Volume 2, Issue 3, September 2024

Journal of Design Service and Social Innovation

Article

Research Article

JDSSI of M-SCHOLAR
NO.00020 (A-2024)

Thinking Medium: A Design-based Critique of Nudge Theory

Lawrence Marcelle , Matthew Robb *

Parsons School of Design, New York City, USA

Abstract: Mainstream economic thought rests on a picture of human actors as rational, calculating, and selfish. Since the 1980's, the sub-discipline of behavioral economics has challenged mainstream economic thought by revealing limitations in the ordinary person's ability to reason as orthodox theory predicts. People's choices are influenced by the manner in which options are presented. And since many of the choices people make involve basic dimensions of moral life, including the chooser's autonomy and general welfare, the structuring of choice situations is an irreducibly ethical enterprise. These topics are explored in "nudge theory," which represents an attempt to understand the implications of behavioral economics for people involved in the shaping of others' decisions. Nudge theory proposes that professional designers should structure choice situations on behalf of others in a paternalistic way: leveraging the designers' expertise and superior knowledge to promote outcomes that serve choosers' long-term interests. In this paper, the co-authors agree that paternalism can be a valid approach to ethical design practice in a limited range of cases. But in general, the program of "libertarian paternalism" advocated by nudge theorists represents a failure of imagination that arbitrarily forecloses opportunities for collective decision-making and reinforces unhealthy social divisions. Instead, a greater sensitivity to the potential of design thinking, specifically its capacity to incorporate a wide range of heterogeneous factors and perspectives, is a key to the ethical deployment of nudges. Design itself is the best answer to the question: how can economic thought register the results of the behavioral sciences in a thoughtful, creative, and defensible way? Through a series of examples drawn from the literature, the co-authors present a design-based critique of nudge theory, illustrating the interdisciplinary potential of design and aspects of the role design thinking may play in creating more harmonious and unified societies.

Keywords: Behavioral economics; Behavioral science; Nudge theory; Design ethics

https:// doi.org/10.59528/ms.jdssi2024.0906a20

UPDATED

Received: August 12, 2024
Accepted: August 20, 2024
Published: September 6, 2024

Principal Author(s):
Lawrence Marcelle, PhD, Part time assistant professor.
Matthew Robb, PhD, assistant professor of design and sociology.

Correspondence email: *
robbm@newschool.edu

Publication Statement:
Journal of Design Service and Social Innovation focuses on design research and cultural dissemination, but does not involve any political views or cultural biases.

Assistant Editor: Devyn Zhao

The Behavioral Critique of Economics

Conventional economic thought has long rested on a picture of the human person as independent, selfish, and rational, and on an axiomatic understanding of rationality generally referred to as "rational choice theory." On this picture, human life is a ceaseless quest for personal advantage, defined as the result of an ongoing cost-benefit calculus aimed at the maximization of individual ("marginal") utility. Individual human actors are understood to maintain a stable list or "schedule" of desires or preferences, to have a better understanding of what they want than other entities such as government agencies, and to be motivated to satisfy as many of their preferences as possible. The individualistic character of orthodox economic theory has a methodological function, in that the theory operates on the assumption that individuals are the real entities, whose actions explain the macro-level behavior of groups, communities and entire societies. These assumptions of

mainstream economic thought – rational choice theory corresponding to an axiomatic account of ideal reason, methodological individualism, and the utilitarian definition of advantage or benefit in terms of preference satisfaction, are summarized in the figure of "economic man," or *homo economicus* [1]. This account of human nature has yielded an overarching commitment to efficiency as the ultimate goal of economic life and as the measure of public policies bearing on people's lives. The methodological implications of these assumptions are so robust that, for much of the past century, mainstream economic thought has tended to minimize or even ignore extended discussion of actual people's concrete behavior, leading to the development of economic theory as a highly technical discipline largely divorced from the norms of empirical accountability characteristic of the natural sciences.

Since the 1980's, an insurgent tendency within economics has sought to reposition the discipline on a more solid empirical basis, specifically by incorporating findings from a body of research on the psychology of choice and decision-making. Although still quite controversial especially among orthodox economists, this subdiscipline, now generally referred to as behavioral economics, has significantly and permanently impacted the field – one sign of this impact being the awarding of the Nobel Prize to one of its leading advocates, Richard Thaler, in 2017. What makes behavioral economics controversial is that it disrupts or undermines some of the fundamental assumptions of conventional economic thought – in particular, the assumption that individual human actors typically behave in a rational manner, where once again, "rational" means calculative and self-serving. Drawing on the aforementioned research on the psychology of decision-making, behavioral economics argues that we cannot, in fact, trust reason to consistently direct our choices toward our preferred outcomes. Instead, a variety of distortions, errors and biases leads us to make decisions that seem right at the time, but which can lead over the long term to disappointment and the frustration of our long-term goals. For a simple example, consider the role of temptation in a typical everyday experience: a person may resolve to maintain a healthy diet, but let that person see or smell fresh-baked cookies coming out of an oven, and their discipline may well be sacrificed on the altar of impulse.

Giving in to momentary temptation in defiance of our settled preferences and commitments suggests an image of the human individual as divided against itself, as if there are actually two personalities or voices within each of us, giving divergent suggestions about what to do in a given situation. The familiar experience of having conflicting commitments or motives has been represented in many different ways, ranging from the popular image of a person with a devil on one shoulder and an angel on the other, to the sophisticated distinction developed in Freudian psychology between conscious and unconscious psychic processes. In the past several decades, behavioral scientists such as Daniel Kahneman have attempted to bring more empirical or descriptive rigor to the account of the divided self, while avoiding strong physiological, moral, or metaphysical claims. Kahneman's version of this account makes use of the neutral vocabulary of "System 1" and "System 2." In this usage, System 1 refers to mental processes that are immediate, automatic, unbidden, and experienced as effortless. The dieter who unthinkingly grabs a fresh-baked cookie and eats it is exhibiting System 1 processes at work [2]. System 2, by contrast, refers to mental processes experienced

as slow, deliberate, effortful, linear, and logical. System 2 is the locus of self-consciousness, the "I," or ego in Freudian terminology, and is the center of executive function and self-control [3].

In Kahneman's scientific writings, the conceptual framework of System 1 and System 2 serves as an important tool for clarifying the kinds of deviation from individuals' stated preferences that behavioral economists have pointed to in their critique of mainstream economic thought. His work has thus formed an important recent contribution to a genre of psychological research that has been deployed in criticisms of conventional economics, often referred to under the rubric "dual-process theory." Kahneman signals his allegiance to the dual-process psychological model with the memorable slogan "thinking fast and slow," which he uses as a sort of nickname for the conceptual framework of System 1 and System 2.

But regardless of the precise vocabulary used, various advocates of dual-process approaches to human cognition have advanced a broadly dualistic account of human mental life as a way of explaining the troubling ways that temptation and other psychological and cognitive foibles can interfere with our sincerely held commitments. Like the more literary, folklore-based image of the little devil and angel seeking to influence the individual's choices, the problematic of System 1/System 2 suggests a kind of "psychodrama with two characters" [4] unfolding in the experience of individual agents, as they confront and resolve the dozens of choices large and small that each of us must make in even the most uneventful day. Just as the characters in any drama can come into conflict, so it is with System 1 and System 2. The dieter, in a calm moment relying on System 2 reflection, resolves to cut down on snacks. But at the moment of temptation, triggered by the tantalizing sight or aroma of the freshly baked cookies, System 1 intercedes, in effect convincing the individual that having a cookie "just this once" is, as it were, compatible with the broader commitment to eating fewer cookies. Although we typically identify with our System 2, the locus of executive functioning, it is as if our System 1 can step in and take over our self-understanding or self-consciousness, if only for a few seconds. But that's all the time needed to cause an alteration in the behavior we had previously committed to.

Kahneman introduces the concept of "cognitive illusion" to further clarify the results of the psychodrama between System 1 and System 2, and the ways that conflicts between these two sources of belief and action can undermine our discipline or – to allude to other biases and distortions widely discussed in the literature – distort our estimation of risk or chances of success. He develops the concept of cognitive illusions on analogy with more familiar optical illusions. As everyone knows, a pencil in a glass half-filled with water appears bent. If we remove the pencil from the glass, and especially if we line up the pencil against a straightedge ruler, we can verify that, in fact, it is not bent – even though, upon returning it to the glass, it once again appears to be bent. In Kahneman's account, common experiences like this illustrate "the autonomy of System 1, as well as the distinction between impressions and beliefs."[5] They also reveal the capacity of rationality to correct mistaken sense impressions, which are evidently subject to certain distortions, including that partially submerged pencils (or sticks, rods, etc.) appear bent. Behavioral economics, then, rests on a similar but deeper and more troubling claim: *rationality itself is subject to systematic distortions or illusions*; in these, System 1 processes interfere with or supervene on System

2 processes. Classic examples of cognitive illusions, explored by Kahneman and his collaborators and followers, include predictable distortions in the perceived likelihood of an event caused by irrelevant but salient factors.

Many people believe crime rates to be much higher than they are as a result of the sensational treatment of specific crimes in the news media. Many people consider shark attacks to be a greater danger than excessive exposure to sunlight, although as a statistical matter, the reverse is true. Purchases of flood insurance policies spike after a major flood, only to dwindle away as people's memory of the flood fades – a pattern at variance with general weather patterns, according to which the risk of a serious flood in a given area remains more or less stable year-to-year. And so on. Examples like these, in which proximity in space or time, the affective valence, or the elevated salience caused by sensational media accounts alter widespread perceptions of risk or likelihood, generally involve a reliance on sample sizes too small to support valid statistical generalizations – the sort of robust information that *homo economicus* supposedly relies on. In Kahneman's words, such distorted ascriptions of risk involve a cognitive illusion whereby "we pay more attention to the content of messages than to information about their reliability."[6] More specifically, when, for example, the terrifying prospect of a shark attack leads to an inflated estimate of the likelihood of such an attack on swimmers, "the amount of concern is not adequately sensitive to the probability of harm; you are imagining the numerator – the tragic story you saw on the news – and not thinking about the denominator."[7]

Such examples of "denominator neglect"[8] are legion in daily life, and for Kahneman they exemplify a category of cognitive illusion that is analogous to the optical variety. To use the term he used with his early research partner Amos Tversky, the many ways that affective salience can alter our judgment about the likelihood of an event's occurring reveals our reliance on "heuristics," or rules of thumb, to navigate situations calling for estimates of risk. Much of the time, such heuristics work quite well, providing accurate, or accurate enough, information to facilitate the successful realization of our goals. But heuristics can also lead us into error, as the examples of shark attacks and flood risks illustrate. And these errors – "cognitive illusions" – are more troubling than those affecting sense perception, because there is no higher intellectual authority or court of appeals to which our distorted judgments may be elevated for further review and revision. Whereas optical illusions can be corrected by an application of reason, distortions of reason itself are, it would appear, incorrigible. Responding to the obvious question whether such distortions can be corrected or overcome, Kahneman claims that "the message of these examples is not encouraging. Because System 1 operates automatically and cannot be turned off at will, errors of intuitive thought are often difficult to prevent."[9]

