CrossMark

Nudge Versus Boost: How Coherent are Policy and Theory?

Till Grüne-Yanoff[1] · Ralph Hertwig[2]

Abstract If citizens' behavior threatens to harm others or seems not to be in their own interest (e.g., risking severe head injuries by riding a motorcycle without a helmet), it is not uncommon for governments to attempt to change that behavior. Governmental policy makers can apply established tools from the governmental toolbox to this end (e.g., laws, regulations, incentives, and disincentives). Alternatively, they can employ new tools that capitalize on the wealth of knowledge about human behavior and behavior change that has been accumulated in the behavioral sciences (e.g., psychology and economics). Two contrasting approaches to behavior change are *nudge* policies and *boost* policies. These policies rest on fundamentally different research programs on bounded rationality, namely, the *heuristics and biases* program and the *simple heuristics* program, respectively. This article examines the policy–theory coherence of each approach. To this end, it identifies the necessary assumptions underlying each policy and analyzes to what extent these assumptions are implied by the theoretical commitments of the respective research program. Two key results of this analysis are that the two policy approaches rest on diverging assumptions and that both suffer from disconnects with the respective theoretical program, but to different degrees: Nudging appears to be more adversely affected than boosting does. The article concludes with a discussion of the limits of the chosen evaluative dimension, policy–theory coherence, and reviews some other benchmarks on which policy programs can be assessed.

✉ Till Grüne-Yanoff
gryne@kth.se;
http://home.abe.kth.se/∼gryne/

[1] Department of Philosophy and History of Technology, Royal Institute of Technology (KTH), Teknikringen 78 B, 100 44 Stockholm, Sweden

[2] Max Planck Institute for Human Development, Berlin, Germany

Introduction

It is not difficult to find evidence that people seem prone to making decisions that are
detrimental to their own welfare or to public welfare. Many people fail to save enough
for retirement (e.g., Börsch-Supan 2004). Consumers frequently make unhealthy food
choices, as a result of which overweight and obesity are among the leading risk factors
for deaths worldwide (World Health Organization 2013). Many Europeans endorse
protecting the environment as a worthy personal goal, but only few adopt green
behavior—for instance, by switching from individual to public transportation (European
Commission 2011). And despite the high seismic risk of the region, 90 % of Californians
have no earthquake insurance (Kunreuther and Michel-Kerjan 2011). According to
some behavioral scientists, these and related behaviors raise "serious questions about
the rationality of many judgments and decisions that people make" (Thaler and Sunstein
2008, p. 7). Experimental evidence seems to buttress such dire conclusions: "hundreds
of studies confirm that human forecasts are flawed and biased" and that people's
"decision making is not so great either" (p. 7). Indeed, proponents of the *heuristics and
biases* (H&B) research program (e.g., Kahneman et al. 1982; Gilovich et al. 2002) have
catalogued a long list of what are widely considered systematic cognitive biases and
flawed (e.g., temporally inconsistent) motivations which, they argue, lead to poor
choices (see Thaler and Sunstein 2008, Chapters 1–3).

Proponents of the H&B program, whose groundbreaking work reaches back to
the early 1970s, have more recently interpreted their quest for cognitive biases (and
the heuristics prone to producing them) in terms of the goal of drafting a map of
human *bounded rationality*, a notion proposed by Simon (1956, 1978). In
Kahneman's (2003) words, the goal of his research with Tversky and other
proponents of the H&B program was:

> [T]o obtain a map of bounded rationality, by exploring the systematic biases
> that separate the beliefs that people have and the choices they make from the
> optimal beliefs and choices assumed in rational-agent models. (p. 1449)

Indeed, Simon (1978) himself cited experimental demonstrations of "striking
departures" from the behavior implied by perfect rationality as evidence that classic
theories of rationality (e.g., subjective expected utility theory) do not "provide a good
prediction—not even a good approximation—of actual behavior" (p. 362). He did not,
however, side with the H&B program's view of human choice as systematically flawed.
He suggested that we "understand today many of the mechanisms of human *rational*
choice" and how the "information processing system called Man, faced with complexity
beyond his ken, uses his information processing capacities … to find ways of action that
are sufficient unto the day, that *satisfice*" (p. 368; emphasis added; see also Simon 1956).

The H&B program interprets bounded rationality in terms of a compilation of
systematic biases. But this is only one of several possible interpretations. Another one
took its inspiration from Simon's (1956) emphasis on environmental structures and the

interplay of cognition and environment. The simple heuristics (SH) research program (e.g., Gigerenzer et al. 1999, 2011) has aimed to explore the cognitive mechanisms that a boundedly rational decision maker—one operating under conditions of limited computational capacity, limited information, and uncertainty—employs to make satisficing, that is, *good enough* decisions. Of course, SH does not deny that people sometimes make poor decisions. Unlike the H&B program, however, it does not attribute these behaviors to profoundly flawed mental software. Instead, it presents a vision of bounded rationality according to which human reasoning and decision making can be modeled in terms of SH that—despite ignoring some information and eschewing computationally extensive calculations—produce good (and good enough) inferences, choices, and decisions when applied in the appropriate contexts (see also Hertwig et al. 2013). SH still permits that choices detrimental to individual and collective welfare can arise for various reasons, including the use of heuristics in environments that have changed—as a result of which the cognitive strategy no longer interlocks properly with the environmental structures and affordances (e.g., Wegwarth and Gigerenzer 2013)—or the provision of information that is, by error or design, profoundly confusing (e.g., Gigerenzer et al. 2007).

The SH program and the H&B program share an objective, namely, to build on Simon's notion of bounded rationality. To this end, they seek to uncover how real mortals—rather than idealized agents with boundless time, information, and computational capabilities—make decisions. The similarities end here, however. By treating deviations from normative behavior as deficiencies and cognitive biases, the H&B program accepts classic economic rationality as the normative standard of rational human behavior (which Simon 1978 certainly did not). The SH program, in contrast, builds on the empirical finding that simple strategies can sometimes be as good as or even better than optimizing strategies that need more information and computation, and calls these often coherence-based norms into question (e.g., Arkes et al. 2014)—though, in the rare cases where risks are measurable, it accept some such norms, such as Bayes' theorem (Gigerenzer and Hoffrage 1995). The two programs also disagree about how bounded rationality should be conceptualized and modeled (Katsikopoulos 2014). Admittedly, both programs have focused on heuristics as the key cognitive process of decision making (with the notable exception of cumulative prospect theory; Kahneman and Tversky 1979; Tversky and Kahneman 1992). Yet the heuristics advanced by the H&B program have typically not been developed into process models (Gilovich et al. 2002), do not take the structure of the environment into account (an aspect strongly emphasized by Simon 1956, 1990), and rest on the assumption of a general accuracy–effort trade-off.[1] In contrast, the SH program has proposed computational models of heuristics (i.e., with fully explicated search, stop, and decision processes; see, e.g., Gigerenzer et al. 2011), heuristics whose architecture can match the statistical structure of specific environments (ecologically rational heuristics; Todd et al. 2012), and heuristics offering an existence proof that the accuracy–effort trade-off is not a universal law of human cognition (see Gigerenzer et al. 2011).

---

[1] According to the accuracy–effort trade-off, the less information, computation, or time a decision maker uses, the less accurate his or her judgments will be (see Payne et al. 1993). From this perspective, heuristics are construed to be less effortful, but never more accurate, than more complex strategies.

The two research programs have each resulted in a behavior change program: the *nudge* and *boost* approaches, respectively. Building on the map of systematic biases catalogued in the H&B program, Thaler and Sunstein (2008) argued that the biases' detrimental consequences for people's health, wealth, and happiness both justify and render possible policy interventions that seek to remedy those consequences. Specifically, the goal is to design policies that, by co-opting systematic biases, nudge individual behavior toward a different, more beneficial outcome. The SH program, in contrast, postulates that policies should aim to extend the decision-making competences of laypeople and professionals alike. To this end, interventions can target the individual's skills and knowledge, the available set of decision tools, or the environment in which decisions are made. We refer to this approach as the *boost* approach. This term captures a key difference between the two approaches. The nudge approach assumes "somewhat mindless, passive decision makers" (Thaler and Sunstein 2008, p. 37), who are hostage to a rapid and instinctive "automatic system" (p. 19), and nudging interventions seek to co-opt this knee-jerk system or behaviors such as myopia, loss aversion, and overconfidence to change behavior. The boost approach, in contrast, assumes a decision maker whose competences can be improved by enriching his or her repertoire of skills and decision tools and/or by restructuring the environment such that existing skills and tools can be more effectively applied.[2]

The policy prescriptions of the nudge and boost approaches are clearly rooted in the respective research programs, both of which explore bounded rationality. But how close and coherent is the connection between theory and policy? This is the overarching question addressed by our inquiry. Specifically, our goal is to characterize the policies proposed within both behavior change approaches, to identify their implicit assumptions about cognition, the decision maker, and the policy maker, and to examine how these assumptions map onto the respective research programs (H&B vs. SH).[3] We proceed as follows. Sections 2 and 3 outline the nudging and boosting interventions, illustrating each with a few examples. Section 4 discusses the necessary assumptions of each policy, and Sect. 5 investigates to what extent the theoretical commitments of two programs, H&B versus SH, are consistent with the respective policy assumptions.

Nudging in Practice

There is no precise definition of a nudge, and policies subsumed under the heading differ widely. The likely reason for this diversity is that the founding document of the policy approach, Thaler and Sunstein's (2008) book, *Nudge: improving decisions about health, wealth, and happiness*, offered a sweeping compilation of

---

[2] Some authors have suggested the term "educate" for these kinds of policies (Bond 2009; Katsikopolous 2014). In our view, boosting goes beyond education and the provision of information. For example, in order to boost decision makers' skills, policy designers need to identify information representations that match the cognitive algorithms of the human mind, thus using the environment (e.g., external representations) as an ally to foster insight and decision-making skills. We therefore prefer the term "boost" to "educate".

[3] In the interest of full disclosure, let us point us that the second author has contributed to the SH program (see, e.g., Hertwig et al. 2013).

supposed illustrations of nudges, including such general constructs as "social influence" (p. 55) and "social pressures" (p. 59). Enlisting these social factors as drivers of behavior change is, of course, nothing new (see Cialdini 2001); neither is the provision of information (Thaler and Sunstein, pp. 93–94) or warnings (Sunstein 2014, p. 59).

In what follows, we assume a nudge intervention to be defined by the following properties (see also Rebonato 2012, p. 32):

1. A nudge is intended by the policy maker (choice architect) to steer the chooser's behavior away from the behavior implied by the cognitive shortcoming and toward her ultimate goal or preference (e.g., healthier food choices).
2. A nudge seeks to realize this influence by exploiting empirically documented cognitive shortcomings in human deliberation and choice, without changing the financial incentives (disincentives).
3. A nudge does not affect those features over which people have explicit preferences (e.g., money, convenience, taste, status, etc.), but rather those features that people would typically claim not to care about (e.g., position in a list, default, framing).
4. The behavior change brought about by the nudge should be easily reversible, allowing the chooser to act otherwise.

This definition is not uncontroversial, and Sunstein (2014, p. 59), somewhat ironically, objected to a related definition because he deemed it "imprecise." We nevertheless use this definition here for the following reason: What is genuinely novel about the nudging approach, relative to inventions such as incentive-changing, norm-changing, informing, educating, and de-biasing policies (all subsumed under nudging by Thaler and Sunstein 2008) is the idea of exploiting people's cognitive and motivational deficiencies in ways that help them to make decisions that their better self (or superego) would make. In the following, we provide three examples of nudges that epitomize bias-exploiting interventions: setting defaults for retirement schemes, changing the time horizon of savings programs, and framing incentives for exercise choices.

