10.3758/BF03196730

Osman, M. (2007). Can tutoring improve performance on a reasoning task under deadline conditions? *Memory & Cognition*, *35*, 342–351.

Osman, M. (2013). A case study: Dual-process theories of higher cognition-commentary on Evans & Stanovich (2013). *Prospect Psychological Science*, *8*, 248–252. doi:10.1177/1745691613483475

Osman, M. (2014). *Future-minded: The psychology of agency and control*. New York, NY: Palgrave Macmillan.

Osman, M. (2016). Nudge: How far have we come? *Œconomia. History, Methodology, Philosophy*, *6–4*, 557–570. Retrieved from https://oeconomia.revues.org/2490

Osman, M., & Stavy, R. (2006). Development of intuitive rules: Evaluating the application of the dual-system framework to understanding children's intuitive reasoning. *Psychonomic Bulletin & Review*, *13*, 935–953. doi:10.3758/bf03213907

Pearson, M., & West, P. (2003). Drifting smoke rings: Social network analysis and Markov processes in a longitudinal study of friendship groups and risk-taking. *Connections*, *25*(2), 59–76.

Perkins, H. W., Meilman, P. W., Leichliter, J. S., Cashin, J. R., & Presley, C. A. (1999). Misperceptions of the norms for the frequency of alcohol and other drug use on college campuses. *Journal of American College Health*, *47*(6), 253–258. doi:10.1080/07448489909595656

Puig-Ribera, A., & Eves, F. F. (2010). Promoting stair climbing in Barcelona: Similarities and differences with interventions in English-speaking populations. *European Journal of Public Health*, *20*(1), 100–102. doi:10.1093/eurpub/ckp059

Rayner, G., & Lang, T. (2011). Is nudge an effective public health strategy to tackle obesity? No. *BMJ*, *342*, d2177–d2177. doi:10.1136/bmj.d2177

Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and gestalt principles: (Re)introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review*, *1*, 26–53.

Reisch, L. A., & Sunstein, C. R. (2016). Do Europeans like nudges? *Judgement and Decision Making*, *11*(4), 310–325.

Reisch, L. A., Sunstein, C. R., & Gwozdz, W. (2016). Better than a whip? European attitudes toward health nudges. *Food Policy*, *69*, 1–10. doi:10.1016/j.foodpol.2017.01.007

Russell, W. D., Dzewaltowski, D. A., & Ryan, G. J. (1999). the effectiveness of a point-of-decision prompt in deterring sedentary behavior. *American Journal of Health Promotion*, *13*(5), 257–259. doi:10.4278/0890-1171-13.5.257

Sacks, G., Rayner, M., & Swinburn, B. (2009). Impact of front-of-pack "traffic-light" nutrition labelling on consumer food purchases in the UK. *Health Promotion International*, *24*, 344–352. doi:10.1093/heapro/dap032

Saghai, Y. (2013). Salvaging the concept of nudge. *Journal of Medical Ethics*, *39*(8), 487–493. doi:10.1136/medethics-2012-100727

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H. M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*, 414.

Schwarz, N. (1998). Warmer and more social: Recent developments in cognitive social psychology. *Annual Review of Sociology*, *24*, 239–264. doi:10.1146/annurev.soc.24.1.239

Selinger, E., & Whyte, K. P. (2012). Nudging cannot solve complex policy problems. *European Journal of Risk Regulation*, *3*(1), 26–31.

Simon, D., & Holyoak, K. J. (2002). Structural dynamics of cognition: From consistency theories to constraint satisfaction. *Personality and Social Psychology Review*, *6*, 283–294.

Simon, D., Snow, C. J., & Read, S. J. (2004). The redux of cognitive consistency theories: evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology*, *86*, 814–837.

Sinclair, S. E., Cooper, M., & Mansfield, E. D. (2014). the influence of menu labeling on calories selected or consumed: A systematic review and meta-analysis. *Journal of the Academy of Nutrition and Dietetics*, *114*, 1375–1388. doi:10.1016/j.jand.2014.05.014

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22. doi:10.1037/0033-2909.119.1.3

Soler, R. E., Leeks, K. D., Buchanan, L. R., Brownson, R. C., Heath, G. W., Hopkins, D. H. & the Task Force on Community Preventive Services. (2010). Point-of-decision prompts to increase stair use a systematic review update. *American Journal of Preventive Medicine*, *38*(2, Suppl), S292–S300. doi:10.1016/j.amepre.2009.10.028

Sonnenberg, L., Gelsomin, E., Levy, D. E., Riis, J., Barraclough, S., & Thorndike, A. N. (2013). A traffic light food labeling intervention increases consumer awareness of health and healthy choices at the point-of-purchase. *Preventive Medicine*, *57*, 253–257. doi:10.1016/j.ypmed.2013.07.001

Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Mahwah, NJ: Erlbaum.

Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and Brain Sciences*, *23*(5), 701–717. doi:10.1017/S0140525x00623439

Sun, R., Slusarz, P., & Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: a dual-process approach. *Psychological Review*, *112*(1), 159.

Sunstein, C. R. (2014). *Why nudge? The politics of libertarian paternalism*. New Haven, CT: Yale University Press.

Sunstein, C. R. (2015). Nudges, agency, and abstraction: A reply to critics. *Review of Philosophy and Psychology*, *6*(3), 511–529. doi:10.1007/s13164-015-0266-z

Sunstein, C. R. (2016a). Do people like nudges? *Administrative Law Review, Forthcoming*. doi:10.2139/ssrn.2604084

Sunstein, C. R. (Ed.). (2016b). *The ethics of influence: Government in the age of behavioral science*. Cambridge, UK: Cambridge University Press.

Sunstein, C. R. (2016c). People prefer System 2 nudges (kind of). *Duke Law Journal*, *66*, 121–168.

Sunstein, C. R., Reisch, L. A., & Rauber, J. (2017). A worldwide consensus on nudging? Not quite, but almost. *Regulation and Governance*. Advance online publication. doi:10.2139/ssrn.2955693

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

Thombs, D. L., Dotterer, S., Olds, R. S., Sharp, K. E., & Raub, C. G. (2004). A close look at why one social norms campaign did not reduce student drinking. *Journal of American College Health*, *53*(2), 61–68. doi:10.3200/JACH.53.2.61-70

Thompson, V. A. (2009). Dual process theories: A metacognitive perspective. In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 171–196). Oxford, UK: Oxford University Press.

Wakefield, M., Coomber, K., Zacher, M., Durkin, S., Brennan, E., & Scollo, M. (2015). Australian adult smokers' responses to plain packaging with larger graphic health warnings 1 year after implementation: results from a national cross-sectional tracking survey. *Tobacco Control*, *24*(Suppl. 2), ii17–ii25. doi:10.1136/tobaccocontrol-2014-052050

Wansink, B. (2004). Environmental factors that increase the food intake and consumption volume of unknowing consumers. *Annual Review of Nutrition*, *24*, 455–479. doi:10.1146/annurev.nutr.24.012003.132140

Wansink, B., & van Ittersum, K. (2005). Shape of glass and amount of alcohol poured: comparative study of effect of practice and concentration. *BMJ*, *331*, 1512–1514. doi:10.1136/bmj.331.7531.1512

Wansink, B., & van Ittersum, K. (2013). Portion size me: Plate-size induced consumption norms and win–win solutions for reducing food intake and waste. *Journal of Experimental Psychology: Applied*, *19*(4), 320–332. doi:10.1037/a0035053

Webb, K. L., Solomon, L. S., Sanders, J., Akiyama, C., & Crawford, P. B. (2011). Menu labeling responsive to consumer concerns and shows promise for changing patron purchases. *Journal of Hunger & Environmental Nutrition*, *6*(2), 166–178. doi:10.1080/19320248.2011.576210

Webb, O. J., & Eves, F. F. (2007). Promoting stair climbing: Intervention effects generalize to a subsequent stair ascent. *American Journal of Health Promotion*, *22*(2), 114–119. doi:10.4278/0890-1171-22.2.114

Weber, E. U. (2017). Breaking cognitive barriers to a sustainable future. *Nature Human Behaviour*, *1*, 13. doi:10.1038/s41562-016-0013

World Health Organization. (2015). *Health warnings*. Retrieved from http://www.euro.who.int/en/health-topics/disease-prevention/tobacco/policy/who-framework-convention-on-tobacco-control-who-fctc/key-areas-of-tobacco-control-policy/health-warnings.

World Health Organization. (2016). *WHO | Tobacco*. Retrieved from http://www.who.int/mediacentre/factsheets/fs339/en/

World Health Organization. (2017a). *NCDs | Noncommunicable diseases and their risk factors*. Retrieved from http://www.who.int/ncds/introduction/en/

World Health Organization. (2017b). *Plain packaging of tobacco products: evidence, design and implementation*. Retrieved from http://apps.who.int/iris/bitstream/10665/207478/1/9789241565226_eng.pdf?ua=1

Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, *106*(2), 231–248.

Wood, W., Quinn, J. M., & Kashy, D. A. (2002). Habits in everyday life: Thought, emotion, and action. *Journal of Personality and Social Psychology*, *83*(6), 1281–1297. doi:10.1037//0022-3514.83.6.1281

Young, L. R., & Nestle, M. (2002). The contribution of expanding portion sizes to the US obesity epidemic. *American Journal of Public Health*, *92*, 246–249.