In a different discussion of the effortful nature of System 2 processes, Kahneman explores the errors that many people make in responding to simple quizzes and tests, claiming that these common errors reveal a certain laziness pervasive in human intellectual nature. Here too, Kahneman sees "discouraging implications for reasoning in everyday life," specifically that "when people believe a conclusion is true, they are also very likely to believe arguments that appear to support it, even when these arguments are unsound. If System 1 is involved, the conclusion comes first and the arguments follow."[10] Because there are numerous

contexts in which System 1 and System 2 processes regularly conflict, Kahneman glumly concludes that

> *biases cannot always be avoided .... Even when cues to likely errors are available, errors can be prevented only by the enhanced monitoring and effortful activity of System 2. As a way to live your life, however, continuous vigilance is not necessarily good, and it is certainly impractical. Constantly questioning our own thinking would be impossibly tedious, and System 2 is much too slow and inefficient to serve as a substitute for System 1 in making routine decisions. The best we can do is a compromise: learn to recognize situations in which mistakes are likely and try harder to avoid significant mistakes when the stakes are high.* [11]

This has more or less become the position of many leading behavioral economists – a position that therefore underwrites a distinctive pessimism about human nature and our ability to consistently execute decisions that in fact advance our settled long-term goals. Most behavioral economists today believe that, to become genuinely rational, choice-making must in many cases be assisted by means of a prior structuring apparatus that makes the "right" choice – i.e. the choice that best promotes the overall health, wealth and wisdom of the chooser – more salient, noticeable, or attractive. Behavioral economics therefore insists that the broadly libertarian commitment of orthodox economic thought – to maximize individual freedom of choice – must be supplemented by a paternalistic treatment of choice situations that increases the likelihood of rational choice-theoretical outcomes actually being realized in specific cases. Accordingly, the paradoxical-sounding policy program known as "libertarian paternalism" has become a main product of the attempt to integrate behavioral science findings into conventional economic thought. The libertarian paternalist program has been articulated and popularized under the rubric of the theory of "nudges," a topic we turn to next.

Nudge Theory: The Behavioral Critique Applied

Nudge theory is a collection of recommendations for structuring choice situations in concrete contexts in ways intended to improve people's lives. In general, nudges maintain the overarching goals of orthodox economic thought while taking into account the disruptive findings of recent behavioral research. Advocates of nudge theory, including the originators of the term "nudging," the economist Richard Thaler and his sometime co-author, legal theorist Cass Sunstein, take their point of departure from the concept of cognitive illusions stemming from conflicts between System 1 and System 2 processes, a concept they call "the key insight"[12] that behavioral scientists have introduced into the field of economic thought. Combining our susceptibility to cognitive illusions with the observation that every choice situation has some structure which tends to facilitate or prioritize certain outcomes over others, Thaler and Sunstein offer the important insight that choice situations can be designed. Moreover, they can be designed in ways that benefit various stakeholders to the

choice situation: the choosers themselves, or another party such as a commercial entity seeking to sell a product or service.

An important benefit of Thaler and Sunstein's 2008 book *Nudge* is the way it makes clear that behavioral science has important implications for the design disciplines, as it has for the economics field. To be maximally effective in their profession, while observing appropriate ethical limits, designers in all the disciplines – graphic, spatial, or experiential – need to know something about the deliverances of the behavioral science of choice-making. *Nudge* represents an early step in clarifying some of those deliverances, and their book makes clear that professional designers represent an important subclass of what they call "choice architects" – that is, people involved in shaping the choices of others. Thaler and Sunstein show how the careful design of choice situations can help to realize the goal of their general program for choice architects, "libertarian paternalism." This is the name they give to their general recommendation that behavioral experts should be involved in the design of choice situations so as to "make choosers better off, *as judged by the choosers themselves.*"[13] As an overarching policy program, in other words, libertarian paternalism seeks to "help people make the choices that they would have made if they had paid full attention and possessed complete information, unlimited cognitive ability, and complete self-control."[14] Three well-known examples drawn from the book will serve to illustrate what nudges are, the range of their application, and the varying scales and stakes of their deployment.

A simple and relatively noncontroversial example of a nudge initiative involves the painting of stripes across a road on the approach to a sharp corner. Reducing the distance between the stripes as they get closer to the corner gives rise to an illusion of increasing speed for someone approaching the corner in a car, leading the driver to put their foot on the brake and thus reducing the risk of an accident. Here, an automatic System 1 process, vision, engaging evolutionarily ancient and primitive neural pathways in the brain, is leveraged in a manner intended to overcome a common hazard of driving, lapses of attention that can lead drivers to enter sharp corners too fast. The paint stripe nudge sets up a sort of optical illusion in service of a universally shared System 2 goal: arriving at one's destination safely. The illusion exemplifies, and trades on, the same distinction between appearance and reality, between *seems* and *is*, as the submerged pencil that appears to be bent. The practical upshot of the transverse paint stripe intervention, which has been adopted in many cities and municipalities, is a reduction in traffic accidents near risky corners. Crucially, the paint stripes require no additional policy or legal apparatus to have their beneficial effect: they do not correspond to any new laws, bans or restrictions on the freedom of drivers, who may choose to override the evidence of their eyes and enter the sharp corners at high speed, without committing any infraction beyond possibly exceeding posted speed limits. By activating a kind of illusion – in this case a literal, optical kind – without preventing or punishing any action by drivers, the paint stripe intervention exemplifies a nudge, and illustrates the broader program of libertarian paternalism: it encourages a certain, undeniably beneficial, outcome, through facilitating minor adjustments in individual behavior, and without imposing new mandates, bans or costs.

Another category of nudges, also introduced in Thaler and Sunstein's book and heavily discussed in the subsequent literature, involves the layout and design of a cafeteria. This

proposal rests on an important observation, which applies to all choice situations and hence to all nudges: there is no neutral way to present a range of options, and this includes the presentation of the various foods in a cafeteria. Something will have to be in front, be the most noticeable, the first – that is, the most *salient* – thing that patrons see, while other foods will be relegated to subsequent or less obvious positions in the patrons' navigational experience of the cafeteria space. This non-neutrality of the cafeteria layout illustrates the key point that choices have a structure, an "architecture."

In this context, Thaler and Sunstein suggest that cafeterias should be designed so as to prioritize healthy foods, for example by locating a salad bar near the entrance, increasing the visibility and availability of fruits and vegetables. Beyond adjustments to a cafeteria's floor plan, other design interventions that can reliably help to promote healthier eating include reducing the size of plates as well as the serving utensils made available to the patrons [15]. Initiatives such as these have been implemented numerous times, and are reliably associated with desirable dietary behaviors such as greater consumption of fruits and vegetables and lower overall caloric intake. As regards the conceptual framework of System 1 and System 2, meanwhile, these cafeteria nudges reveal a much more complex situation than that pertaining to the road stripes example. For it is clear that people's dietary choices involve both System 1 impulse and long-term System 2 commitments. Many of us can sympathize with the cafeteria patron who has resolved for health reasons to eat a salad, but then notices the shelf full of tempting desserts and ends up putting a slice of cake on the cafeteria tray. In orthodox economic thought, this behavior might be interpreted as showing that the cafeteria patron in fact prefers, all things considered, to eat the cake [16].

Behavioral economics, by contrast, claims to show that this temporary suspension of their perceived preferences involves an "illusion," albeit an illusion of a cognitive, not purely optical, nature. The cafeteria patron is not conscious that the enhanced salience of the cake – the fact that it is made visually and emotionally more "available" than the salad by means of its prominent placement – is actually the cause of the patron's choosing the cake over the salad. (Of course, in almost all theories of rationality within economic debates, as in everyday common sense, the fact that the cake is several inches in front of the salad is not considered a good reason for preferring the cake.) In addition to its deviation from orthodox economic explanations of preferences that simply infer desires from individual actions, the behavioral-economic analysis of such common failures to live up to our stated goals also diverges from traditional philosophical explanations of such failures in terms of incontinence or weakness of the will. Here, the data is explained not simply in terms of the struggle between an impulsive desire for immediate pleasure versus a commitment to long-term health and well-being. Rather, the outcome is here shown to be determined by an altogether trivial cause (the cake's being made more salient to the patron through sheer proximity or visual prominence).

In comparison with the road stripe nudge, the "healthy cafeteria" provides a more compelling example of the way that the "psychodrama" of System 1 and System 2 processes can lead to conflict. But as with the road stripes, we see here the possibility of using design interventions to encourage choices that leverage System 1 processes in service of System 2 goals. People cannot always be trusted to behave in ways that honor their

settled goals, but they can – and should – be nudged to do so by means of careful choice architecture. In the present case, that architecture might include locating desserts at the back of the cafeteria, outside of the main customer flows, or around a corner and out of sight. Once again, design decisions such as these can influence choices, behaviors and outcomes without the cafeteria, or a governmental agency that regulates the cafeteria's operations, needing to implement bans, mandates, or even conventional incentive mechanisms like higher taxes on disapproved foods. Thus the "healthy cafeteria," too, like the road stripes, exemplifies the program of libertarian paternalism, by which individual choices can be influenced without being coerced. Yet despite the absence of overt coercion – after all, patrons can simply walk around the salad bar to the shelf of desserts, or take their small plates back for a second helping, if they choose (most do not) – efforts to design cafeterias in service of health-promoting goals have been met with alarm. Commentators claiming to represent a libertarian perspective, in particular, have observed that being influenced in what we choose to eat, even if those choices are not coerced, can feel like a kind of manipulation. In general, then, efforts to act on people's stomachs have been less well received than interventions such as the paint stripe initiative, which acts on evolutionarily ancient and primitive mechanisms of visual perception.

The third type of nudge initiative has also generated widespread controversy, even as it has been widely adopted. Like many industrialized countries, the United States has a taxpayer-funded retirement savings and social insurance system intended to help retired persons avoid sinking into poverty once they are no longer generating an income; however, this system is insufficiently generous to fully support the majority of citizens throughout their old age, especially given long-term trends in the cost of living as well as in longevity. So Americans are strongly encouraged to set up and fund their own retirement savings plans, and the government has developed various tax-advantaged investment vehicles for encouraging such savings by individuals.