Setting Defaults

By setting a default, a choice architect determines the option that an individual will receive if she forgoes making a choice of her own. Four steps can be distinguished in the process of setting a default. First, one or more individuals face a choice between two or more options in a choice set. Second, a choice architect determines one element in the choice set to be the default option. Third, the choice architect makes this default known to the individual before the choice. Fourth, the individual receives either the option she has actively chosen or, if she has not made a choice (after a grace period or after her deliberations have ended), the default option.

There is considerable evidence that setting a default can increase the probability of the default option being chosen (for a qualitative review of evidence, see Smith
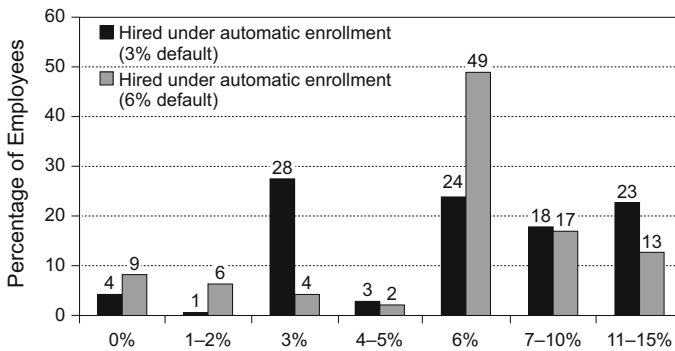
Fig. 1 Distribution of 401 k contribution rates under two defaults (Beshears et al. 2009, p. 173)

et al. 2013). Figure 1, for example, shows how changing the default from a 3 to a 6 % contribution to a pension fund (Beshears et al. 2009) affected participation rates in the respective programs. The first plan automatically enrolled employees who did not make an explicit choice into a 3 % contribution rate; 28 % of employees ended up with this rate. The second plan also automatically enrolled nonchoosers, but at a contribution rate of 6 %; 49 % of employees ended up with this rate. The policy recommendation derived from this and similar investigations is to set the default (e.g., contribution rate) to the option considered optimal, thus ensuring that people end up with that option.

Defaults have been advocated as a "powerful" means (Thaler and Sunstein 2008, p. 35) of steering behavior across a wide range of domains, including organ donation (Johnson and Goldstein 2003), e-mail marketing (Johnson et al. 2002), car purchase decisions (Park et al. 2000), and decisions with environmental impact (Pichert and Katsikopoulos 2008). Setting defaults is, perhaps, the most paradigmatic example of a nudge.

Why do defaults work? Various mechanisms have been proposed (Smith et al. 2013), one of which—consistent with the definition of a nudge provided above—postulates cognitive biases that can be exploited. Specifically, Thaler and Sunstein (2008) wrote:

> [M]any people will take whatever option requires the least effort, or the path of least resistance. … inertia, status quo bias, and the 'yeah, whatever' heuristic. All these forces imply that if, for a given choice, there is a default option … then we can expect a large number of people to end up with that option, whether or not it is good for them. (p. 83)

A related cognitive-bias account stresses the role of loss aversion: People may perceive the default option as something they somehow possess, and giving up this possession as a loss. From this perspective, losing the default option matters more than the (equivalent) gain represented by switching to the nondefault option (Smith et al. 2013).

Changing the Time Horizon of Savings Programs

Another paradigmatic bias-exploiting nudge is the *Save More Tomorrow*[TM] program by Thaler and Benartzi (2004), which likewise aims to increase employees' contributions to retirement savings accounts. In a standard savings decision, employees are asked to choose their optimal trade-off between money available now and money saved for retirement later. In other words, the contribution rate is determined by how much the employee values an extra dollar to spend now over an extra dollar to spend as a retiree, in 20 or 30 years' time.

Some behavioral economists have argued that people typically overvalue the immediate present over the future, but that this overvaluation decreases quickly once a certain delay is factored into the equation. For example, although people tend to prefer one apple now to two apples next week, many people prefer two apples next week to one apple in 3 days. The *Save More Tomorrow*[TM] program exploits this cognitive bias (i.e., time-inconsistent behavior; Loewenstein and Prelec 1992). Instead of asking people to choose a trade-off between consumption now and later, the policy makers ask people to choose between consumption in the near future (say, a year from now) and later. Furthermore, the program takes advantage of people's assumed inertia (increases in savings are automatic) to ease them into self-control restrictions (by projecting those restrictions into the future).[4] Among people who said that they could not afford a cut in pay now, a large majority (79 %; see Thaler and Sunstein 2008) joined the program and agreed to increase their future contribution with every pay rise.

Framing

A final paradigmatic bias-exploiting nudge is framing. Messages about the outcomes of certain actions (on health, financial stability, etc.) are effective not only because of their informational content, but also because of the way they present—or 'frame'—this information. Consider, for instance, the case of a patient with a life-threatening heart condition who needs to decide whether to undergo surgery (see Thaler and Sunstein 2008, pp. 36–37). The patient can be presented with the odds of success in two ways. A 'gain' frame says: "Of 100 patients who have this surgery, 90 are alive after 5 years." A 'loss' frame describes (supposedly equivalent information) thus: "Of 100 patients who have this surgery, 10 are dead after 5 years." These two descriptions are assumed to be "extensionally equivalent descriptions" (Kahneman 2003). Any systematic differences in patient choices caused by such framing, so the interpretation, would violate an essential aspect of rationality, namely, that preferences should not be affected by inconsequential variations in the descriptions of outcomes.

Yet numerous studies have demonstrated that such supposedly inconsequential variations do affect preferences (as manifested, for instance, in terms of 'irrational'

---

[4] To be precise, the original *Save More Tomorrow*[TM] program consisted of two stages. In the first, a consultant discussed possible retirement plans with the employees, based on their own stated preferences. Only if they were reluctant to accept the consultant's advice did the consultant switch to the program (Thaler and Benartzi 2004, p. 172).

preference reversals): For example, participants have been found to evaluate ground beef more favorably when it is described as 75 % lean than when it is described as 25 % fat [Levin and Gaeth 1988; see Rothman et al. (2006) and Gerend and Cullen (2008) on the effectiveness of framed messages in influencing decisions, particularly in the realm of health behavior].

Thaler and Sunstein (2008, p. 37) proposed the following illustration of framing as a policy intervention. A government aiming to encourage energy conversation could inform the public in two different ways:

1. If you use energy conversation methods, you will save $350 per year;
2. If you do not use energy conversation methods, you will lose $350 per year.

Thaler and Sunstein suggested that campaigns that employ the loss frame are much more effective than those using a gain frame. In their view:

> Framing works because people tend to be somewhat mindless, passive decision makers. Their Reflective system does not do the work that would be required to check and see whether reframing the questions would produce a different answer. One reason they don't do this is that they wouldn't know what to make of the contradiction. This implies that frames are powerful nudges, and must be selected with caution (p. 37).

Boosting in Practice

The common denominator behind boost policies is the goal of empowering people by expanding (boosting) their competences and thus helping them to reach their objectives (without making undue assumptions about what those objectives are). These competences can be context-transcending—for instance, statistical literacy— or relatively context-specific, such as making fast and good decisions in a professional (e.g., medical) context. At least three classes of boost policies can be distinguished: Policies that (1) change the environment in which decisions are made, (2) extend the repertoire of decision-making strategies, skills, and knowledge, or (3) do both. In what follows, we describe each of these classes in more detail.

Decisions Under Risk and Risk Competence

Many consequential decisions are based on known statistical information. Numerous studies have concluded that laypeople and professionals alike (see Berwick et al. 1981; Koehler 1996) make poor diagnostic inferences on the basis of statistical information. In particular, their statistical inferences do not follow Bayes' theorem—a finding that prompted Kahneman and Tversky (1972, p. 450) to conclude: "In his evaluation of evidence, man is apparently not a conservative Bayesian: he is not Bayesian at all." The studies from which this and similar conclusions were drawn presented information in the form of probabilities and percentages. From a mathematical viewpoint, it is irrelevant whether statistical

information is presented in probabilities, percentages, absolute frequencies, or some other form, because these different representations can be mapped onto one another in a one-to-one fashion. Seen from a psychological viewpoint, however, as the proponents of the boost approach have argued, representation does matter: some representations make people more competent to reason in a Bayesian way in the absence of any explicit instruction (Gigerenzer and Hoffrage 1995; Hoffrage et al. 2000).

One possible reason for why representation matters is that the cognitive algorithms that perform statistical inference are not adapted to probabilities or percentages: The notion of mathematical probability and mathematical tools such as Bayes theorem were proposed by the classical probabilists of the Enlightenment, and percentages did not become common notations until the 19th century. From this it follows that cognitive algorithms must be tuned to other input formats. Gigerenzer and Hoffrage (1995) hypothesized that algorithms evolved to be adapted to frequencies as actually experienced in series of events. Furthermore, they suggested that Bayesian algorithms are computationally simpler when the input format is natural frequencies rather than probabilities or percentages.[5] Consistent with this hypothesis, Gigerenzer and Hoffrage (1995) and Hoffrage et al. (2000) showed that statistics expressed as natural frequencies improve the statistical reasoning of experts and non-experts alike. For example, as Fig. 2 shows, advanced medical students asked to solve medical diagnostic tasks performed much better when the statistics were presented as natural frequencies than as probabilities. Similar results have been reported for medical doctors (in a range of specialties), HIV counsellors, lawyers, and law students (Hoffrage et al. 2000; Lindsey et al. 2003; Akl et al. 2011; Anderson et al. 2012). These effects commonly occurred without respondents being given training or explicitly instructed to plug probabilities into mathematical formulas, such as Bayes theorem. Sedlmeier and Gigerenzer (2001) trained people to actively construct natural frequency representations and compared the success of this *representation training* to explicit *rule training* (how to insert probabilities into Bayes theorem). Rule training proved to be as good as representation training in terms of transfer to new problems, but representation training had a higher immediate learning effect as well as greater temporal stability.

Changing the representation of statistical information in health brochures, textbooks, and patient–doctor interactions—from probabilities to natural frequencies, from relative to absolute risks (Gigerenzer et al. 2007), or from numerical to graphical representations (Kurz-Milcke et al. 2008; García-Retamero et al. 2010), and generally from representations that mislead and confuse to representations that match the cognitive algorithms of the human mind—can be understood as an act of framing information (see Gigerenzer et al. 2007, p. 58). The key difference to the nudge approach, however, is that statistical information is framed in such a way that people *understand* it and can, by extension, make well-informed decisions (e.g., as patients, jurors), rather than falling prey to confusing and misleading information.

---

[5] Natural frequencies refer to the outcomes of natural sampling—that is, the acquisition of information by updating event frequencies without artificially fixing the marginal frequencies. Unlike probabilities and relative frequencies, natural frequencies are raw observations that have not been normalized with respect to the base rates of the event in question.
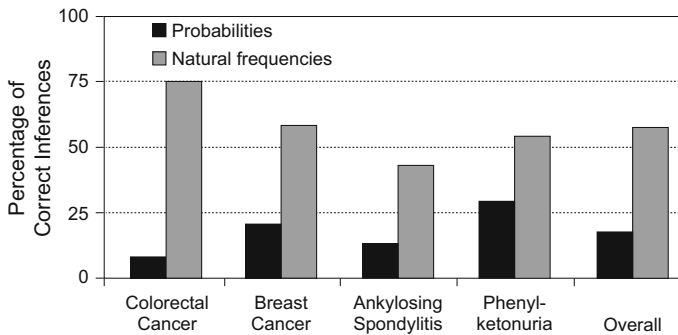
Fig. 2 Boosting medical students' percentage of correct inferences in four realistic diagnostic tasks (adapted from Hoffrage et al. 2000)

For instance, treatment efficacy can be conveyed in terms of relative or absolute risk reduction:

> *Relative risk reduction*: if you have this test every 2 years, it will reduce your chance of dying from this cancer by around one third over the next 10 years
> *Absolute risk reduction*: if you have this test every 2 years, it will reduce your chance of dying from this cancer from around 3 in 1000 to around 2 in 1000 over the next 10 years. (from Sarfati et al. 1998, p. 138)

The benefits of testing are identical in both cases. Yet when the information was presented in the form of relative risk reduction, 80 % of respondents stated they would likely undergo the test, compared with only 53 % when it was presented in the form of absolute risk reduction (Sarfati et al. 1998). A recent review of (experimental) studies showed that many patients do not understand the difference between relative and absolute risk reduction, and that they tend to judge a treatment alternative more favorably if its benefits are expressed in terms of relative risk reduction (Covey 2007).

There are two ways to respond to this finding. The nudge approach might add it to the repertoire of cognitive deficiencies that can be exploited to prompt a specific change in behavior. Relative risk reduction could, for instance, be used to persuade women above age 50 to participate in mammography screening by stating that early detection reduces breast cancer mortality by 20 %. This number operates as a nudge by exploiting individuals' statistical illiteracy, prompting them to accept the number and its behavioral suasion. In absolute numbers, screening reduces mortality from 5 to 4 in 1000 women (after 10 years), tantamount to an absolute risk reduction of 1 in every 1000. But this figure is typically presented as a relative risk reduction, thus causing women to evaluate it more favorably (Gigerenzer 2014).

The boost approach, in contrast, would aim to enhance people's statistical literacy, enabling them to understand and see through confusing and misleading representations by making those representations less manipulative and opaque, rendering them less computationally demanding (Gigerenzer and Hoffrage 1995), making them semantically and pragmatically less ambiguous (Hertwig and

Gigerenzer 1999), and by teaching them to scrutinize potentially manipulative information (e.g., relative risk information). From the boost perspective, difficulties understanding statistical information are seen not as an incorrigible mental deficiency of, say, doctors or patients, but as largely attributable to poor or intentionally misleading information. Moreover, the goal is not to push people toward a particular goal (e.g., to seek or not seek a particular treatment), but to help everybody (e.g., doctors and patients) to understand statistical information as the first critical step toward figuring out one's preference.

The boost approach does not stop at improving the format of statistical information. It also has an explicit educational goal (see Sedlmeier and Gigerenzer 2001). National and international surveys have shown that many people lack the basic numerical skills essential to make informed medical decisions (Nelson et al. 2008; Reyna et al. 2009). Against this background, proponents of the boost approach have argued that school and university curricula should be reformed to improve the basic statistical literacy and numeracy skills of laypeople and professionals alike:

> We need to change school curricula. Our children learn the mathematics of certainty, such as geometry and trigonometry, but not the mathematics of uncertainty, that is, statistical thinking. Statistical literacy should be taught as early as reading and writing are. (Gigerenzer 2010, p. 469; see also Bond 2009)

Statistical illiteracy, if not addressed, leaves citizens, patients, and doctors vulnerable to "techniques that deliberately and insidiously exploit limited statistical literacy in order to convince the audience that they are at high risk of illness" (Gigerenzer et al. 2007, p. 71), and ultimately renders the ideals of informed consent and shared decision making unattainable (Gigerenzer 2010, p. 469).

Teaching Core Competences

Another class of boost policies identifies and corrects specific skill and knowledge deficits with far-reaching consequences for health, wealth, and happiness. These policies do not aim to inform and educate on anything and everything. Rather, they offer ways of achieving a high return on a small amount of smart information in domains in which everybody makes choices (e.g., medical, dietary, health, or financial affairs), and in which the ABCs of key factual and procedural knowledge can be identified and taught. Take, for example, the medical domain. Worldwide, cardiovascular diseases are the leading cause of death and a significant cause of chronic disability. To be most effective, medical interventions need to be performed within a few hours of a heart attack or stroke. Ideally, citizens should therefore both be able to recognize the symptoms of heart attack and stroke and know how to respond. However, as a new representative survey in nine European countries shows, most respondents recognized only few of the respective symptoms (Mata et al. 2014). Furthermore, immediately calling an ambulance—the option with the highest chance of ensuring effective, timely treatment—was the most frequently endorsed response in only three of the nine countries. Proponents of boost policies argue that providing people with elementary knowledge of the cardinal symptoms of stroke and heart attack and of the single most important response—calling 911—

will considerably decrease death or loss of quality-adjusted life years from these diseases. In such cases, a limited but targeted skills has high leverage.

Provision of information is, of course, a policy intervention that on the face of it is not specific to the boost approach but is a widely employed social policy instrument. The emphasis, however, is not the typical one in many information campaigns, namely, *persuading* people to change their preferences, thus, for instance, declaring their willingness to donate their organs, drink less or reduce their carbon footprint by using environmentally friendly methods of getting around. Instead, the objective is to equip people with key skills to perform actions such as responding to a medical emergency.

Decisions Under Uncertainty: Designing Smart Strategies

Changing representations of statistical information such that they match the algorithms of the human mind (e.g., natural frequencies) works in situations in which the risks are known and can be explicitly stated (e.g., decisions from description; Hertwig and Erev 2009). It does not work in the many situations in which risks have not been measured or are not even measurable. For these situations, the boosting approach proposes the design of simple, but highly efficient, cognitive strategies to support better decisions. These include fast-and-frugal trees (FFT), a special kind of decision tree with at least one end node after each decision node (e.g., Martignon et al. 2003, 2008). FFTs have several advantages over full trees (Luan et al. 2011), including frugal demands on information, a simple decision rule, and robustness (i.e., stable performance when predicting on the basis of low-quality or unknown data); they thus simplify and speed up the decision process.

Take, for example, Jenny et al.'s (2013) FFT, designed to boost the performance of physicians screening patients for clinically depressed mood. General practitioners commonly use a 21-question depression questionnaire that generates a weighted average of all available cues. Jenny et al. simplified this time-consuming tool into a decision tree consisting of just four binary (yes vs. no) questions (see Fig. 3). When tested against two compensatory models (which, on average, use almost five times as many cues), the FFT performed as well as both competitors in detecting depressed mood from epidemiological data.

Another class of strategies supports better decisions by harnessing the wisdom of the crowd (e.g., Hertwig 2012) or the wisdom of crowds within one mind (Herzog and Hertwig 2009, 2013, 2014). Let us consider another medical example. Surrogate decisions are decisions in which someone is asked to make potentially consequential decisions on behalf of an incapacitated patient (e.g., a patient with a brain injury). Frey et al. (2014) found that respondents clearly prefer to make these decisions collectively (i.e., together with other family members), rather than by applying the individual decision-making approach implemented in the nearest-relative hierarchy that is prescribed in many countries (see legend in Fig. 4).[6] Furthermore, they found

---

[6] It could be argued that the collective approach is consistent with the legal approach of a hierarchy of nearest relatives to the extent that the legally assigned (or patient-designated) surrogate can always consult others. This is correct, but the surrogate is not obliged to consult anybody else, nor does he or she need to take others' opinion into account should their opinion differ from his or hers.
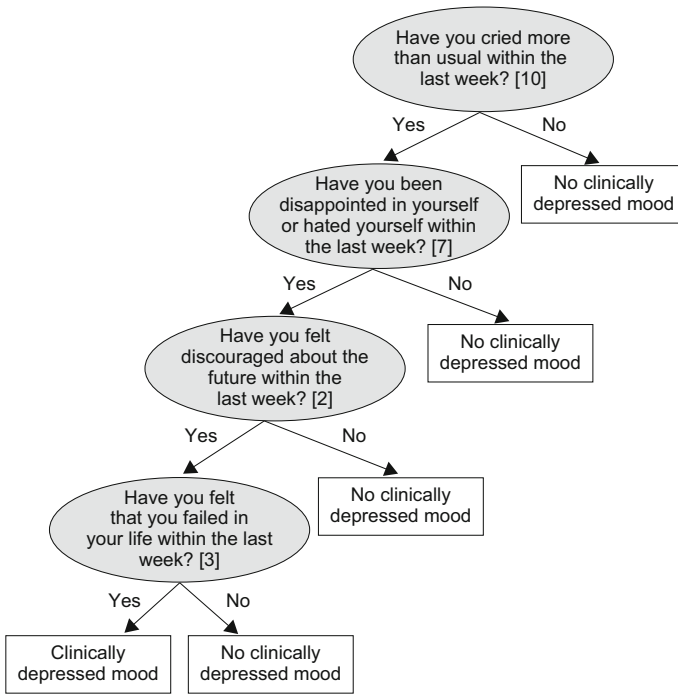
Fig. 3 Boosting general practitioners' ability to screen for depressed mood based on a fast-and-frugal decision tree. The tree poses four simple questions (e.g., "Have you cried more than usual within the last week?"). Each question offers at least one end node. If any of the first three questions is answered with "no," the tree categorizes this person as not having clinically depressed mood, and all subsequent questions are then ignored. The tree can stop cue inspection at any level (as there is at least one exit on each level)

that implementing the collective approach did not compromise accuracy, which was measured in terms of correct predictions of a hypothetical patient's treatment preferences (stated explicitly in an advance directive or living will). In fact, accuracy increased when the prediction of the treatment preference (i.e., treatment vs. no treatment) proved especially difficult (i.e., preferences were equally divided, with half of hypothetical patients preferring treatment and the other half, no treatment; Fig. 4).

What follows from these results? In light of people's evident high regard for shared surrogate decision making and the fact that it does not compromise accuracy, legal frameworks could explicitly acknowledge the collective approach in at least two ways: In their advance directives, people could be prompted to designate not only a single surrogate, but (if they wish) a group of family members or friends who would be asked to share the burden of making a surrogate decision. In addition, the framework for legally assigned surrogates could be modified such that it encourages the legally assigned surrogate to make an *active choice* between either a collective surrogate decision (involving close relatives and based on a recommended
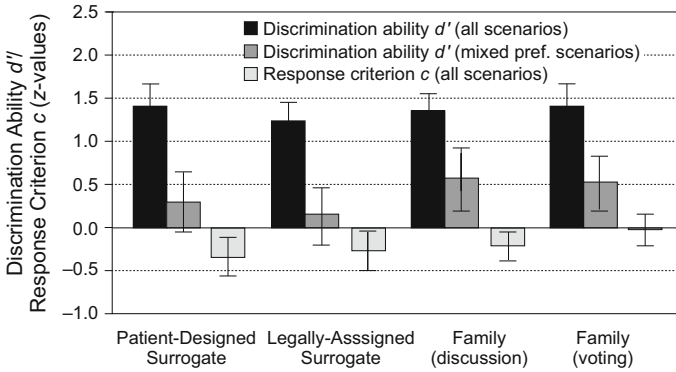
Fig. 4 Four approaches to making surrogate decisions: (1) by a person previously designated by the patient; (2) by a person who is legally assigned according to a hierarchy, starting with the patient's spouse, followed by the patient's adult children, parents, and adult siblings; (3) by multiple family members who collectively discuss the situation and aim to find a consensus; (4) by multiple family members who first individually consider the patient's situation and then vote privately to find a majority decision. The *bars* show the discrimination ability (d') of the different approaches to distinguish between treatment and no-treatment preferences. In difficult medical scenarios with mixed preferences, only the shared approaches (the "family" conditions) performed clearly better than chance. In addition, only the family-voting approach displayed a neutral (rather than a biased) response criterion c. The response criterion reflects a decision-maker's tendency to infer either a treatment preference or a no-treatment preference, independent of his or her discrimination ability (d')

procedure) or an individual surrogate decision. Alternatively, the collective approach could be declared the *default*.