Nevertheless, many people in the United States do not believe they are accumulating sufficient retirement savings [17]. To remediate this problem, Thaler and Sunstein include in *Nudge* a recommendation that employers in the United States promote retirement savings among their employees by switching from an opt-in to an opt-out mechanism for administering the retirement plans the employers are required to create and contribute to on behalf of their employees. Instead of waiting for the employee to proactively request the establishment of such a plan, stipulating a certain percentage or dollar amount of each paycheck to be diverted into the plan, they argue, employers should take that first step on behalf of their workers, diverting a small amount of the employees' starting paychecks into retirement account and investing the monies in some noncontroversial investment vehicle such as a broad-based stock fund, while also making matching contributions in accordance with the employer's policy. Switching from an opt-in to an opt-out mechanism like this overcomes – indeed, leverages – a familiar and well-established deficit of actual human choice makers: inertia. Most employees, goes the reasoning, *would* opt in to their employers retirement savings plan if they thought clearly enough about the matter; the fact that many employees do not set up their retirement savings plan upon being hired shows not that they prefer to have less money in retirement, but instead reveals a breakdown in

their self-serving rationality: a kind of blindness or cognitive illusion that undermines their long-term interests. A nudge in the form of automatic enrollment in a retirement savings plan, from which the employee can of course opt out at any time, is a way to remediate the lack of retirement savings among American workers, using the cognitive deficit of inertia for people's own good, instead of, for example, seeking to educate the workers to take greater responsibility for their own retirement.

Adverting once again to the System 1/System 2 conceptual scheme, we can see that the utilization of System 1 processes through the manipulation of defaults can serve the long-term, System 2-based, interest of the individual. But the retirement savings initiative differs from both the road stripes and the health-promoting cafeteria in the focus of its intervention. Far more than avoiding an accident and maintaining a diet, saving for one's future is a purely or almost purely System 2-based goal. Formulating and acting on that goal requires the most long-term focus and the most abstract reasoning abilities among the three nudge examples. And as mentioned previously, System 2 is the locus of the individual's ego or sense of executive decision-making. It lies at the heart of personal identity. It is therefore not surprising that the opt-out retirement savings mechanism, along with related recommendations including that future raises be automatically diverted to retirement savings [18], have been among the most controversial nudge proposals, being criticized by commentators across a wide range of political perspectives.

Some defenders of "individual liberty" (or of related concepts such as "consumer sovereignty"[19]) may find that the automatic retirement savings plans, like the health-promoting cafeteria, is unacceptably paternalistic, curtailing individual choice, initiative and responsibility even though it does not mandate or prevent any specific outcome [20]. Meanwhile, some critics on the political left have criticized the way the automatic savings proposals accept and reinforce the putative American consensus that avoiding poverty in old age is largely the responsibility of individuals, potentially distracting and sapping the will of the public to address a moral matter at the level of public policy by institutionalizing "solutions" perceived as being better than nothing and thus avoiding important questions about the status of 'citizen' in a representative polity [21]. By intervening not at the level of sense perception as with the road stripes, nor even at the level of appetite like the cafeteria, but in the domain of reason itself, the proposals for nudging people into saving more for their own retirements penetrate to the heart of uniquely human behavior, namely, reasoned action in service of long-term goals. This is a main reason why these proposals have been subject to the most serious criticisms and across a wider range of political orientations than the other examples noted above.

The Self-Misunderstanding of Behavioral Economics

In its careful cataloging of the vagaries of choice-making by actual human choosers – in particular the many ways that choices can be altered through the role of heuristics, emotions, and other sources of cognitive illusions – behavioral science threatens some of the

fundamental assumptions of orthodox economic thought. In particular, the behavioral turn in economics challenges the picture of human nature summarized in rational choice theory: the account of human beings as selfish, utility-maximizing, calculative creatures able to identify and maintain a ranked schedule of individual preferences, and to strive to realize those preferences over time. The depth of the challenge is revealed in the amount of controversy behavioral approaches have generated in the economics field, with many representatives of the orthodox view continuing to refute, qualify, or simply ignore the empirical research in favor of continued work of a more strictly "theoretical" and nonempirical nature [22]. But despite the profoundly disruptive bearing of behavioral economics on conventional economic theory, many leading proponents of behavioral economics, and of its practical application in the strategic development of nudges, retain a strong allegiance to the overall conception of "rational choice" contained in the orthodox account. And this allegiance leads to serious tensions within the behavioral approach to the design of choice situations. Evidence of these tensions can be seen in the role that the concept of cognitive illusion plays in key contributions to the behavioral science literature, including Kahneman's. A brief return to the example of the submerged stick will help clarify this point.

In an optical illusion we distinguish between the perceived content (that the stick looks bent) and the judgment (that the stick is in fact bent). Our judgment may be corrected through learning (experience or the study of optics). As we have seen, Kahneman, like other advocates of dual-process psychological models, attempts to draw an analogy between optical illusions and cognitive illusions. However, whereas optical illusions can be corrected through an appeal to reason – that is, to System 2 processes – there is no further higher court of appeals, no "System 3" or super-rational level of cognition in which cognitive illusions might be put on trial and corrected in the way that optical illusions can be corrected by System 2-level review of System 1 processes.

Kahneman wishes to make an analogy between optical and cognitive illusions, but there is no analogous mechanism in his account to the cognitive division of labor between, on one hand, a perception of a state of affairs, and on the other, a final judgment as to the accuracy of that perception, rendered through an agreed-upon system of measurement. In its attempt to develop a theory of systematic human irrationality, behavioral science still relies on System 2 as the arbiter of rationality itself. In other words, System 2 alone is in a position to claim that we are justified, based on the evidence, in believing that we are prone to systematic failures in judgments. Since for Kahneman the source of cognitive illusion is System 1 intervening in System 2 processes, System 2 is assigned a double role in his empirical psychology: *both* of being led astray by cognitive illusions, *and* of being able to recognize that it is being led astray and thus to take some kind of corrective action. System 2, in other words, is expected to serve as a judge in its own case. This double status of System 2 creates a deep ambiguity in the critique that behavioral economics advances against the axiomatic account of rationality on which orthodox economics depends. The problem is not primarily that the behavioral science literature appears to solicit a sophisticated philosophical conceptualization of reason that is nowhere provided (certainly not in Kahneman's work). Rather, the main point is that behavioral economics undermines

all appeals to ideal reason – that is, to any conception of reason free from the distorting effects of System 1 – *even as* it continues to hold up such an ideal as the horizon against which our judgments are to be tested. More specifically, behavioral economics continues to affirm the norms of reasoning stemming from the theory of rational choice as the measure of rationality as such, even after it has shown, through its cataloging of cognitive biases, that rational choice theory is an empirically unfounded account of the reasoning process.

The tension in the behavioral response to standard economic theory is also evident in the practical application of behavioral economics, nudge theory, which, as noted previously, seeks "to help people make the choices that they would have made if they had paid full attention and possessed complete information, unlimited cognitive ability, and complete self-control."[23] Here too, in other words, the advocates of nudge theory respond to the empirical undermining of a traditional understanding of the role of reason in human behavior with advice on how that traditional understanding of reason might, in effect, be rehabilitated in specific cases. It is almost as if, having delivered their comprehensive and devastating critique of rational choice theory, showing its utter inadequacy as a model of human thought and action, the authors of the critique experience a sense of remorse or of longing for the lost model – and proceed to offer a series of workaround ideas for intervening in people's lives in ways that would vindicate the rational choice model of human thinking after all, by steering people into behavioral patterns they would already exhibit if rational choice theory were true.

Dewey's Pragmatism: From Rehabilitation to Reconstruction

In the view of its leading theorists including Daniel Kahneman, the behavioral sciences eventuate in a deep pessimism about human beings' ability to make choices rooted in reason and thereby tending toward their long-term self-interest. In turn, the pessimism of the behavioral scientists underwrites the development of the libertarian paternalist program, mainly conceived as the development of a series of nudge initiatives. Ordinary human actors, unable to reliably advance their own interests through rational choice-making, must be guided in their decisions through repeated encounters with choice situations intentionally designed to encourage appropriate, advantageous outcomes – precisely the outcomes that people would always choose if they were thinking and acting "rationally." The behavioral science of choice first undermines the foundations of orthodox economic thought, methodological individualism and rational choice theory, then proceeds to attempt a rehabilitation or restoration of these foundations, effectively making them "true" again through the paternalistic intervention of choice architects.

The arbitrariness of this effort to rehabilitate rational choice theory has already been highlighted. But if not such a program of rehabilitation, then what *is* the appropriate response to the behavioral-scientific findings regarding our (unquestionable) susceptibility to cognitive bias and distortion? Here the insight of nudge theorists, that choices are never neutral or purely open-ended, but always delimited and structured, yields an important clue.

In the majority of situations in which members of modern societies must make a choice, those choices have been, or could be, intentionally designed. Practice, action, and intervention impinge on and condition the mental tasks of observing and thinking, picking and choosing. This means that desires are never developed, let alone acted upon, in isolation. Individual agents with their beliefs and desires do not just find themselves in a choice situation experienced as an empty stage for the enactment of their pre-existing plans and dreams. Instead, plans are circumscribed and choices restricted, whether through natural forces or human-made interventions. Choice situations, in other words, are never neutral, and individual freedom is never the "negative liberty" discussed by Western political philosophers and understood as the absence of all constraints or barriers to action. The structured nature of freedom suggests that there is something deeply wrong about the methodological individualism at the heart of the account of human nature summarized in the figure of *Homo economicus*. Individual human agents, with their beliefs and desires, are not after all the basic building blocks of complex social institutions and structures.

A second important clue pointing the way toward a more adequate understanding of human cognition and action in the wake of the recent behavioral research lies in the empirical analysis of choice-making in terms of the dual-process account of human psychology. In this account, as we have seen, System 1 and System 2 are presented as independent and autonomous systems of cognition, with System 1 functioning as an analog for desires that can act as direct prompts for action and System 2 representing our capacity to apply logical structures in service of appropriate means-end reasoning. In the current context it is crucial to notice that the dual-process explanation of human cognition, at least in Kahneman's version, *also assigns to System 1 a role in belief and goal formation*. System 1 has, if not conscious intentional states, at least proto-conscious states. For it is responsive to certain cognitive demands associated with being able to understand language. Kahneman himself calls attention to this fact in his list of automatic activities that he ascribes to System 1, which among other items includes the following achievements:

- Complete the phrase "bread and …"
- Answer to 2 + 2 =
- Read words on large billboards
- Understand simple sentences
- Recognize that a "meek and tidy soul with a passion for detail" represents an occupational stereotype [24].