Proponents of the nudge approach also recommend employing social influences (social nudges such as peer pressure). The aim of this approach is not to co-opt the wisdom of the crowd to empower more accurate inferences, however, but to prompt behavior change by, for instance, providing or withholding social information. For example, Thaler and Sunstein (2008, p. 68) suggested: "If you want to nudge people into socially desirable behavior, do not, by any means, let them know that their current actions are better than the social norm."

To conclude, at least three types of policies can be distinguished within the boost approach. The first is to foster risk competence in situations in which risks are known and measurable. This can be achieved by choosing representations of statistical information that match the cognitive algorithms of the human mind ("natural frequencies") or by enabling people to translate statistical risk information (often presented in terms of probabilities and percentages) into transparent and understandable form. The second is to identify the limited core of factual and procedural smart knowledge that constitutes health, risk, dietary, or financial literacy and to boost people's competence by teaching them these domain-specific ABCs. The third policy approach is to build and teach simple, intuitive, and efficient heuristics (e.g., FFTs) to support decisions in a wide range of situations in which knowledge about risks is incomplete and uncertain. All three policy approaches, but especially the first and third, can be used to boost the competence of experts and non-experts alike (see for instance, Hautz et al. 2015).

The very fact that these policies have been proposed indicates that proponents of the underlying SH program do not deny that people are not perfect thinkers and, at times, make bad decisions (for a variety of reasons). However, the difference to the H&B program is that these difficulties are not assumed to be so impervious to change that they have to be exploited rather than overcome. The starting premise of the boost approach is that these difficulties can be addressed by training, information, education, better decision strategies, and better representations. The nudge approach, in contrast, presupposes that these cognitive deficiencies are difficult or costly to overcome, and therefore recommends their skillful manipulation to facilitate better choices.

Boosters and nudgers sometimes appear to propose identical policies. Take, for example, the default-setting policy, which uses the practice of setting a default option (i.e., an option that the chooser receives if she does not make an active choice) to influence choice (see Sect. 2). Studies conducted by proponents of both approaches—of choice situations as diverse as organ donation (Johnson and Goldstein 2003), retirement savings, and energy provider choice (Pichert and Katsikopoulos 2008)—have shown that setting a default significantly increases the probability that an agent will end up with that option. On the same lines, one conceivable policy change in response to the analysis of surrogate decision making reported above (Frey et al. 2014) could be a change in the default setting.

In these cases, it might seem that nudge and boost policies share the same means. Yet a more detailed analysis shows that this is not the case. The SH and H&B programs tend to disagree about the underlying *causal mechanism* that explains the relationship between default setting and changes in choice distribution. H&B researchers tend to explain default effects in terms of "inertia, status-quo bias, or the 'yeah, whatever' heuristic" (Thaler and Sunstein 2008, p. 83). This kind of explanation stresses the biasing features of setting any default, thus revealing the policy to be *re*biasing. SH researchers, in contrast, explain default effects in terms of the implicit recommendation or endorsement effect (e.g., Gigerenzer and Brighton 2009, p. 130; see McKenzie et al. 2006). This kind of explanation stresses the genuine social information contained in the default, thus describing the behavioral change in response to the default as consisting in a learning effect, and hence revealing the policy to be *de*biasing. Thus, even in cases where nudgers and boosters propose the same policy, their respective mechanistic interpretation of the intervention distinguishes the distinct goals they pursue with it.

Necessary Policy Assumptions

In the previous section, we described paradigmatic nudge and boost policies. We now systematically investigate the implicit and explicit assumptions of these policies. In particular, we discuss their assumptions about (1) the errors they seek to counteract; (2) the goals they address; (3) the characteristics of the people they aim to help; and (4) the characteristics of the people who design and employ them. We focus especially on those assumptions (all listed in Table 1) that distinguish between the two kinds of policies—that is, assumptions that are made by one kind

Table 1 Eight assumptions of the nudge and boost approaches

	Nudge	Boost
Cognitive error awareness		
Must the decision maker be able to detect the influence of error?	No	Yes
Cognitive error controllability		
Must the decision maker be able to stop or override the influence of the error?	No	Yes
Information about goals		
Must the designer know the specific goals of the target audience?	Yes	No
Information about the goals' distribution		
Must the designer know the distribution of goals in the target audience?	Yes	No
Policy designer and cognitive error		
Must experts be less error-prone than decision makers?	Yes	No
Policy designer and benevolence		
Must the designer be benevolent?	Yes	No
Decision maker and minimal competence		
Must the decision maker be able to acquire trained skills?	No	Yes
Decision maker and sufficient motivation		
Must the decision maker be motivated to use trained skills?	No	Yes

of policy, but not the other. The resultant set of assumptions will provide the basis for Sect. 5, in which we investigate the link between the respective policies and the theoretical commitments of the H&B program versus the SH program. To what extent do the respective policy approaches follow from and show theoretical coherence with the two research programs?

What is the Nature of Cognitive Errors?

Nudge and boost policies reflect differing assumptions about the errors they intend to counteract.[7] In particular, they disagree on whether and to what extent cognitive errors such as base rate neglect and overconfidence, the use of (sometimes) fallacious heuristics such as representativeness, or phenomena such as inertia occur *automatically* due to the way human cognition works. Automaticity of error is suggested by the analogy between cognitive and visual illusions that is often drawn in the H&B program. It is also echoed in Kahneman's (2011) System 1–System 2 view of human cognition. In this view, errors are the result of a System 1 that "operates automatically and quickly, with little or no effort, and no sense of voluntary control" (p. 20). Automaticity of judgment and behavior can be decomposed into (among other components) *awareness* and *controllability* (Bargh

---

[7] The SH program does not endorse the strong language used by (some) proponents of the H&B approach to suggest that human reasoning is at times severely deficient (see Lopes 1991). One reason is that terms such as "cognitive illusions" presuppose the existence of a clear and unambiguous normative benchmark—an issue that has been hotly debated between the two programs (e.g., Gigerenzer 1996; Kahneman and Tversky 1996). Here, we use the more descriptive term "error" instead of "illusion" or "bias".

1994). Awareness concerns the degree to which people are conscious of how stimuli are categorized and how these categorizations influence judgments and behavior. Applied to the context of cognitive error, the question is: Are people aware or could they become aware of the error they are about to exhibit and, by extension, of the error-producing process (error awareness, Table 1)? For illustration, consider Frederick's (2005) example of a nearly inevitable bias:

> A bat and ball cost $1.10. The bat costs one dollar more than the ball. How much does the ball cost?

In order to realize that the answer that immediately comes to mind ($0.10) is wrong, an individual must resist accepting the 'automatic' answer as correct. To this end, some awareness of and insight into the error and the error-producing process appears necessary—the individual may otherwise see no need to replace it with a better process and response. Nudge policies do not demand that kind of awareness. The effectiveness of a default change, time-horizon shift, or reframing does not call for—and may in fact be negatively affected by—the target person's ability to consciously access the cognitive process that produces the error. In the boost approach, in contrast, which seeks to equip people with better ways of reasoning and deciding, lack of such awareness poses a problem. We return to this conceptual issue in the next section.

Another attribute of automaticity, according to Bargh (1994), is controllability—in the context of our discussion, the controllability of error (Table 1); that is, the ability (or lack thereof) to stop or override the influence of a process on judgment and behavior. For example, many visual illusions cannot be 'undone.' Although people may become aware of them and come to intellectually understand that what they are 'seeing' is not in fact a veridical reflection of the world (e.g., the checker shadow illusion; see http://persci.mit.edu/gallery), the illusions typically persist. If the cognitive system, and in particular, the error-producing processes were, to use Fodor's (1983) terminology, as informationally encapsulated and inaccessible as the visual system, then people might intellectually understand that their reasoning is wrong, but remain captive to the error. They would be unable to stop or override the influence of the process that produces it. Nudge policies, which treat biases as useful allies, would not mind such lack of controllability; boost policies would, however, be challenged by it. It is thus not surprising that proponents of boost policies typically do not subscribe to the analogy of cognitive illusions with visual illusions or to Kahneman's (2011) System 1–System 2 view of human cognition. In sum, the two research programs make drastically different assumptions about the nature of cognitive errors.

What are People's Goals and are They Homogenous?

Policies have goals, and policy designers implicitly or explicitly make assumptions about the goals of their targeted audience. Both nudgers and boosters *require* information about people's goals, but to a very different extent (Table 1). Nudge policies are concerned with specific, context-dependent goals—namely the goals toward which a targeted audience is to be nudged (e.g., employees are nudged

toward contributing more to their retirement savings; the public is nudged toward a higher consent rate for organ donation). Consequently, designers of nudge policies need to know whether the assumed goal is indeed one that the target audience has (for more detail, see Grüne-Yanoff 2012). Without this information, the nudge would lack legitimacy and would risk being arbitrary or implementing special interests. This information can take various forms, including revealed preferences, happiness indices, and general welfare considerations.

Information about the existence of people's goals does not suffice, however. Nudge designers also need to know how widely they apply across the target population. If a nudge is a one-size-fits-all measure, it requires that everybody in the target population share the same goal. If there is goal heterogeneity, however, and, for instance, only some people truly prefer contributing more to their retirement savings, whereas others prefer trading off retirement savings against investments in their children's education, property, or fewer working hours—the nudge needs to be able to discern between people with different goals and to steer them differentially toward their 'optimal' option.

To what extent do boosters need to know their target audience's goals? The goals of boost policies vary from the broad, such as fostering statistical literacy, to the more specific, such as designing strategies to help doctors to screen people for depressed mood (Jenny et al. 2013), prescribe antibiotics to children (Fischer et al. 2002), or allocate emergency room patients swiftly and accurately (Beglinger et al. 2015). Specific strategies are typically designed for professionals (e.g., general practitioners, emergency physicians), and it is they who typically identify the need for a decision aid and outline its goals. In fact, the existence of an explicit goal is the starting point of the design process; without it, no tool would be commissioned. Identification of a highly consequential knowledge gap may also suggest a goal. In contrast, the broad boost policy goal of fostering statistical literacy does not pursue a specific objective, such as changing participation rates in cancer screenings. Rather, it aims to increase laypeople's statistical reasoning competence as a prerequisite for realizing their own goals in light of the often misleading, confusing, or even manipulative information conveyed, for instance, by a health-care system in which patients' and health-care professionals' interests are no longer (fully) aligned (see Gigerenzer et al. 2007). Here, the choice of goals ultimately resides with the informed patient and, more generally, the citizen. In sum, boost policies, unlike nudge policies, require only minimum knowledge of goals (e.g., whether statistical literacy and shared decision making are desired goals), and no knowledge of how they are distributed in the population. This is because boost policies typically leave the choice of goals to people themselves (e.g., it remains the sole decision of the statistically literate person whether or not to participate cancer screening) or design effective tools in response to a demand expressed by professionals.

How 'Error Free' is the Policy Designer?