Although Kahneman refers to System 1 as an "associative machine,"[25] it is clear from these examples that System 1 can also process language at least at a rudimentary level. But to credit System 1 with "understanding a simple sentence," even as an associative and thus not a fully conceptual operation, only makes sense insofar as we already know how to understand complex sentences (for example, the sentence "System 1 understands simple sentences"). But if "understanding" has the same meaning whether it describes an ability of System 1 or System 2, then we are really crediting them with sharing some overlapping cognitive function or domain – and the two systems cannot be truly independent and

autonomous after all. The dual-process account does offer an intuitively plausible way of understanding a range of typical errors of cognition in terms of conflicts between the two systems. But the ascription of even elementary proto-conceptual processing abilities to System 1 hints at a problem with the binary and polarized framing of the dual-process theory as presented by Kahneman.

A third clue suggesting a way out of the impasse created by the behavioral scientists' pessimism about human rationality and the consequent program of libertarian paternalism is indicated by Kahneman's repeated acknowledgements that "most of the time,"[26] System 1 and System 2 do not conflict, but yield their respective contributions to belief and action in ways that harmonize, producing coherent, reasonable and functional behaviors. This acknowledgement appears to qualify or call into question the *practical* bearing of the System 1/System 2 distinction, much as the acknowledgement of System 1's proto-linguistic capacities calls into question the *conceptual* distinction between these psychological layers or "processes." The essential point is this: if System 1 and System 2 mostly get along, working in harmony to advance people's understanding and effective navigation of the world, then it would appear that the pessimism of the behavioral scientists is largely unwarranted and the program of libertarian paternalism via nudges mostly unnecessary. On the other hand, if System 1 and System 2 are in fact independent and autonomous systems, each impacting human belief and action, then the existence of any harmonious or integrated behavior at all becomes a mysterious anomaly and deeply puzzling coincidence. As with the recognition of System 1's quasi-conceptual capacity, the upshot of attending to the acknowledgement that the two systems frequently work in concert is to question the depth and meaning of the dualistic way that these systems are presented by Kahneman and other dual-process theorists. These considerations reinforce the outcome of the previous section of this paper, that there is something mistaken about attempting to retain the individualism and the axiomatic account of reason typical of the orthodox tradition in economics in the wake of the behavioral critique of that tradition. The findings of that critique point in a different direction than the critique's authors themselves recommend. And these three clues suggest that the first step in that new direction must involve moving definitively away from the dualistic tendencies of the tradition.

One thinker whose work represents a sustained effort to overcome unhelpful conceptual dualisms of all kinds is John Dewey, a major exponent of philosophical pragmatism. Dewey's discussion of theory and practice is representative of his general approach, and relevant for our purposes in this paper. In a manner that empirical psychologists like Kahneman would readily agree with, Dewey asserts that science is a preeminent intellectual achievement of contemporary civilization. But the scientific enterprise, in Dewey's view, suffers from a widespread misunderstanding about its own status, a misunderstanding that reflects the influence of the broader metaphysical tradition in Western philosophy. Since ancient times, Western thought has rested on a dualistic framing of basic ideas. Dewey claims that Greek thinkers in particular contributed to that tradition

> *the idea of a higher realm of fixed reality of which alone true science is possible and of an inferior world of changing things with which experience and practical matters*

> *are concerned. They glorified the invariant at the expense of change, it being evident that all practical activity falls within the realm of change. It bequeathed the notion, which has ruled philosophy ever since the time of the Greeks, that the office of knowledge is to uncover the antecedently real, rather than, as is the case with our practical judgments, to gain the kind of understanding which is necessary to deal with problems as they arise.* [27]

Both ontologically and epistemologically, Western thought has been predicated on an invidious distinction between true being, admitting of no change and hence the object of certain knowledge, and the empirical world, full of contingency, change and decay, and of which true knowledge is not possible, but only mere "opinion." Dewey's critique of the distinction between being and becoming, and of its many correlates including theory and practice, mind and body, and ends and means, centers on the actual success of the enterprise of science, especially over the past four centuries. If not *the* most, science surely constitutes one of the most successful domains of human intellectual endeavor during that recent period. And what accounts for that success, in Dewey's view, is science's transformation from a broadly observational to an experimental enterprise – that is, its transformation from a discipline of passive *looking at* nature to one of active *intervening* focused on the prediction and control of natural processes. As he puts the point, "science in becoming experimental has itself become a mode of directed practical doing,"[28] a "doing that manipulates"[29] natural forces and events in service of particular ends (or as he calls them, "ends-in-view"[30]). The rapid progress of modern science is a powerful illustration of the priority of practical reasoning over abstract theorizing. But despite its legacy of practical success, Dewey claims, science remains for the most part trapped within a metaphysical self-understanding, as an enterprise that seeks to accumulate facts about the "antecedently real." This self-misunderstanding of science has a range of negative consequences, including the fact that it impedes further social progress that might otherwise issue from the application of scientific (that is, experimental) procedures to other domains of human concern and activity.

Seen against this broad philosophical backdrop, we can understand that the self-misunderstanding of the behavioral scientists, in claiming to identify disappointing gaps or hiatuses in ordinary people's reasoning abilities and in calling for the rehabilitation of ideal reason through the careful design of people's choices, appears as a specific instance of the more general misunderstanding of science as such pointed out by Dewey. What is needed, in both the general and specific contexts, is a transition from an understanding of science as an accumulation of facts that enables a subsequent formation of individual desires and corresponding practical actions aimed at the realization of those desires, to an understanding of science as itself an inherently practical enterprise, essentially interventionist, and seeking not an accumulation of facts per se, but the more effective navigation of experience. With respect to the problematic exposed by the behavioral turn in economic thought, what is needed is not the *rehabilitation* of reason as an instrument for individual preference satisfaction based on true beliefs about an objective world, but rather the *reconstruction* of reason, now framed as "the kind of understanding which is necessary to deal with problems as they arise." Dewey's name for a rationality liberated from the task bequeathed to it by

Western metaphysics – the task of uncovering antecedent reality – is "intelligence." He means by this term *an adaptive capacity for organizing and directing behavior within an environment or context*. A hallmark feature of intelligence, as Dewey uses this term, is its resolutely non-dualistic character. In Dewey's pragmatist view, reasoning is neither caught up in, nor renders its judgments in terms of, the familiar dualities of mind and body, self and other, or means and ends. Dewey uses the word "transaction" to clarify what he means by intelligence, drawing attention to the connections that unite an organism and its environment, as opposed to highlighting the gaps between them. Together, these terms reinforce a sense of rationality that correlates this term with the kind of homeostasis or dynamic equilibrium characterizing an ecosystem, in which organisms and environment interact to yield a functioning, thriving whole.

The conceptual capacities we associate with rationality get fetishized in rational choice theory and misunderstood in the main currents of behavioral economics. They do not, and never did, allow us to transcend our creaturely status in apprehending timeless truths. Thus, "failures to be rational" are not well understood in terms of cognitive illusions, but as stemming from a misunderstanding of reason as primarily an instrument for having true beliefs. These points are summarized in Dewey's claim that today,

> *the work and office of philosophy rests upon a romantic exaggeration of what can be accomplished by "intelligence." [This] ... word names something very different from what is regarded as the highest organ or "faculty" for laying hold of ultimate truths. It is a shorthand designation for great and ever-growing methods of observation, experiment and reflective reasoning which have in a very short time revolutionized the physical and, to a considerable degree, the physiological conditions of life, but which have not as yet been worked out for application to what is itself distinctively and basically human.* [31]

Dewey's account of rationality helps to illuminate problems in the common roots of orthodox economic theory and of behavioral economics, namely methodological individualism and rational choice theory, and in the dual-process psychology that informs the behavioral-economic critique of the orthodox theory. In the Deweyan view, "reason" is best understood not in terms of the satisfaction of individual preferences based on true beliefs, but as an adaptive capacity for organizing and directing behavior. In the next section, we argue that, seen through the lens of the pragmatist understanding of rationality as adaptation, the findings of the behavioral turn in economic thought point toward design as a fruitful way forward in understanding and incorporating behavioral-scientific findings regarding human psychology.

Design as Intelligence

In its rejection of dualistic thinking of all kinds, philosophical pragmatism provides intellectual grounding for a common criticism of nudges: that they focus attention on minor details or aspects of problematic institutional arrangements, when a more comprehensive

analysis of the issues raised by the nudge is what is really needed [32]. Indeed, in its rejection of dualism and reconstructive treatment of rationality as adaptive intelligence, the pragmatist philosophy of Dewey suggests that even the simplest nudges, like the road stripe intervention, reside at a nexus of heterogeneous forces and factors, ranging from the physical laws governing the motion of objects through space, to social expectations around orderly behavior frequently enforced by laws, to cultural norms of many modern societies prioritizing efficiency in reaching one's destinations quickly and safely. The cafeteria design and retirement savings plan examples, meanwhile, show how much more complex things can get when factors such as individuals' emotional relationship to food or the abstract demands of planning for one's future become part of the picture. We will return to a discussion of our three examples in the final section of the paper. For now, it is enough to recognize that, although advocates of nudge theory such as Thaler and Sunstein stress the simplicity or elegance of many nudge solutions, a closer look at many of their examples reveals just the opposite: a tangle of interconnected habits, customs, norms, expectations and traditions, many of which are simply bypassed or occluded by the initiative being promoted.

The complexity revealed in even relatively straightforward nudges aligns with the idea, now widely shared in the design fields, that design problems are systemic in nature. To use language that has become familiar in design contexts, many if not most of the problems that occupy working designers are "wicked problems."[33] They exhibit not only great complexity but a fundamental indeterminacy, such that the very description of the problem at hand is likely to be controversial and subject to political disagreements.

For design theorist Horst Rittel, who coined the expression "wicked problems," for complex systemic challenges there is often no such thing as *the* problem that all parties will recognize as such. Rather, "the formulation of a wicked problem *is* the problem!"[34] For example, it is not universally agreed that promoting fruits and vegetables or reduced caloric intake should after all be the basic principles governing the design of cafeterias, or how exactly to relate these goals to others including the maximization of revenues, profits, or reported "customer satisfaction," the reduction in user wait times and other ease-of-use factors, and so on. Given their indeterminacy, the wicked problems comprising the majority of design challenges today are often described in ways reflecting ideological commitments or the "world view" (*Weltanschauung*) of the person doing the describing. Is the failure of many workers in the United States to save for retirement evidence of a fundamental laziness or inertia in people, for example, or does it reflect broader social failures of American society with respect to the issue of aging and retirement? Again, wicked problems are understood to have no definitive conclusion or "stopping rule": specific solutions cannot be understood as the "best" or "worst," but only as better or worse than other concrete alternatives. "The good" reduces in practice to the "good enough, at least for now."