If cognitive biases are ubiquitous, automatic, and thus difficult to evade, to what extent can policy designers be assumed to be immune to the very cognitive errors they seek to exploit and counteract in others (Table 1)? In nudge policies, the choice architect

needs to be fully informed about the target population's goals. Specifically, nudge policies assume that the choice architect is familiar with the target population's own better selves—their goals and the distribution of those goals—and with the effect of the nudge on that population. Without this knowledge, the policy can neither be designed nor justified, nor can its effectiveness be assessed. Yet, information about goals, the behaviors instrumental to achieving them, and the evaluation of effectiveness are precisely the domains in which the H&B program has identified cognitive errors. In order to implement the preferences of a population's own better selves, policy designers (choice architects in the terminology of the nudge approach) thus need to be immune to cognitive errors or at least less likely to succumb to them than laypeople are. Only then can they detect errors, redesign the choice environment to counteract them, and evaluate the extent to which the nudge is effective.

Admittedly, the policy maker's decision problem is often different from that of the subjects of the policy: citizens have to decide, for instance, how much to save for retirement; policy makers, what savings policy it is best to adopt. The fact that people are myopic about the former does not necessarily imply that they are myopic about the latter. However, the concern is that the diagnosed errors also apply, mutatis mutandis, to the decision problems typically facing policy makers. For example, anchoring and insufficient adjustment affects evaluation of ends (hence possibly also policy goals), loss aversion affects choice of means (hence possibly also policy means), and confirmation bias affects evaluation of effectiveness (hence possibly also policy effectiveness; see, e.g., Kahneman et al. 1999).

For boost policies to work, policy designers need to be (1) competent at identifying domains in which people experience difficulties in reasoning statistically, understanding numbers, and evaluating risks, or in which they operate under highly consequential knowledge gaps and misconceptions, and (2) able to devise and implement policies that develop people's competences. A bad designer of boost policies may create a flawed decision tool or bungle a teaching tool for statistical literacy. The difference to the nudge tradition is this: The target audience of a boost policy may, in principle, be able to detect that the policy is defective and ask for it to be remedied. In contrast, nudge policies in at least some settings can be considered "hidden persuaders" (Smith et al. 2013). The target audience, unaware of being nudged and otherwise biased, will not be able to offer feedback on a dysfunctional nudge. This asymmetry assigns more responsibility (and power) to the policy designer in the nudge approach than in the boost approach. Moreover, it generalizes from dysfunctional interventions to ill-intentioned interventions. Seeing through ill-intentioned interventions is much more difficult if the intervention is designed to work without the explicit awareness of the targeted individuals. Therefore, the nudge approach rests more than the boost approach on the assumption that policy designers have the welfare of the population at heart—that is, that policy designers are benevolent.

Are Citizens Minimally Competent and Motivated?

What kinds of assumptions do the two policy approaches make about the targeted individuals? Boost policies require a minimum of cognitive abilities from the agent whose competences are to be improved (Table 1). In particular, the individual is

assumed to possess some arithmetic skills, the ability to generalize from the training to new contexts (see representation training; Sedlmeier and Gigerenzer 2001), and the ability to execute simple decision strategies (e.g., decision trees). Furthermore, the individual is assumed be able to muster enough motivation to implement the skills and cognitive tools acquired—that is, to overcome inertia, old habits, or the desire to not engage in shared decision making (e.g., between patient and doctor). If the target audience is not motivated to apply new skills and cognitive tools (e.g., decision trees), boost policies are unlikely to be effective (though exceptions are conceivable). Nudge policies require no such minimal competence and motivation, as they make no attempt to equip people with new skills and cognitive tools.

To conclude, we have analyzed what may be deemed the *necessary* assumptions underlying nudge and boost policies and, by extension, the requirements that need to be met for these policies to be successful. Table 1 summarizes our findings. This list—to our knowledge, the first of its kind—may not be complete, and our reading may not be uncontroversial. However, the list offers a contrastive approximation of the differences between the two policy approaches, and it provides a basis for the next step in our analysis: We now investigate the link between the policies and the theoretical commitments expressed in the respective research programs of bounded rationality (H&B vs. SH) by examining the extent to which the research programs imply the necessary policy assumptions.

Coherence of Theory and Policy

Do the proponents of nudge and boost policies propose drastically different policy approaches because of their respective theoretical commitments? We seek to answer this question by investigating to what extent the H&B and SH research programs, respectively, support the assumptions underlying the two kinds of policies. In particular, we ask whether the assumptions listed in Table 1 are implied by the theoretical commitments of the H&B and SH programs. In other words, we investigate the *theoretical coherence* of the H&B/nudge and SH/boost combinations. In the following, we discuss the theory-to-policy mapping for each assumption in turn.

Error Awareness and Controllability

The first two assumptions in Table 1 concern awareness and controllability of errors. Boost policies aim to help people overcome errors—for instance, by enhancing their statistical literacy (when risks are measurable and can be communicated). This aim seems to imply that people do, in principle, realize when they are about to commit a bias. Yet awareness of error depends on how one conceptualizes its origins. If, as in the H&B program, an error is regarded as an inevitable product of the fast and automatic System 1 (see Kahneman 2011), then its (pending) occurrence can reach awareness only if the decision maker engages the more reflective and effortful System 2. This system acts like a teacher identifying a student's errors and rectifying the error-producing process. But System 2 is also

assumed to be quiescent, and it tires easily. Too often, instead of analyzing the proposals of System 1, System 2 is content with the solutions it offers, and decision makers remain unaware that they are about to commit an error. In theory, decision makers are assumed to be able to stop or override heuristic processes resulting in error; in practice, the languor of System 2 means that they often fail to do so, as highlighted in the following quote:

> What is perhaps surprising is the failure of people to infer from lifelong experience such fundamental statistical rules as regression toward the mean … Statistical principles are not learned from everyday experience. … people do not learn the relation between sample size and sampling variability, although the data for such learning are abundant. (Tversky and Kahneman 1974, p. 1130)

How does the SH program respond to this challenge? To the extent that its proponents maintain that error can be corrected, they have to address both key assumptions of the H&B program, namely that people remain largely unaware of errors they commit and that they are unable to stop or override the error-producing processes. There are different ways to make this argument. One would be to retain the distinction between System 1 and System 2 and to argue that System 2 is not as quiescent as portrayed above. The proponents of the SH program do not adopt this system distinction (see Kruglanski and Gigerenzer 2011); rather, they suggest that not all heuristics, and not all processes enlisted by a heuristic, are executed *automatically*. Some heuristics can be used without the awareness of the user (e.g., the gaze heuristic; McLeod and Dienes 1996); others can be explicitly learned and deliberately employed (e.g., imitate-the-majority). The level of awareness can be such that people are, in principle, aware that not all answers produced by a heuristic are correct, and that a heuristic's judgments sometimes need to be stopped or overridden. Take, for illustration, the recognition heuristic (Goldstein and Gigerenzer 2002):

> If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value with respect to the criterion.

This heuristic relies on two competences: the ability (or lack thereof) to recognize objects and the ability to evaluate whether the recognition information should be used as the criterion to select between two objects, given the structure of the environment. Pachur and Hertwig (2006) found evidence that people frequently overrule the inference suggested by the recognition heuristic in environments in which there is little relationship between recognition and the criterion. Furthermore, the decision to suspend the recognition heuristic appears to exact the cost of longer response times (relative to judgments based on the recognition heuristic), suggesting that the heuristic is not used automatically. Similarly, Volz et al. (2006) provided neuroimaging evidence for the operation of an evaluation process that checks whether the solution recommended by the recognition heuristic should be accepted in a given task. Specifically, the authors observed activation in the anterior frontomedian cortex (aFMC), linked in earlier studies to evaluative judgments and

self-referential processing. This activation provides evidence against the interpretation that the heuristic's use is purely 'automatic.'

A second response to the assumption that people are largely unaware of cognitive errors and generally unable to override them is to challenge the interpretation that errors are the paradigmatic product of a cognitive "machine for jumping to conclusions" (Kahneman 2011, p. 79). Specifically, an error may not (or not only) be the first answer that leaps to mind, but a last-resort response produced once all other attempts at solving a problem have failed. Consider, for instance, investigations of people's competence to reason in accordance with Bayes' rule (e.g., Kahneman and Tversky 1972). Respondents are typically provided with a set of data including prevalence information (e.g., the prevalence of colorectal cancer, 0.3 %), likelihood information (e.g., the sensitivity and false-alarm rate of a screening test, say, 50 and 0.3 %, respectively), and a test result (e.g., positive). They are then asked: What is the probability that someone who tests positive actually has colorectal cancer? Typically, no immediate answer comes to mind. Indeed, there is considerable variability in the quantitative answers respondents ultimately give (Gigerenzer and Edwards 2003; see also Brunswik 1956). In visual biases, in contrast, a single (erroneous) response typically dominates.

In sum, according to the SH program, heuristics do not reside in an automatic System 1; the program suggests that evaluative processes intervene between the production of a (heuristic) answer and its final acceptance and, last but not least, that cognitive errors (such as base-rate neglect) may be a last-resort solution, rather than the first response that springs to mind. The theoretical commitments of the SH program thus imply that people can, at least in principle, become aware of an error before committing it, and are hence consistent with the assumption that erroneous cognitive process can be recognized and overridden (Table 1).

Goals and Heterogeneity of Goals

The third and fourth assumptions concern the goals of the people who are supposed to be nudged and the distribution of those goals in the population. Nudge policies require information on both in order to justify the choice of the option toward which people are to be nudged. The nudge approach often appears to follow the logic of standard welfare economics, identifying policy goals consistent with preference satisfaction and wealth maximization. In other cases, its proposed outcomes are gauged in terms of happiness measures (Kahneman and Krueger 2006) or material indices (Posner 1979). In the following, we restrict our discussion to preferences as goals. We argue that the very theoretical commitments of H&B imply that identifying decision makers' goals is, for a variety of reasons, an error-prone undertaking.

There are two major approaches in the literature on preferences. The first, dominant in mainstream economics, is that people have well-defined preferences that are revealed in the choices they make, and that they are aware of those preferences (e.g., Freeman 1993, p. 7). The second, dominant in psychology, is that preferences or objects of any complexity and novelty are often constructed—not merely revealed—in the generation of a response to a judgment or choice task. From this perspective, people use a variety of methods (heuristics) to construct

preferences. Which method is used is contingent, among other factors, on the problem (task and context), person (knowledge, ability, goals), and social context (accountability, group membership). The constructive nature of preferences implies, and is implied by, the fact that expressed judgments and choices are highly contingent on seemingly minor changes in response modes (e.g., choice vs. pricing), information displays, descriptions of choice options (e.g., Fox et al. 1996; Rottenstreich and Tversky 1997), frames, context, and emotions.

Proponents of the H&B program have often doubted the existence of well-defined preferences and advocated for the notion of constructed preferences. Tversky (1996), for instance, wrote that "violations of context dependence indicate that people do not maximize a precomputed preference order, but construct their choices in light of the available options" (p. 17). But even if the H&B program were to endorse the existence of well-defined preferences, the theoretical commitment of the program is that people tend to violate the basic axioms of rational choice models (e.g., transitivity), which makes inferring their preferences from their choices highly problematic. Measuring constructed preference is inherently reactive, and the process of preference construction is subject to cognitive errors (e.g., framing). Its results are therefore to some extent arbitrary. One way to deal with these problems would be a theory of well-constructed preference in combination with tests of predictive validity, construct validity (e.g., multitrait–multimethod approach), and process validity. To the best of our knowledge, no such theory exists, although steps have been made toward it (e.g., Payne et al. 1999).

Apart from the problems inherent in identifying goals, nudge policies require information about the distribution of those goals in the population, so that each person can be steered toward his or her 'optimal' option. Such person-to-nudge matching works trivially if each member of the population has the same goal and is afflicted by the same error, and if the same nudge steers everybody towards the optimal option. Matching become less trivial if goals differ between members of the population, and people are afflicted by errors to different degrees. In this case, the nudge policy needs to be able to trigger differential effects contingent on different goals and/or degrees of error-proneness.