But in the context of complex, politically fraught problems, this aspect of progress in design, which the political scientist Herbert Simon referred to as "satisficing,"[35] should make apparently simple, quick-fix solutions such as switching from an opt-in to an opt-out default a source of suspicion, or at least ongoing reflection, not of uncritical celebration followed by a rapid forgetting of the original issue. In other words, the fact that a solution may be

preferable, advantageous, or better in some measurable way compared with what went before, should not become an excuse to stop seeking additional incremental improvements: satisficing is best understood as part of the structure of ongoing progress, not as a name for the discovery of a once-and-for-all solution. Seen against the backdrop of the literature on wicked problems, behavioral economics and nudge theory tend to bypass or suppress the complexity revealed in all but the simplest design problems, attempting instead to impose a "linear" and straightforward solution onto the problem situation. Indeed, we can say that nudge theory represents a classic example of trying to "tame" the wickedness of complex problems, offering simplistic and reductive "solutions" whose plausibility rests on a series of implicit assumptions that mostly ignore the intractability of the problem situation.

What would an approach to complex problems which seeks to register, not suppress, that complexity look like? The initially surprising answer offered by design theorist Richard Buchanan is that the design process must move in a direction informed by *rhetoric*. That is to say, design must become more self-aware of its status as a mode of argumentation. A design solution is a kind of proposal, an invitation to address our needs in a particular manner. And design is uniquely well situated to confront concrete problems – "problems as they arise" – in non-reductive ways that that avoid the pitfalls of dualistic thinking highlighted by Dewey. In his classic essay "Wicked Problems in Design Thinking," which itself acknowledges the relevance of Dewey's philosophy, Buchanan summarizes his point about the rhetorical character of design in the following way:

> *The power of design as deliberation and argument lies in overcoming the limitations of mere verbal or symbolic argument – the separation of words and things, or theory and practice that remains a source of disruption and confusion in contemporary culture. Argument in design thinking moves toward the concrete interplay and interconnection of signs, things, actions, and thoughts. Every designer's sketch, blueprint, flowchart, graph, three-dimensional model, or other product proposal is an example of such augmentation.* [36]

Design practice today contains great potential for addressing complex problems in a way that does not deny that complexity and lapse into a feckless solutionism. For Buchanan, the reason involves the development of design over the past century, from a series of discrete trades rooted in handcraft traditions, to an expansive disciplinary domain that initiates its own research – an expansion that leads him to refer to design as "a new liberal art of technological culture."[37] The liberal ideal in education is an understanding of the world that *integrates* diverse and heterogeneous sources of knowledge to provide a comprehensive grasp. Such an integrated view is especially needed today, given the advanced degree of fragmentation in knowledge born of increasing rates of specialization in the natural and human sciences. With its nonlinear, opportunistic approach to problem solving, design is uniquely well suited to addressing the wicked problems that confront modern societies. By contrast, and despite its adjacency to the design fields as indicated in its constant references to "choice architecture," nudge theory, and the larger program of libertarian paternalism, ignores the argumentative or interrogative character of design, declining to engage the user or subject in favor of a much narrower practice of "noncoercive" manipulation. Far from

inviting a conversation, nudges seek to foreclose it, instead hurtling toward a goal whose ascription to the subject (the nudgee) is only presumed. By definition, nudges assign all the key decisions to a small group of experts to be made on behalf of the rest of us, and executed along a narrow range of possible outcomes.

Buchanan's emphasis on the rhetorical dimension of design – on design as a mode of argument and communication – points beyond nudge theory, which seeks to bypass reflection, conversation, an participation and to tame the complexity of the problems it addresses. But in Buchanan's programmatic essay, the discussion of the rhetorical character of design remains highly compressed and underdeveloped. In this connection, the sociologist Jenny Davis adds detail, nuance, and analytical power to the rhetorical understanding of design. In a critical assessment of the concept of "design affordances" popularized by the industrial psychologist Donald Norman [38], Davis makes points also relevant to the current discussion about Buchanan. She rejects the stark binary fashion in which the concept of design affordance is typically interpreted, noting that "objects afford in varying degrees, and their effects are exerted with differing levels of force. … Affordances are not present *or* absent but present *and* absent, by degree."[39] For Davis, this fact suggests that the real question is not: "*What* does this artifact afford?" but rather: "*How* does this artifact afford?"[40] Adding texture to Buchanan's elliptical allusion to design as a mode of argumentation, Davis's proposed analytical categories, or "mechanisms" of affordance, include *demand* and *request*; *encourage*, *discourage*, and *refuse*; as well as *allow*. Although not presented as an exhaustive list, these terms already suggest, more directly than Buchanan does in his classic essay, how design, and designers, can be understood as being *in communication* with users. In effect, a designed artifact (or interface, service, environment, etc.) is experienced as a series of suggestions: "do this, don't do that"; "this is what we recommend"; "wouldn't you like to try that?" Davis illustrates her expansive analysis of design affordances with a discussion of swipe-based dating apps. Such apps, she points out, "*request* that users consider a high volume of potential partners and *discourage* users from slow considerations."[41] Thus far, her discussion is close to how a nudge theorist might describe the app interface: a dating app user wants to find an appealing date, and the app greatly facilities the work of finding one. But Davis's analysis goes much further:

> *The swipe feature may then shape how individuals evaluate potential partners and how they present themselves as romantically appealing – placing emphasis on quickly identifiable markers such as physical attractiveness and income. The glut of potential partners and ease of selection and dismissal may shape how those who use the apps interact during dates, perhaps moving quickly to intimacy to establish commitment within a crowded pool or keeping distance to avoid foreclosing the full range of romantic options. These micro interactions can affect romance and intimacy at a cultural-structural level by normalizing serial dating, detaching a single date from future romantic engagements, and empowering those who feel dissatisfied in current relationships to explore the abundant field. In short, swipe apps don't just offer another way to date but reshape the meaning and practice of finding love.* [42]

Again, although her own goal in the passage involves extending and updating Donald Norman's account of design affordances, Davis's analysis makes clear that the current design of most dating apps is experienced as a series of nudges. The app emphasizes the "availability" of potential partners in an obvious way, and plays on consumers' aversion, demonstrated in the research, to being presented with single or limited options [43]. At the same time, Davis's example hints at what a deeper understanding of design as a mode of argumentation might yield by way of the further development of such apps.

For the purposes of this paper, the key point is that Davis points both to the world-shaping power of design as a mode of argument, and to the possibilities for users to participate in that shaping work. Regarding the app example, we might point out that finding love, and in general navigating the conditions of intimacy in conjunction with the other demands of contemporary life, is one of many wicked problems confronting people in modern societies. For users of dating apps to understand themselves as being in a conversation, in communication, with the designed artifact (the app) and thus with its designers, invites reflection on what it might mean for the users to "talk back" to the design or to the designers, directly adjusting or else offering feedback and suggestions about the design of the interface in ways that could alter the rhetoric of its engagement with users. There is of course no guarantee that a more participatory app design process would lead in the end to better romantic relationships. But such an approach to design would be both more respectful and more empowering to the users than the reductive approach of nudge theory, which seeks to leverage our cognitive deficits toward a desired end, even if we have expressed that desire ourselves. The wicked problems understanding of design and its potential, supporting a rhetorical understanding of what design is and how it is experienced, holds far more promise than nudge theory for the development of artifacts and experiences truly worthy of our admiration and investment.

Rooted in the pessimism about human educability bequeathed to it from the behavioral sciences, nudge theory trades on a sense that nudges are the best we can hope for – and that guiding people to make decisions in their own best interest, without literally mandating or forcing those decisions, is the best way to promote general well-being while maintaining the basic freedoms that citizens of modern liberal societies expect. To behavioral psychologists like Kahneman, any alternative to the program of libertarian paternalism will appear to rest on a misunderstanding of their research findings, or to traffic in in pie-in-the-sky fantasies about the redemption of human limitations. To them, improving human decision-making in some other way looks impossible. But from design theorist Buchanan's point of view,

> *what many people call "impossible" may actually only be a limitation of imagination that can be overcome by better design thinking. This is not thinking directed toward a technological "quick fix" in hardware but toward new integrations of signs, things, actions, and environments that address the concrete needs and values of human beings in diverse circumstances.* [44]

What Buchanan in this passage refers to as "design thinking" is a contemporary instance, indeed a paradigm, of what Dewey means by "intelligence." At least in a broad construal of the term, design, as a preeminent manifestation of practical reason, is a main locus of intelligent conduct in the world today. As such, design, understood as a widely shared "liberal art of technological culture," represents the most promising way forward: it is the name for the best available way to "address problems as they arise" in the wake of the behavioral research on human decision-making.

Thinking Medium

John Dewey's reconstruction of reason as adaptive intelligence assigns a central role to the concept of transaction, the intersection of organism and environment that functions as the locus of adaptation (that is, intelligence) or its opposite, maladaptation. The contemporary psychologist Gerd Gigerenzer, a prominent critic of the libertarian paternalist program, invokes essentially the same conception when he speaks of "ecological rationality."[45] For both thinkers, intelligence is a type of behavior stemming from the transactions of organisms in specific environments. To refer to rationality as adaptive or ecological means that the discrete terms "organism" and "environment" are themselves analytical products or outcomes of reflection on the contexts in which intelligence or its opposite – adaptation or maladaptation – emerge.

Although there is nothing wrong with treating organisms and their environments as separate foci of investigation, it is misleading to conceive of these terms of analysis as pointing to items that preexist their surfacing as terms of analysis. The temptation to do so reflects the ongoing influence of Western metaphysical philosophy, with its separation of being and becoming and its conception of knowledge as grasp of "the antecedently real." But as we have seen in connection with Kahneman's dual-process problematic of System 1 and System 2, the dualistic thinking that results from succumbing to the temptation of metaphysics creates insuperable conceptual problems. Once System 1 and System 2 are treated as independent and autonomous layers of a person's psychology, the possibility that those layers might work in harmony becomes a deep mystery. Similarly, once "organism" and "environment" are treated conceptually as independent entities, the mutually constitutive shaping and conditioning of each by the other becomes impossible to grasp. Beyond purely natural environments, the shaping of individuals by social and historical circumstances, as well as the possibility of cooperative endeavors manifesting social intelligence, also become incomprehensible mysteries if "organism" and "environment" are understood as ontologically independent entities. These are some of the pernicious legacies of dualism that Dewey sought to overcome through a reconstructive treatment of reason as adaptive intelligence.