The H&B program appears to support a trivial person-to-nudge matching. Specifically, there is some evidence that (at least some) advocates of the H&B program believe that errors are largely uniform. Thaler (1991), for instance, argued that "mental illusions should be considered the rule rather than the exception" (p. 4). Relatedly, the assumption is that representativeness, availability, and anchoring-and-adjustment are heuristic principles that "*people* rely on" (Tversky and Kahneman 1974, p. 1124; emphasis added), not because they are insufficiently motivated or prone to wishful thinking (p. 1130), but because of the way the cognitive architecture works: "The subjective assessment of probability resembles the subjective assessment of physical quantities such as distance or size" (p. 1124). Similarly, a "defining property of intuitive thoughts is that they come to mind spontaneously, like percepts" (Kahneman 2003, p. 1452). In other words, in the same way as the visual system produces visual illusions that nobody can escape, the cognitive system produces low-variance cognitive illusions (errors). Further assuming that goals are uniformly distributed in the population would liberate

nudgers from worrying about goal or bias heterogeneity in a population. In this ideal world, a-one-size-fits-all nudge could be prescribed.[8]

In sum, the H&B program offers ample reasons to doubt that people's goals are easily and veridically determinable. Consequently, it does not appear to be consistent with the nudge approach's crucial assumption that this information is available. Moreover, if information about goals is hard to come by, it will, by extension, be (even more) difficult to find out about the goals' distribution in the population. Even assuming uniformity of errors, the challenge of good person-to-nudge matching persists to the extent that goals are heterogeneous. Consequently, H&B's theoretical commitments do not unambiguously imply that information about goals and goal distribution is easily available and accessible.

Policy Designers, Errors, and Benevolence

Nudge policies assume that the designers of error-exploiting policies do not to succumb to errors—or at least they are not as error-prone as average citizens. Their considerable leverage over the targeted audience is otherwise difficult to justify. If experts were subject to the same cognitive errors or errors in affective forecasting (e.g., focalism, immune neglect; Meyvis et al. 2010) as average Joe, there would be no grounds for assuming that nudgers are able to accurately infer the chooser's preference or to steer the chooser toward that preference (for a critical review of proposed policies from this perspective, see Glaeser 2006).

Interestingly, proponents of the H&B program have often highlighted that experts also succumb to cognitive errors (e.g., Tversky and Kahneman 1971). Early on, Tversky and Kahneman (1974), for instance, wrote:

> The reliance on heuristics and the prevalence of biases are not restricted to laymen. Experienced researchers are also prone to the same biases—when they think intuitively. For example, the tendency to predict the outcome that best represents the data, with insufficient regard for prior probability, has been observed in the intuitive judgments of individuals who have had extensive training in statistics. (p. 1130)

The crucial qualification in this statement is, of course, "when [experts] think intuitively." Following the logic of the H&B framework, experts' System 2 has the same shortcomings as laypeople's, leaving them similarly reliant on intuition (System 1) rather than on effortful deliberation. Judging from the numerous publications in the tradition of the H&B program suggesting that experts in business, medicine, or politics cannot escape the reach of biases and bias-producing heuristics (e.g., Bornstein and Emler 2001; Heath et al. 1998; Kahneman and Renshon 2007; Klein 2005; Malmendier and Tate 2005), the framework evidently assumes experts to be

---

[8] Some proponents of nudge policies consider the possibility that some people may be immune to certain errors, thus admitting a kind of population heterogeneity. *Asymmetric* paternalism (Camerer et al. 2003), for example, assumes that some members of a population may be fully rational, and hence not need a nudge that others require. Consequently, it seeks to devise policies that affect only those whose judgments are erroneous. Yet even asymmetric paternalism assumes that those who are subject to error are affected in such a way that a uniform nudge can steer them toward their optimal option.

afflicted by the same or similar shortcomings as average citizens. Consequently, the design of nudges is also vulnerable to the influence of errors.

The benevolence (or lack thereof) of policy designers also warrants consideration. Nudge policies are orthogonal to the H&B research program on bounded rationality on this issue; thus, coherence between theory and policy cannot be assessed on this dimension. Thaler and Sunstein (2008, p. 240) have acknowledged the possibility of evil or ill-intentioned nudges, but they seem to envision primarily policy designers in the private sector. With regard to the public sector, they appear to rely on the power of the democratic system to check and control public officials' actions and thus ensure that nudge policies are (mostly) in line with collective welfare. Rebonato (2012) discussed this reliance critically and pointed out that "for the controls that the libertarian paternalists mention to be effective, the full engagement of the deliberate System-II is required" (p. 221).

In the boost approach, in contrast, there is less need for a functioning system of democratic checks. Indeed, some boosts are meant to enhance the ability of consumers and patients to see through misleading private sector interventions. Take, for instance, the phenomenon of "mismatched framing": reporting the benefits of an intervention (e.g., a new medical treatment) as relative risk reduction, but reporting harms as absolute increases in risk, thus making the benefits look much bigger and the harms much smaller. Mismatched framing can already happen before a treatment reaches the market. According to a study of three major medical journals (*British Medical Journal*, *The Lancet*, and the *Journal of the American Medical Association*), a third of the studies published from 2004 to 2006 that reported both benefits and harms used a different metric for each (Sedrakyan and Shih 2007). If, as Thaler and Sunstein (2008) have argued, nudges are indeed omnipresent (because the private sector seeks to nudge people in directions that promote its selfish goals), then the boost approach's answer is not to enter a race in which (supposedly) benevolent governmental nudgers scramble to compete against millions of private sector nudgers. Instead, a boost policy aims to enhance individuals' ability to detect misleading communication, such as mismatched framing, thus honing the meta-ability to recognize nonbenevolent policies and policy designers.

The Target Person's Minimal Competence and Sufficient Motivation

Boost policies require minimal abilities on the part of the person whose competences are to be improved. The SH program is clearly committed to the existence of such abilities. It suggests that a substantial segment of the cognitive architecture can be described in terms of a repertoire of simple decision strategies (the adaptive toolbox) of medium range (in contrast to domain-general algorithms, such as expected utility theory). These heuristics often permit good-enough or in fact surprisingly accurate performance (see Gigerenzer et al. 2011). Their computational and informational demands are simple because they are masterful exploiters—tapping into both environmental regularities and evolved cognitive, visual, motor, or other capacities of the mind and body, such as recognition, forgetting, movement tracking, numerical discrimination, and the ability to feel empathy (Gigerenzer et al. 2011; Hertwig et al. 2013). The SH program assumes

that everybody is equipped with these capacities and the ability to learn and execute SH that exploit them. It also assumes that people can learn to adapt to new problems, although these adaptation processes are not instant and often involve "good errors" (Gigerenzer 2005), and that people are able to select the appropriate heuristics for the environment thanks to basic learning processes, e.g. reinforcement learning (Rieskamp and Otto 2006) or individual cognitive representations of the structural properties of the environment (e.g., see Marewski and Schooler 2011). Moreover, there is some empirical evidence that the knowledge and skills relevant for improving judgment and decision making can indeed be taught (Sedlmeier and Gigerenzer 2001; Kurz-Milcke et al. 2011; Latten et al. 2011). The SH program is theoretically committed to the existence of a minimal competence enabling people to understand and learn, and therefore to benefit from transparent information, education, and the provision of new simple decision tools (Table 1).

Boost policies implicitly make another assumption about the decision maker, namely, that he or she is sufficiently motivated to implement the newly acquired skills (e.g., to translate one representation into another). The SH program has little to say about individual motivation. It implicitly assumes that people offered tools for better judgment and decision-making will be motivated to acquire them and employ them appropriately. This assumption is less problematic when professionals commission a specific decision tool (e.g., a decision tree). It is more problematic when the boost policy is an unsolicited bid to a potentially interested target audience (patients, parents, doctors, lawyers, etc.). Patients, for instance, may attach little value to statistical literacy and informed judgment and prefer to rely on a less effortful "trust-the-doctor" heuristic (Wegwarth and Gigerenzer 2013), even though modern health-care systems, fraught as they are with conflicts of interests, compromise this heuristic. Similarly, doctors may not be motivated to provide a patient with evidence-based information due to conflicts of interest (Wegwarth and Gigerenzer 2013). The SH program does not deny the existence of such motivations, misaligned incentives, and conflicts of interest. It has suggested ways of overcoming these obstacles in the medical domain (Gigerenzer and Muir Gray 2011; Hertwig et al. 2011). Yet it seems fair to say that the theoretical commitments of the SH program do *not* stringently imply that individuals have sufficient motivation to implement the acquired skills.

How Coherent are Policy and Theory?

Our analysis of nudge and boost policies has produced three results. First, we identified profound divergences in the assumptions underlying the two policy approaches. These assumptions, which are listed in Table 1, concern the nature of cognitive errors, the targets of policies (the individuals' goals, motivation, and minimal competence), and the designers of policies (their error-proneness or lack thereof). Second, we analyzed the extent to which the two research programs from which the two policy programs stem are theoretically committed to the assumptions on which the policy approaches rest. Specifically, nudging and boosting are informed by two opposing theoretical perspectives on bounded rationality, and their particular set of assumptions ought, in theory, to be consequences of the respective

foundational theories. Yet, and this is the third result of our analysis, this link is not always in place: The H&B program does not imply all assumptions built into nudge policies, neither does the SH program imply all assumptions built into boost policies. However, this partial disconnect between theory and policy appears to be more pronounced in the nudge approach than in the boost approach.

With respect to the nudge approach, the critical policy assumptions that are not or do not appear to be sufficiently rooted in the H&B research program concern (1) the policy designer's awareness of the target audience's goals and their distribution in the population; (2) the policy designer's (expert's) immunity to cognitive errors; and (3) the policy designer's benevolence. With respect to the boost approach, the critical policy assumption that is not sufficiently rooted in the SH research program is that people have sufficient motivation to benefit from a boost—for example, to acquire and execute a new skill. Nothing in the SH program explicitly speaks to this issue (though nothing contradicts it). Ultimately, this assumption appears to represent a reliance on an enlightened and autonomy-inspired citizenry that may or may not be justified.

There are several ways in which the proponents of the research programs and policy programs, respectively, can respond to our findings and interpretations thereof. One possible response is to simply acknowledge these disconnects and treat the respective policies as to some extent (which may differ for nudge and boost) autonomous from the theoretical parent programs. Another is to acknowledge these disconnects and examine them empirically. For instance, do people have sufficient motivation to learn, execute, and benefit from boosts? Can multitrait–multimethod approaches be found to unambiguously measure (constructed) preferences? A further possible response is to argue that the coherence between policy and theory is just one benchmark against which policies can be evaluated, but by no means the most important.

Of course, we acknowledge that theory–policy coherence is but one of several benchmarks. Others include the empirical validity of the assumptions made in the respective theoretical frameworks, and the empirical success of the policies proposed. Unfortunately, analyses of these benchmarks go far beyond the scope of this article (but see, e.g., House of Lords Science and Technology Select Committee 2011). Moreover, any single benchmark may fail to suffice. For instance, the empirical success of a policy program does not necessarily capture its overall 'success.' A nudge policy may prove effective in changing behavior in the intended direction; yet, it can count as successful only if the (auxiliary) assumption that people's true ends were unambiguously discerned also holds. Another benchmark on which a policy approach can be evaluated is its ethical implications. We conclude our inquiry with a brief discussion of this issue, showing how our analysis can also be employed in this context.