On the naturalistic account developed by Dewey, human beings are to be understood as organisms, i.e. as animals that have evolved alongside all the other animal species currently populating the planet. What distinguishes us from other animals is not a transcendent power of reason which puts us in touch with permanent truths, but rather our extraordinary capacities for operating within the transactional contexts in which all creatures live out their

lives. Two capacities accounting for the extraordinariness of human beings stand out: our ability to *coordinate* and align our efforts with others of our kind – that is, our capacity to use communication to advance, modify, and abandon discrete goals; and our evident ability to *shape* the environment in which we find ourselves – that is, our capacity for design. But in both our communicative and world-shaping (designerly) capacities, intelligence or success remains adaptive in nature. In other words, neither of these aptitudes of the human being reveal human nature as in any way escaping or transcending the domain of the natural world described by the sciences. Moreover, no human intellectual or cognitive aptitude reveals human beings as individuals first and foremost. On the contrary, individuality is always a precipitate of communication, of social and cultural norms that shape and delimit the historical horizon of possibility for individual human organisms. The designed world, too, in both its physical infrastructure and in abstract creations like company policies and political constitutions, is logically and temporally prior to the individual human agent who confronts, not an open-ended world of unlimited possibility and negative liberty, but a discrete set of options and parameters that condition what the individual will recognize as possible and desirable. The functional unity of individuals is internally connected to group endeavors, manifest in the communication contexts and in the designed physical (and now digital) spaces in which we first encounter ourselves as individuals.

The adaptive account of rationality as intelligence developed in American pragmatism, of which design thinking is a paradigmatic instance, encourages an interpretation of the individual's success and failure in reasoning that diverges from the axiomatic account of the economists. The norms of rationality are in fact biological and social adaptations, not transcendent standards that apply in all times and places. Irrationality – or what Kahneman calls "cognitive illusion" – is on this view the product of a certain maladaptation in the transaction encompassing organism and environment. But maladaptations can occur in different ways and for different reasons. It's necessary to examine, in a comprehensive and holistic manner, the various factors – biological, socio-cultural, historical, and so on – comprising a certain problematic situation before any idea of how to resolve the problem can be responsibly formulated. Above all, on the adaptive understanding of rationality, it is not possible to ascribe problems of human social or institutional life to the failure of individuals to meet an arbitrary standard of rationality or behavior – least of all when such standards are obviated by the very nature of the critique in question, as in the case of the behavioral critique of rational choice theory. However, this is precisely what behavioral economics does. And its rush to judgment, identifying the errant individual as the main source of the problems challenging modern societies, amounts to a sophisticated and insidious form of victim-blaming. While obesity, for example, is a real problem in many modern societies, the behavioral-economic way of framing that problem is as an individual failing: Overweight people eat too much, and need to be induced to eat less – and since we can't trust anyone to do that on their own, given the susceptibility to temptation that we all share, we should help them meet the goal we have identified for them by tricking them through design, including the design of cafeterias.

But if design is not to be understood simply as a set of instrumentalities for nudging people into behaving more "rationally," how *do* we understand design and the relevance of design

thinking to addressing genuine human needs, including in the light of the biases, distortions and cognitive heuristics that sometimes lead us astray? The answer comes from the literature on design in relation to wicked problems. As we have seen, most design problems, properly understood, have wicked characteristics: that is, they are systemic, complex, indeterminate, political, and iterative, having no criteria for identifying a discrete end or stopping point. This means that the great majority of design problems are not problems of optimization. There is no "best" solution or outcome that can be sketched or even inferred from the nexus of conflicting factors and forces making up the contexts for design intervention. Thus, the failure of individuals to respond to such contexts in the anticipated or normatively recommended way cannot straightaway be ascribed to a failure or breakdown in the cognitive apparatus of the individual in question. In particular, the identification of "fast" and "slow" cognitive processes, as valuable as it is, can only be understood as one more input for the designer to consider in attempting to realize a better resolution.

Indeed, from a design perspective, breakdowns in the transactions between individuals and designed environments call for a much broader and more holistic analysis than the one offered by behavioral economists and their intellectual allies in the psychology field. Informed by the broader understanding of adaptive intelligence bequeathed by Deweyan pragmatism, a designerly approach to a problematic situation will start by stepping back, considering larger timescales and the historical development of the problem. It will encourage the adoption of various perspectives vis-à-vis the problematic situation – perspectives ranging from those of concrete stakeholders to more abstract positions such as "the average user" or "the reasonable person," which may represent political ideals in a contemporary society. Design thinking in relation to a problem is sensitive to cultural norms and traditions and the ways that these may bear on individuals' thoughts, desires, and behavior. Crucially, design thinking seeks not a definitive "solution" that solves the problem once and for all, but a satisficing approach that on one hand seeks an arrangement that is "good enough for now," but at the same time recognizes that no concrete solution is perfectly or permanently satisfactory, that there is always and in principle more to be done, further improvements to be made. In sum, the norms of judgment that design accommodates, in thinking about a problematic situation and developing recommendations for improvement, are far broader than those which economists have tended to prioritize, even after the intervention of behavioral sciences in the field.

Modern societies confront their members with a range of characteristic problems. For example, many people get into car accidents, overeat, and enter retirement without sufficient resources to secure a comfortable old age. A design approach to these problems, attuned to their great complexity and backed by a reconstructive understanding of rationality as intelligence, sees them not a priori as failures of the individual stemming from cognitive illusions, but as instances of maladaptation – and promotes investigation of the broadest possible range of factors that might help explain the lack of fit between organism and environment, and what changes to either might help to ameliorate the problem. Instead of capitulation to pessimism, what the problems of modern societies actually call for is communication. Various constituencies with a stake in the current situation should be

involved, whether literally or in principle, in a wide-ranging discussion of the situation in all its complexity. Communication encompasses the task of giving and asking for reasons, naming and describing problems from a range of perspectives, educating one constituency about the challenges faced by others. This task is essentially social and participatory in nature, and thus it is fundamentally opposed to the kind of expert-driven paternalism represented by nudges. Finally, a design approach focused on complex social problems does not seek a definitive "solution," but rather an incremental amelioration or satisficing intervention that makes things better than they were.

But what does it mean to make things better, after we have abandoned the criteria of optimization stemming from the belief in ideal reason and the axiomatic account of individual utility maximization? How can we make sense of "better," having forsaken the conception of "best" bequeathed by the tradition which we must now abandon? To address this final question, let's recall the important "clues" that guided us toward a re-examination of John Dewey's non-dualistic philosophy as a better way to interpret the real upshot of behavioral science's critique of orthodox economic theory. In that context, we referred to the fact that choices do not take place on an empty stage or neutral field, but only emerge in discrete, and for the most part designed, contexts; to Kahneman's examples of automatic System 1 processes, which nevertheless exhibit proto-linguistic or proto-conceptual abilities; and finally, to Kahneman's acknowledgment that "most of the time" System 1 and System 2 do not in fact conflict, but work in harmony, each playing a role in advancing individuals' aims and meeting their needs.

Again, these clues point toward a non-dualistic account of the human organism-in-an-environment, summarized by Dewey under the rubric of "transaction." And as we have argued, they reveal the characteristically opportunistic, nonlinear, and comprehensive sensibility exemplified by designers, and recommended for broader adoption under the rubric of design thinking focused on complex wicked problems, as a paradigm of rationality reconstructed as intelligence. In light of the behavioral-scientific critique of conventional economic thought, and in light of the pragmatist and design-based critique of the self-misunderstanding of the behavioral sciences, we can now see that the question to be asked in light of the characteristic problems of modern living is not: "How can we realign our actions with the recommendations of mainstream economists?" But rather, "How can we restore a state of healthy equilibrium or homeostasis, with all facets of the human agent – both System 1 and System 2 included – contributing in appropriate ways to the successful navigation of the individual organism's experience? That is, we must no longer ask: how can individuals maximize their utility? and must now substitute the question: how can we more fully manifest intelligence in the current problematic situation? "Intelligence" is another name for the kind of situation – the kind that obtains "most of the time" – in which organism and environment are in a state of healthy homeostasis or dynamic equilibrium.

Notwithstanding Kahneman's caveat that the terms System 1 and System 2 are merely offered as "useful fictions,"[46] we have shown that his own use of these terms remains mired in ontologically freighted dualistic thinking, which confuses his understanding about what his research has actually shown. Above all, we have exposed the arbitrariness of his pessimism about human nature, rooted in his belief that System 1 and System 2 represent

two independent and "autonomous" domains in the human cognitive apparatus, which can never be fully and definitively integrated. In the face of the resulting "psychodrama" between the two domains of cognition, between our disparate capacities for "thinking fast and slow," Kahneman lays the groundwork for the program of libertarian paternalism via nudges, understood as the only viable way to remediate a human nature divided against itself. Against this, the philosophical and design-theoretical resources discussed in this paper show that we are not in fact divided against ourselves in the permanent and ontological manner described by Kahneman, his collaborators and his followers. *Most of the time*, the human organism-in-its-environment is able to proceed on its way, advancing goals and meeting needs by drawing on abstract or empirical System 2 knowledge and by applying System 1-level automatic processes as well as ingrained skills, in a mixed and organic way dictated by the immediate needs of the moment. Moving through experience, in other words, is a matter of thinking fast and slow all at once, in complex and shifting arrangements that adapt as conditions change and new circumstances emerge. And when our projects do get frustrated, our choices stymied, and we are spurred to more deliberative or theoretical reflection, the end goal is to return once again to the kind of integrated homeostasis we were in before our projects were interrupted.

Our name for the state of integration, in which all of us live when we don't have a problem and to which we all return once we have resolved "problems as they arise," is *thinking medium*. This phrase is a way of acknowledging the power of what behavioral scientists including Kahneman have discovered in their research, while rejecting the dualism implicit in their own framing of that discovery. "Thinking medium" is both a *description* of the psychology of ordinary experience, and a *prescription*, as the name of the psychological status quo which we seek to reestablish through the satisfactory resolution of a problem. It is both a fact and a value. And it is a better way to understand the bearing of the behavioral critique of orthodox economics, and the role that design must play in articulating the next stage of human social and cultural evolution, than the authors of the behavioral critique themselves have understood.

The Three Nudge Examples Revisited

A brief review of our three nudge examples will help to concretize the points we have made. As we saw, nudges involving stripes painted onto roadways utilize optical illusions in service of a desirable result, a reduction in automobile accidents. This class of traffic management devices, widely deployed around the world, is among the simplest, most obvious, and least controversial of design interventions classifiable as nudges. In the language developed by Kahneman and others, the road stripe intervention appears as a purely System 1 affair – that is, in trading on a literal optical illusion, it seeks to elicit the type of instinctive or instantaneous behavioral response that Kahneman thematizes under the rubric of "thinking fast." The nudge acts directly on the driver's eyes and visual processing mechanism – that is, on the body – and bypasses the higher-order reasoning functions we typically associate with the "mind." Given its physiological focus, it is not surprising that the road stripe nudge has generated the least concern or controversy among

our three examples: this intervention does not threaten anything perceived to be essential to people's core identities, while it also plays an obviously helpful role in improving traffic safety [47].