Comparing the Ethical Implications of Nudge and Boost Policies

To shape behavior, governmental policy makers can recruit the established tools of the governmental toolbox, including laws that eliminate (e.g., prohibiting goods or services) or restrict choice (e.g., outlawing smoking in public places), fiscal disincentives (e.g., taxation on cigarettes, congestion charges), fiscal incentives (e.g., tax breaks on pension contributions), nonfiscal incentives and disincentives,

and provision of information and warnings. In addition, a huge variety of behavioral tools can be derived from the rich stock of findings from behavioral research. Preceding the nudge approach, Skinner (1975) strongly advocated the use of scientific evidence about human behavior to promote better outcomes:

> The choice is clear: either we do nothing and allow a miserable and probably catastrophic future to overtake us, or we use *our knowledge about human behavior* to create a social environment in which we shall live productive and creative lives and do so without jeopardizing the chances that those who follow us will be able to do the same. (p. xvi, emphasis added).

The great achievement of the nudge movement is to have re-raised awareness that policy tools can be derived from empirical evidence about human behavior. Like Skinner's methods of behavioral modification and his controversial interpretations of 'freedom' (e.g., in terms of the absence of punishment; Skinner 1971), however, the nudge approach has faced ethical challenges, some of which are vaguely reminiscent of arguments raised against Skinner's social philosophy (e.g., Staddon 1995).

One line of criticism is that, because nudge policies intentionally take advantage of nonrational factors to influence choice, they undermine autonomy (Wilkinson 2013). Critics have conjectured that "only rational persuasion fully respects the sovereignty of the individual over his or her own choices" (Hausman and Welch 2010, p. 135). A related argument is that nudge policies exploit weaknesses of the human mind to influence choices, and that such exploitations violate conceptions of dignity (Saghai 2014). These arguments consist of two steps: The first diagnoses specific properties of the policy; the second argues that those properties (or a subset of them) violate human autonomy or dignity. Our comparative analysis shows that nudges and boosts differ with respect to the first step. As discussed in Sects. 4 and 5, in a nudge, the policy maker does not rely on the agent's ability to stop or override a targeted behavior or cognition. Instead, the nudge intervention can 'remedy' an individual's actions without the individual making an active contribution. It is this feature that leads critics to argue that nudges are manipulative, and that they violate autonomy and dignity. Boost policies do not rely on the same logic. Rather, they (mostly) depend on a person's ability to stop or override the targeted behavior and cognition; otherwise, they cannot be effective. Furthermore, boosts require minimal competences and motivations on the part of the target audience to be effective. Consequently, the criticism that nudge policies infringe on human autonomy and dignity does not apply (or applies less) to boost policies.

Defenders of nudge have countered this criticism. We next review three of their counterarguments and show that, for each of them, boost policies come off well. Thus, even if the nudge approach could respond satisfactorily to concerns about autonomy and dignity, the boost approach would appear to be preferable. The first counterargument stresses the purported inevitability of nudges. According to Thaler and Sunstein (2008), governments and organizations inevitably find themselves in the role of choice architects. Many of their mandated actions affect the context in which people make choices—even when they do not intend to influence those choices. Consequently, impediments to autonomy and dignity are already in place before nudge policies are implemented with intent:

> In many cases, some kind of nudge is inevitable, and so it is pointless to ask government simply to stand aside. Choice architects, whether private or public, must do something. (Thaler and Sunstein 2008, p. 337)

However, nudges are not the only possible intervention policy; there often is, or could be, a boost alternative. Where there is a boost alternative, the argument of inevitability thus fails.

A second response to the criticism that nudge policies infringe on autonomy and dignity is to stress their 'softness.' As Thaler and Sunstein put it:

> Libertarian paternalism is a relatively weak, soft, and non-intrusive type of paternalism because choices are not blocked, fenced off, or significantly burdened. If people want to smoke cigarettes, to eat a lot of candy, to choose an unsuitable health care plan, or to fail to save for retirement, libertarian paternalists will not force them to do otherwise—or even make things hard for them … To count as a mere nudge, the intervention must be easy and cheap to avoid. (Thaler and Sunstein 2008, pp. 5–6)

Critics of nudge have pointed out that the easy-avoidance argument fails to addresses the approach's lack of transparency (Bovens 2008; Hausman and Welch 2010; Grüne-Yanoff 2012; Wilkinson 2013). An intervention that lacks transparency can be difficult to sidestep, even if the costs of doing so are minimal. Lack of transparent and comprehensible information impedes the ability to make autonomous decisions.

Our analysis of nudges reinforces this concern. In order to be nudged, people do not need to be aware of being biased. Thus, the reason for the intervention may not be transparent to them. In contrast, boosts require people to be motivated to participate in the intervention. Without their active participation, it cannot be effective. Consequently, the boost intervention must be transparent. Thus, boost policies, more than nudge policies, support the transparency condition.

A third response to the autonomy and dignity criticism asserts that nudges do not reduce the individual's autonomy, but contribute to it. According to this argument, nudges seek to influence people's choices to make them better off as judged by themselves. By doing so, they improve the authenticity of a person's behavior and thus provide a practicable method of empowerment (White 2013, pp. 81–102; Sunstein 2014, pp. 123–42). Some critics have countered this argument by insisting that autonomy is never improved by violations of liberty, and that nudging clearly violates liberty (Mitchell 2005; Veetil 2011). This criticism does not necessarily support boosts either, because boosts violate liberty at least in the weak sense that they also aim to improve people's decisions, and seek to intervene in their deliberations to this end. Another objection to the authenticity-through-intervention defense is that it would require information that nudgers seldom have. Policies can only help people to make subjectively better choices if their authentic goals are known—and given that people may not be fully aware of those goals, it is unrealistic that policy makers will be able to access them. Instead, nudgers assume homogenous approximations of welfare-enhancing goals, thus easily turning nudges into standard paternalistic interventions that are inconsistent with the autonomy-

improvement argument (Grüne-Yanoff 2012; Rebonato 2012; Fateh-Moghadam and Gutmann 2013). This concern is also supported by our analysis. Nudge policy designers require information about and justifications of goals, because the choice of means to reach these goals ultimately lies with them. In contrast, boost policy designers seek to improve people's capacity to reach certain goals, but leave the choice of means to the decision makers themselves.

Conclusions

One of the goals of this article is to instigate a comparative debate of the pros and cons of two contrasting evidence-based policies. New arguments buttressed by empirical evidence on the workings (and potential 'side-effects') of nudge and boost policies will enrich this discussion in the future. At this point, many arguments are necessarily conceptual. We have contributed to this conceptual discussion by analyzing the necessary assumptions underlying nudge and boost policies, examining the coherence between policy and theory, and, finally, considering how the respective assumptions can inform the discussion of ethical implications. As we stated at the outset, one of us has contributed to the SH research program and to the design of boosts. We hope that this 'bias' has not overly tainted our analysis. Yet even if that were the case, we are confident that the insights and arguments presented in this article are able to shed fresh light on the new tools in public policy.

References

Akl, E. A., Oxman, A. D., Herrin, J., Vist, G. E., Terrenato, J., Sperati, F., Costiniuk C, Blank D., & Schünemann H. (2011). Using alternative statistical formats for presenting risks and risk reductions. *Cochrane Database of Systematic Reviews, 3*, CD006776.

Anderson, B. L., Gigerenzer, G., Parker, S., & Schulkin, J. (2012). Statistical literacy in obstetricians and gynecologists. *Journal for Healthcare Quality, 36*(1), 5–17.

Arkes, H. R., Gigerenzer, G., & Hertwig, R. (2014). *Coherence cannot be a universal criterion for rational behavior: An ecological perspective*. Manuscript submitted for publication.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer Jr & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1–40). Hillsdale, NJ: Lawrence Erlbaum Associates.

Beglinger, B., Rohacek, M., Ackermann, S., Hertwig, R., Karakoumis-Ilsemann, J., Boutellier, S., Geigy, N., Nickel, C., & Bingisser, R. (2015). The physician's first clinical impression of emergency department patients with non-specific complaints is associated with morbidity and mortality. *Medicine, 94*(7), e374. doi:10.1097/MD.0000000000000374

Berwick, D. M., Fineberg, H. V., & Weinstein, M. C. (1981). When doctors meet numbers. *American Journal of Medicine, 71*, 991–998.

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. R. Brown, J. B. Liebman, & D. A. Wise (Eds.), *Social security policy in a changing environment* (pp. 167–195). Chicago, IL: University of Chicago Press.

Bond, M. (2009). Risk school. *Nature, 461*, 1189–1192.

Bornstein, B. H., & Emler, A. C. (2001). Rationality in medical decision making: A review of the literature on doctors' decision-making biases. *Journal of Evaluation in Clinical Practice, 7*, 97–107.

Börsch-Supan, A. (2004). Mind the gap: The effectiveness of incentives to boost retirement saving in Europe. *OECD Economic Studies, 39*, 111–144.

Bovens, L. (2008). The ethics of nudge. In T. Grüne-Yanoff & S. O. Hansson (Eds.), *Preference change: Approaches from philosophy, economics and psychology* (pp. 207–219). Berlin: Springer.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley, CA: University of California Press.

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania Law Review, 151*(3), 1211–1254.

Cialdini, R. B. (2001). *Influence: Science and practice* (4th ed.). Boston: Allyn & Bacon.

Covey, J. (2007). A meta-analysis of the effects of presenting treatment benefits in different formats. *Medical Decision Making, 27*, 638–654.

European Commission. (2011). *Attitudes of European citizens towards the environment.* (Special Eurobarometer 365). Retrieved from http://ec.europa.eu/environment/pdf/ebs_365_en.pdf.

Fateh-Moghadam, B., & Gutmann, T. (2013). Governing [through] autonomy: The moral and legal limits of "soft paternalism". *Ethical Theory and Moral Practice, 17*, 383–397.

Fischer, J. E., Steiner, F., Zucol, F., Berger, C., Martignon, L., Bossart, W., et al. (2002). Use of simple heuristics to target macrolide prescription in children with community-acquired pneumonia. *Archives of Pediatrics and Adolescent Medicine, 156*, 1005–1008.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Fox, C. R., Rogers, B. A., & Tversky, A. (1996). Options traders exhibit subadditive decision weights. *Journal of Risk and Uncertainty, 13*, 5–17.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives, 19*(4), 25–42.

Freeman, A. M, I. I. I. (1993). *The measurement of environmental and resource values*. Washington, DC: Resources for the Future.

Frey, R., Hertwig, R., & Herzog, S. M. (2014). Surrogate decision making: Do we have to trade off accuracy and procedural satisfaction? *Medical Decision Making, 34*, 258–269.

García-Retamero, R., Galesic, M., & Gigerenzer, G. (2010). Do icon arrays help reduce denominator neglect? *Medical Decision Making, 30*, 672–684.

Gerend, M. A., & Cullen, M. (2008). Effects of message framing and temporal context on college student drinking behavior. *Journal of Experimental Social Psychology, 44*(4), 1167–1173.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review, 103*, 592–596.

Gigerenzer, G. (2005). I think, therefore I err. *Social Research: An International Quarterly, 72*(1), 1–24.

Gigerenzer, G. (2010). Collective statistical illiteracy. *Archives of Internal Medicine, 170*, 468–469.

Gigerenzer, G. (2014). Breast cancer screening pamphlets mislead women: All women and women's organisations should tear up the pink ribbons and campaign for honest information. *British Medical Journal, 348*, g2636.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*(1), 107–143.

Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: From innumeracy to insight. *British Medical Journal, 327*, 741–744.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest, 8*(2), 53–96.

Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior*. Oxford, UK: Oxford University Press.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*(4), 684–704.

Gigerenzer, G., & Muir Gray, J. A. (Eds.). (2011). *Better doctors, better patients, better decisions: Envisioning health care 2020*. Cambridge, MA: MIT Press.

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, UK: Cambridge University Press.

Glaeser, E. L. (2006). Paternalism and psychology. *University of Chicago Law Review, 73*, 133–156.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109*, 75–90.

Grüne-Yanoff, T. (2012). Old wine in new casks: Libertarian paternalism still violates liberal principles. *Social Choice and Welfare, 38*(4), 635–645.

Hausman, D. M., & Welch, B. (2010). Debate: To nudge or not to nudge. *Journal of Political Philosophy, 18*(1), 123–136.

Hautz, W. E., Kämmer, J. E., Schauber, S. K., Spies, C. D., & Gaissmaier, W. (2015). Diagnostic performance by medical students working individually or in teams. *JAMA, 313*, 303–304.

Heath, C., Larrick, R. P., & Klayman, J. (1998). Cognitive repairs: How organizational practices can compensate for individual shortcomings. *Research in Organizational Behavior, 20*, 1–37.

Hertwig, R. (2012). Tapping into the wisdom of the crowd: With confidence. *Science, 336*(6079), 303–304.

Hertwig, R., Buchan, H., Davis, D. A., Gaissmaier, W., Härter, M., Kolpatzik, K., Légaré, F., Schmacke, N., & Wormer, H. (2011). How will health care professionals and patients work together in 2020? A manifesto for change. In G. Gigerenzer & J. A. Muir Gray (Eds.), *Better doctors, better patients, better decisions: Envisioning health care 2020* (pp. 317–337). Cambridge, MA: MIT Press.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences, 13*, 517–523.

Hertwig, R., & Gigerenzer, G. (1999). The "conjunction fallacy" revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making, 12*, 275–305.

Hertwig, R., Hoffrage, U., & the ABC Research Group. (2013). *Simple heuristics in a social world*. New York, NY: Oxford University Press.

Herzog, S. M., & Hertwig, R. (2009). The wisdom of many in one mind: Improving individual judgments with dialectical bootstrapping. *Psychological Science, 20*, 231–237.

Herzog, S. M., & Hertwig, R. (2013). The crowd-within and the benefits of dialectical bootstrapping: A reply to White and Antonakis (2013). *Psychological Science, 24*, 117–119.

Herzog, S. M., & Hertwig, R. (2014). Think twice and then: Combining or choosing in dialectical bootstrapping? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 218–232.

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science, 290*, 2261–2262.

House of Lords Science and Technology Select Committee. (2011). *Behaviour change* (2nd report of session 2010–12, HL paper 179). London, UK: The stationary office. Retrieved from http://www.publications.parliament.uk/pa/ld201012/ldselect/ldsctech/179/179.pdf.

Jenny, M. A., Pachur, T., Williams, S. L., Becker, E., & Margraf, J. (2013). Simple rules for detecting depression. *Journal of Applied Research in Memory and Cognition, 2*, 149–157.

Johnson, E. J., Bellman, S., & Lohse, G. L. (2002). Defaults, framing and privacy: Why opting in-opting out. *Marketing Letters, 13*(1), 5–15.

Johnson, E., & Goldstein, D. (2003). Do defaults save lives? *Science, 302*, 1338–1339.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review, 93*, 1449–1475.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus & Giroux.

Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). *Well-being: The foundations of hedonic psychology*. New York, NY: Russell Sage Foundation.

Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives, 20*, 3–24.

Kahneman, D., & Renshon, J. (2007). Why hawks win. *Foreign Policy, 158*, 34–38.

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

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430–454.

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

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions: A reply to Gigerenzer's critique. *Psychological Review, 103*, 582–591.

Katsikopoulos, K. (2014). Bounded rationality: The two cultures. *Journal of Economic Methodology,*. doi:10.1080/1350178X.2014.965908.

Klein, J. G. (2005). Five pitfalls in decisions about diagnosis and prescribing. *British Medical Journal, 330*, 781–783.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative and methodological challenges. *Behavioral and Brain Sciences, 19*, 1–53.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review, 118*, 97–109.

Kunreuther, H., & Michel-Kerjan, E. (2011). People get ready: Disaster preparedness. *Issues in Science and Technology, 28*, 1–7.

Kurz-Milcke, E., Gigerenzer, G., & Martignon, L. (2008). Transparency in risk communication: Graphical and analog tools. *Annals of the New York Academy of Sciences, 1128*, 18–28.

Kurz-Milcke, E., Gigerenzer, G., & Martignon, L. (2011). Risiken durchschauen: Grafische und analoge Werkzeuge [Understanding risk: Graphical and analog tools]. *Stochastik in der Schule, 31*, 8–16.

Latten, S., Martignon, L., Monti, M., & Multmeier, J. (2011). Die Förderung erster Kompetenzen für den Umgang mit Risiken bereits in der Grundschule: ein Projekt von RIKO-STAT und dem Harding Center [Teaching risk literacy in elementary school: A project of RIKO-STAT and the Harding Center]. *Stochastik in der Schule, 31*, 17–25.

Levin, I. P., & Gaeth, G. J. (1988). How consumers are affected by the framing of attribute information before and after consuming the product. *Journal of Consumer Research, 15*, 374–378.

Lindsey, S., Hertwig, R., & Gigerenzer, G. (2003). Communicating statistical DNA evidence. *Jurimetrics: The Journal of Law, Science, and Technology, 43*, 147–163.

Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics, 107*, 573–597.

Lopes, L. L. (1991). The rhetoric of irrationality. *Theory & Psychology, 1*, 65–82.

Luan, S., Schooler, L. J., & Gigerenzer, G. (2011). A signal detection analysis of fast-and-frugal trees. *Psychological Review, 118*, 316–338.

Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *Journal of Finance, 60*, 2661–2700.

Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological Review, 118*, 393–437.

Martignon, L., Katsikopoulos, K. V., & Woike, J. K. (2008). Categorization with limited resources: A family of simple heuristics. *Journal of Mathematical Psychology, 52*, 352–361.

Martignon, L., Vitouch, O., Takezawa, M., & Forster, M. (2003). Naïve and yet enlightened: From natural frequencies to fast and frugal decision trees. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment, and decision making* (pp. 189–211). Chichester, UK: Wiley.

Mata, J., Frank, R., & Gigerenzer, G. (2014). Symptom recognition of heart attack and stroke in nine European countries: A representative study. *Health Expectations, 17*, 376–387.

McKenzie, C. R. M., Liersch, M. J., & Finkelstein, S. R. (2006). Recommendations implicit in policy defaults. *Psychological Science, 17*, 414–420.

McLeod, P., & Dienes, Z. (1996). Do fielders know where to go to catch the ball, or only how to get there? *Journal of Experimental Psychology: Human Perception and Performance, 22*, 531–543.

Meyvis, T., Ratner, R. K., & Levav, J. M. (2010). Why don't we learn to accurately forecast feelings? How misremembering our predictions blinds us to forecasting errors. *Journal of Experimental Psychology: General, 139*, 579–589.

Mitchell, G. (2005). Libertarian paternalism is an oxymoron. *Northwestern University Law Review, 99*(3), 1245–1277.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: Theory and practice. *Annals of Behavioral Medicine, 35*(3), 261–274.

Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 983–1002.

Park, C. W., Jun, S. Y., & MacInnis, D. J. (2000). Choosing what I want versus eliminating what I don't want: The effects of additive versus subtractive product option framing on consumer decision making. *Journal of Marketing Research, 37*(2), 187–202.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, UK: Cambridge University.

Payne, J. W., Bettman, J. R., & Schade, D. A. (1999). Measuring constructed preferences: Towards a building code. *Journal of Risk and Uncertainty, 19*, 243–270.

Pichert, D., & Katsikopoulos, K. V. (2008). Green defaults: Information presentation and pro-environmental behavior. *Journal of Environmental Psychology, 28*, 63–73.

Posner, R. A. (1979). Some uses and abuses of economics in law. *The University of Chicago Law Review, 46*(2), 281–306.

Rebonato, R. (2012). *Taking liberties: A critical examination of libertarian paternalism*. New York, NY: Palgrave Macmillan.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135*(6), 943–973.

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135*, 207–236.

Rothman, A. J., Bartels, R. D., Wlaschin, J., & Salovey, P. (2006). The strategic use of gain-and loss-framed messages to promote healthy behavior: How theory can inform practice. *Journal of Communication, 56*(s1), S202–S220.

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review, 104*, 406–415.

Saghai, Y. (2014). Salvaging the concept of nudge. *Journal of Medical Ethics, 38*, 487–493.

Sarfati, D., Howden-Chapman, P., Woodward, A., & Salmond, C. (1998). Does the frame affect the picture? A study into how attitudes to screening for cancer are affected by the way benefits are expressed. *Journal of Medical Screening, 5*, 137–140.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General, 130*, 380–400.

Sedrakyan, A., & Shih, C. (2007). Improving depiction of benefits and harms: Analyses of studies of well-known therapeutics and review of high-impact medical journals. *Medical Care, 45*, 523–528.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129–138.

Simon, H. A. (1978). Rationality as process and as product of thought. *American Economic Review, 68*, 1–16.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*, 1–19.

Skinner, B. F. (1971). *Beyond freedom and dignity*. New York, NY: Knopf.

Skinner, B. F. (1975). *Walden two*. Indianapolis, IN: Hackett.

Smith, N. C., Goldstein, D. G., & Johnson, E. J. (2013). Choice without awareness: Ethical and policy implications of defaults. *Journal of Public Policy & Marketing, 32*, 159–172.

Staddon, J. (1995) On responsibility and punishment. *The Atlantic Monthly, 275*(2), 88–94.

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

Thaler, R. H. (1991). *Quasi rational economics*. New York, NY: Sage.

Thaler, R., & Benartzi, S. (2004). Save more tomorrow: Using behavioral economics to increase employee savings. *Journal of Political Economy, 112*, 164–187.

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

Todd, P. M., Gigerenzer, G., & the ABC Research Group. (2012). *Ecological rationality: Intelligence in the world*. New York, NY: Oxford University Press.

Tversky, A. (1996). Contrasting rational and psychological principles in choice. In R. J. Zeckhauser, R. L. Keeny, & J. K. Sebenius (Eds.), *Wise choices: Decisions, games and negotiations* (pp. 5–21). Boston, MA: Harvard Business School Press.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76*(2), 105.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.

Veetil, V. P. (2011). Libertarian paternalism is an oxymoron: An essay in defence of liberty. *European Journal of Law and Economics, 31*(3), 321–334.

Volz, K. G., Schooler, L. J., Schubotz, R. I., Raab, M., Gigerenzer, G., & von Cramon, D. Y. (2006). Why you think Milan is larger than Modena: Neural correlates of the recognition heuristic. *Journal of Cognitive Neuroscience, 18*(11), 1924–1936.

Wegwarth, O., & Gigerenzer, G. (2013). Trust-your-doctor: A simple heuristic in need of a proper social environment. In R. Hertwig, U. Hoffrage, & the ABC Research Group (Eds.), *Simple heuristics in a social world* (pp. 67–102). New York, NY: Oxford University Press.

White, M. (2013). *The manipulation of choice. Ethics and libertarian paternalism.* New York, NY: Palgrave MacMillan.

Wilkinson, T. M. (2013). Nudging and manipulation. *Political Studies, 61*(2), 341–355.

World Health Organization. (2013). *Obesity and overweight.* Factsheet N°311. Geneva, Switzerland: Author. Retrieved from http://www.who.int/mediacentre/factsheets/fs311/en/.