Yet even here, with this simplest and least controversial type of nudge, it is important to note that the context in which such a nudge could be implemented – a sharp corner on a road – implicates far more than the visual perception of the driver considered in isolation. Indeed, what is at stake in the road stripe nudge includes a whole series of institutions in which people's getting quickly from point A to point B is something that needs to be accomplished on a regular and widespread basis. Transportation whether for intermittent travel or for daily commutes to and from work; human communities dense enough to warrant the construction of roads yet which have not created mass transit options functional or attractive enough to create a viable alternative to driving – an entire form of life is implicated in the question what is to be done about a dangerous corner on a certain stretch of road. As design theorist Ariel Guersenzvaig notes in a related context, questions like whether the road stripe nudge is an appropriate design intervention or not "cannot be approached from the perspective of mobility alone. To reach an improved state of affairs, one needs to include other important issues besides mobility: dwelling, commerce, work, leisure, safety, environment, and others perhaps."[48] Thus, even this example, simple as it is, reveals the arbitrariness and insufficiency of the individualistic approach to design problems encouraged by economic thought. Without places to go and people to see, so to speak, the vast majority of drivers will not be on that road in the first place, and the nudge seeking to slow them down ahead of a sharp corner may not even be worth the cost of the paint.

Our other nudge examples illuminate other aspects of dualistic thinking transcended by pragmatism's emphasis on practical reason and reconstruction of rationality as intelligence. For example, the question how to best lay out a cafeteria immediately reveals the extent to which food-related norms are deeply social norms. That is, questions of what we should eat, what we should want to eat, what we should encourage others to eat, and so on, are inextricably bound up with widely shared, widely circulated cultural norms around health and longevity, beauty and attractiveness, and the desirability of eating foods understood to be "natural" as opposed to being products of the industrial systems that provide most of our food today. Individuals' relationship to food, therefore, is never simply a matter of "individual consumer choice," but is heavily overdetermined, incorporating layers of memory, family, economics, and corresponding norms that condition "what we should eat." Eating is at once physiological *and* cultural: it activates System 1 and System 2 at the same time, yielding a far more complex "transaction" than in the case of a car approaching a sharp corner. The psychological, social, and cultural complexity around eating, hence the question of how best to design a cafeteria, thus reveals the unreality of the atomistic individualism, and the commonsense division of self and other, private and public, that underpins both orthodox economic thought and its attempted rehabilitation by behavioral economists and nudge theorists. The assumptions about the individual cafeteria patron on which a proposed cafeteria design might be offered as "best" or "optimal" are bound to be

partial and arbitrary, as will larger policy decisions or recommendations pertaining to food production and consumption based on the conceptual duality of individual and society.

A design thinking approach to the design of a cafeteria offers a different perspective, treating human beings as biologically evolved animals that over time have developed a wider environment for directed activity that includes the power to name and describe, to communicate and discuss goals (ends-in-view), and to design physical and institutional infrastructures in accordance with our agreed-upon goals and aspirations. In such a context, an "ecological" rationality recognizes that a large factor in problems connected to overeating, such as obesity and the health challenges it brings, is that there are structures in the modern environment that aim to profit from people's poor nutritional choices.

The design of food-related nudges in (e.g.) the cafeteria setting involves focusing on a problem that has been created in the wider social ecosystem that includes a food industry which provides foods, including unhealthy foods, on a mass scale and often at competitive price points. A designer of nudges in a space like a cafeteria needs to maintain an awareness of the difference between individual failures (e.g., wanting to lose weight but nevertheless buying the cake) and social pathologies (e.g., companies resolving to market appealing but unhealthy food items because doing so is profitable). Designers (choice architects) who are engaged in laying out a cafeteria need to be aware of that wider social ecosystem, and the way that factors like the commercial incentives behind the consumption of unhealthy amounts of junk food qualify and contextualize the analysis of what is happening, at least much of the time, when cafeteria patrons make a food choices that deviate from their own stated commitments. Recognizing that nudge theory "focuses the blame for societal problems exclusively on the individual mind, closing our eyes to institutions that steer individual behavior so that they can take advantage of it, and … misleadingly suggests that a more sustainable solution, educating people, is a hopeless endeavor,"[49] designers of a cafeteria should aim to make patrons aware of the reasons behind the specific design decisions, promoting transparency.

Even better, the designers should strive to engage the cafeteria's patrons in a conversation about what their eating goals actually are, what a healthy or appropriate relationship to food for them would actually entail, and how the physical infrastructure of the cafeteria and its design might best serve those goals. It is not to be assumed that unanimous agreement about so complex a matter as food and eating would be forthcoming. Accordingly, a well-designed cafeteria might incorporate multiple stations or pathways, corresponding to different sets of preferences: light snacks over here; healthier options incorporating fruits and vegetables over here; fuller meals over there, and so on. A more respectful, less paternalistically designed cafeteria might still incorporate some of the conventional instrumentalities yielded by the behavioral research, such as smaller plates or serving utensils – but these would not be simply perpetrated on the patrons as it were behind their back: instead, design and sourcing decisions around things like plates and utensils could be communicated by means of physical signage, QR codes, and other easily accessible sources of information. Finally, a well-designed cafeteria would operate in a state of continual review, generating feedback both from patrons and through observation, in service of revised iterations within a satisficing logic of continual improvement. In this connection, a

modular approach to cafeteria layout could be one way to support such a logic of continual, iterative improvement. Approaching the task of cafeteria design and management through initiatives like these would not definitively "solve" the problems associated with healthy eating habits. But such an approach would represent a more sensitive, creative and potentially more successful response to the behavioral research about human decision-making. Initiatives like these could help people to be more thoughtful about their relationship to food and eating, and could help them to more effectively maintain their goals around eating over the long term. By contrast, implementing a series of nudges represents a quick-fix "solution" that leaves untouched the largest drivers of unhealthy eating habits, including the commercial imperatives that generally characterize modern industrial economies.

Lastly, as we have seen, the retirement savings plan nudge involves promoting individual employees' long-term financial security through a simple adjustment in the design of the websites by which employers' human resources departments help their employees to initiate and maintain retirement savings plans. Whereas the paint stripe nudge appears to intervene purely on the level of System 1, retirement savings plans in general operate entirely at the level of System 2. Involving long time frames and the highly abstract instrumentality of money, the determination to save for one's old age is an almost entirely intellectual achievement, involving a cool assessment of costs and benefits, and at least a partial grasp of sophisticated concepts such as opportunity cost, the time value of money, and the dynamics of compound interest. In comparison with the road stripes, and even more than the cafeteria, employee retirement savings plans are an almost entirely cultural (i.e. System 2-level) phenomenon. Here too, however, the simplistic way that the suggested nudge (shifting from an opt-in to an opt-out system) is represented, and the reductive account of human nature that underpins the recommendation to implement these systems, reveals the legacy of dualistic thinking under which nudge theory labors. In this case, the retirement savings nudge trades on a rigid conceptual separation of means and ends. Specifically, it rests on a highly conventional and commonsensical end, namely that people desire to avoid poverty late in their lives; and a corresponding means, namely that people can avoid an impoverished old age by starting to save now.

But as unproblematic as the conceptualization may seem, and as prudent as it no doubt is for individuals to save some of their current income if possible, we may well ask whether the goal of financial comfort exhausts the range of ends that people have in regard to their old age – or that they might have if a broader conversation were facilitated about such an important matter. In such a broader conversation, other nonpecuniary goals – living with dignity; being able to cultivate new interests, activities and skills; experiencing honor and the recognition of one's peers; enjoying a sense of fulfillment after a lifetime of contribution – might come to the fore alongside the material minimum of avoiding poverty. Of course, a broader conversation about retirement would almost certainly include pecuniary matters as well, such as whether the relative distribution of responsibility for retirees' financial well-being – among the retirees themselves, institutions including corporate employers, and governments – is warranted, appropriate, and just. It's impossible to know in advance what values or commitments might be surfaced through such a conversation, or, therefore, what

modalities or instrumentalities for promoting such values might come to be sanctioned at the level of the community or society. Even less is it possible to predict how a community's or society's goals for retirement might themselves be modified through the discussion about what it might take to realize those goals, including discussion of the costs, unintended consequences, or painful trade-offs that pursuing such goals might incur. From a Deweyan perspective, the complex and multilayered nature of a large topic such as how a society gives meaning to the process of aging and of retirement from the workforce, to what extent it honors or takes care of its elders, and what role personal responsibility ought to play in securing a happy old age, show the arbitrariness of belief in the commonsense distinction between means and ends. In reality these interpenetrate, as we negotiate, adjust, realize, and take on new plans, on timescales ranging from seconds to decades and beyond. Once again, our goals or "ends" are really ends-in-view, never finally fixed but shifting and conditioned by the available means.

What might a design-led approach to setting up retirement savings plans look like? In this case, in comparison with the other two examples of nudging we have discussed, Gigerenzer's call to educate the user would seem to have its most direct application. Although everyone knows that having more money is better than having less, comparatively few grasp what is at stake in the failure to save for one's own retirement. Even less are most people encouraged to reflect on the broader issues at stake – for example, how in the United States, the responsibility for avoiding an impoverished old age is substantially assigned to individuals themselves, and how this arrangement ultimately reflects a decision that American citizens have effectively made.

For the majority of people in the US, this status of individual responsibility for retirement being a decision or choice is not something that comes readily to mind. Although the question of whether and how much to save is highly abstract and thus activates System 2 processes, the idea that saving is primarily an individual responsibility is, in fact, something more like a heuristic for people in modern societies, especially the US, in which for the most part people just assume that "this is what is done." In other words, whereas the substance of the questions around retirement activate System 2, the assumption that individuals must bear the burden of retirement operates at the level of System 1.

From a Deweyan pragmatist perspective, the pervasiveness of this assumption reveals a failure of our society to promote a deeper reflection on this important phase of life. The real reason that we save for ourselves is that we've accepted this responsibility, so that other entities, such as corporations or governments, do not have to shoulder more of the burden. Recognizing the larger landscape of issues in the questions around retirement, then, a design approach to this wicked problem of modern life would reject the simplistic and reductive "solution" of adjusting the default options on companies' human resources websites as anything close to an adequate response. Switching employees to an opt-out savings regime is unacceptably paternalistic, and moreover represents a quick, passive and rather mindless acceptance of individual responsibility and an endorsement of the status quo. As such, the opt-out system represents a failure of design to live up to its own potential. Again, here above all, education – or a "boost" to people's ability to grasp the broader social and political stakes involved in retirement planning, is called for [50]. A web-based

wizard or flowchart to help employees set up their savings plan would ideally invite a deeper and more mature discussion among people, in their status as employees but also as citizens, regarding why exactly we do things in the way that we do and whether a different way would be better, what the trade-offs would be, what unintended consequences might ensue, and so on. Companies, not naturally set up to advance serious societal discussions especially of a political nature, might be enjoined through a government mandate to promote this kind of deeper conversation through the development of a generic web-based functionality that all would be required to use in helping their employees sign up for retirement savings plans. At a minimum, companies might make available access to other resources that lay out the stakes of different possible approaches to retirement, without claiming to endorse particular perspectives.

A more sophisticated understanding of the bearing of the behavioral turn on economic life, rooted in pragmatism's non-dualistic reconstruction of reason as adaptive intelligence, would make possible a much broader, deeper, more respectful and more creative discussion about a host of difficult challenges facing modern societies such as those we have discussed here. And those discussions could be structured, advanced, facilitated and iteratively resolved through an application of design thinking, suited as it is to grappling with very complex, wicked problems. Design practices, aligned with the theoretical orientation of pragmatism, can establish a better, more coherent response to the findings of the behavioral sciences than the scientists' own accounts do. Daily life indeed thrusts us into a series of challenging choice situations, but for the most part the response that is called for is not libertarian paternalism, i.e. nudges designed by experts and perpetrated on passive and unsuspecting subjects, but rather a type of participatory design.

Conclusion

Recall Kahneman's glum advice noted earlier in this paper: that "biases cannot always be avoided [and therefore] the best we can do is a compromise: learn to recognize situations in which mistakes are likely and try harder to avoid significant mistakes when the stakes are high."[51] This quote is a manifestation of the pessimism that the behavioral science of choice has led its leading advocates to adopt based on their uncritical importation of assumptions about human nature from mainstream economics and the broader philosophical inheritance of Western thought.

This paper has shown the arbitrariness of those assumptions, and sought to show that a more adequate interpretation of that critique, drawing on philosophical themes from American pragmatism, points to design as a deeply constructive approach to the complex problems typical of modern life. Design is well positioned to grapple with complexity without sliding into reductive and unimaginative responses, of which nudge theory and the broader program of libertarian paternalism are prime examples. Our ultimate recommendation, to see "thinking medium" as both a starting point and end point for human agents, pursuing goals and confronting "problems as they arise," represents a critique of the behavioral scientific literature and of its practical application, nudging, in all but the simplest cases. From the standpoint of thinking medium, "compromise" does

not look like such a bad thing after all. That word well describes what often happens as people navigate challenges, communicate their needs and learn about others' needs, and take in a wide range of considerations in developing – designing – satisfactory-for-now resolutions. The rest of Kahneman's counsel also looks to us like wisdom: learning to recognize situations in which mistakes are likely and trying to avoid them sounds like another way of saying that we should facilitate communication and, in the broadest sense, education as a way of helping people to navigate the challenges that occur in every life. A sufficiently broad commitment to education, encompassing both factual *knowledge-that* and, following Dewey and Buchanan, a designerly practical *know-how*, would be the best way to take up the results of the very powerful behavioral critique of economic thought, and to respond to that critique in a creative way that opens new possibilities. The cognitive biases and distortions identified by Kahneman and his colleagues represent an important breakthrough in the study of human beings. But the grim interpretation of those findings that has generally accompanied their presentation is not warranted, and the design disciplines can help us achieve a more appropriate, positive, and imaginative way of responding to them.

Acknowledgements:
The co-authors would like to thank Anis Amir for his support with a literature review and research in connection with this project. A version of this article was presented as a keynote address at the 2024 International Forum on Design Service and Social Innovation; the co-authors gratefully acknowledge the opportunity to share their work in that context.

Author Contributions:
The co-authors collaborated very closely on all aspects of producing this article.

Funding:
Not applicable.

Conflicts of Interest:
The authors declare that they have no conflicts of interest related to this research.

References and Notes

1 The term first emerged in the 19th century as part of a critical reaction to John Stuart Mill's utilitarian account of human nature, which had "raised the indignation of Victorian moralists shocked at the postulation of such blatant selfishness." Joseph Persky, "The Ethology of Homo Economicus," *Journal of Economic Perspectives* 9, no. 2 (1995): 222. [CrossRef]

2 Importantly, skilled behavior born of long and disciplined practice – the surgeon confidently executing a delicate medical procedure; the chess master who accurately assesses the state of play without needing to calculate various possible moves one by one; the trained athlete whose grace and power makes her physical performance "look easy" – also relies on and exemplifies System 1 psychological processes.

3 Other terms for articulating these two layers or aspects of human psychology have included "automatic" and "reflective" or "involuntary" and "voluntary" processes.

4 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 21.

5 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 26. [CrossRef] The author's own example in this passage involves a different optical distortion, the famous Müller-Lyer illusion, in which two lines of the same length appear to be different lengths because of the fins or vanes, pointing in different directions, that extend from the ends of the lines.

6 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 118.

7 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 144.

8 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 329.

9 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 28.

10 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 45.

11 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 28.

12 Richard H. Thaler and Cass R. Sunstein, *Nudge: The Final Edition*, (Yale University Press, 2021 [2008]), 25.

13 Richard H. Thaler and Cass R. Sunstein, *Nudge: The Final Edition*, (Yale University Press, 2021 [2008]), 7.

14 Richard H. Thaler and Cass R. Sunstein, *Nudge: The Final Edition*, (Yale University Press, 2021 [2008]), 7.

15 Paul Rozin, Sydney Scott, Megan Dingley et al., Nudge to Nobesity I: Minor Changes in Accessibility Decrease Food Intake," *Judgment and Decision Making* 6, no. 4 (June 2011): 323–332. [CrossRef]

16 Extreme versions of this idea have sought to bring even obviously self-destructive behaviors, such as those involving drug addictions, under the rubric of personal preference and hence of choice. See Gary S. Becker, *The Economic Approach to Human Behavior* (Chicago: University of Chicago Press, 1976).

17 Richard H. Thaler and Cass R. Sunstein, *Nudge: The Final Edition*, (Yale University Press, 2021 [2008]), 183.

18 Richard H. Thaler and Shlomo Benartzi, "Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy* 112, no. S1 (2004): S164-187. [CrossRef]

19 W. H. Hutt, "The Concept of Consumers' Sovereignty," *The Economic Journal* 50, no. 197, 1940: 66–77. C

20 Robert Sugden, "Why Incoherent Preferences Do Not Justify Paternalism," *Constitutional Political Economy* 19 (2008): 226–248. [CrossRef]

21 Daniel M. Hausman and Brynn Welch, "Debate: To Nudge or Not to Nudge" *The Journal of Political Philosophy* 18, no. 1 (2010): 123–136. [CrossRef]

22 For a sustained critique of behavioral economics from the standpoint of mainstream economic theory, see David K. Levine, *Is Behavioral Economics Doomed? The Ordinary versus the Extraordinary* (Cambridge, UK: Open Book Publishers, 2012). [CrossRef]

23 Richard H. Thaler and Cass R. Sunstein, *Nudge: The Final Edition*, (Yale University Press, 2021 [2008]), 7.

24 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 21.

25 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 50–58.

26 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), e.g. 24, 25.

27 John Dewey, *The Quest for Certainty*, in *John Dewey, The Later Works, 1925-1953, Vol 4: 1929* (Carbondale: Southern Illinois University Press, 1981), 14.

28 John Dewey, *The Quest for Certainty*, in *John Dewey, The Later Works, 1925-1953, Vol 4: 1929* (Carbondale: Southern Illinois University Press, 1981), 20.

29 John Dewey, *The Quest for Certainty*, in *John Dewey, The Later Works, 1925-1953, Vol 4: 1929* (Carbondale: Southern Illinois University Press, 1981), 5.

30 This expression reflects Dewey's non-dualistic treatment of the relation of means and ends, showing that any end is an end only provisionally, and only in relation to practically realizable means. See Brendan Hogan and Lawrence

Marcelle, "The Complementarity of Means and Ends: Putnam, Pragmatism, and the Critique of Economic Rationality," *Graduate Faculty Philosophy Journal* 38, no. 2 (2017): 401–428. [CrossRef]

31 John Dewey, *Reconstruction in Philosophy*, in John Dewey, The Middle Works, 1899-1924, Vol 12: 1920 (Carbondale: Southern Illinois University Press, 1977), 258.

32 Nick Chater and George Loewenstein, "The I-Frame and the S-Frame: How Focusing on the Individual-Level Solutions has Led Behavioral Public Policy Astray," *Behavioral and Brain Sciences* (2022) [CrossRef]; Ruth Schmidt, "A Model for Choice Infrastructure: Looking Beyond Choice Architecture in Behavioral Public Policy," *Behavioural Public Policy* 8, no. 3 (2024): 415–440 [CrossRef]; Tim Harford, "What Nudge Theory Got Wrong," Financial Times, May 6, 2022, https://www.ft.com/content/a23e808b-e293-4cc0-b077-9168cff135e4.

33 Richard Buchanan, "Wicked Problems in Design Thinking," *Design Issues* 8, no. 2 (Spring, 1992): 5–21. [CrossRef]

34 Horst W. J. Rittel and Melvin M. Webber, "Dilemmas in a General Theory of Planning," *Policy Sciences* 4, (1973): 161. [CrossRef]

35 Herbert A. Simon, "Rational Choice and the Structure of the Environment," *Psychological Review* 63, no. 2 (1956): 129–138. [CrossRef]

36 Richard Buchanan, "Wicked Problems in Design Thinking," *Design Issues* 8, no. 2 (Spring, 1992): 5–21, 20. [CrossRef] In this paper we generally retain the expression "design thinking," while acknowledging the many criticisms directed against it in the recent literature. Although the term has been associated with precisely the sort of formulaic reduction and simplistic "solutionism" that we here accuse nudge theory of perpetuating, Buchanan's own use of "design thinking" clearly points toward the more open-ended, participatory role that he believes design can and should play in helping to address complex contemporary problems. We hope to retain the spirit of Buchanan's own use of this now-controversial expression in this paper.

37 Richard Buchanan, "Wicked Problems in Design Thinking," *Design Issues* 8, no. 2 (Spring, 1992): 5–21, 5. [CrossRef]

38 Donald Norman, *The Design of Everyday Things* (New York: Basic Books, 2013 [1988]).

39 Jenny L. Davis, *How Artifacts Afford: The Power and Politics of Everyday Things* (Cambridge: MIT Press, 2020), 64. [CrossRef]

40 Jenny L. Davis, *How Artifacts Afford: The Power and Politics of Everyday Things* (Cambridge: MIT Press, 2020), 65. [CrossRef]

41 Jenny L. Davis, *How Artifacts Afford: The Power and Politics of Everyday Things* (Cambridge: MIT Press, 2020), 84. [CrossRef]

42 Jenny L. Davis, *How Artifacts Afford: The Power and Politics of Everyday Things* (Cambridge: MIT Press, 2020), 84. [CrossRef]

43 Daniel Mochon, "Single-Option Aversion," *Journal of Consumer Research* 40, 2013: 555–566. [CrossRef]

44 Richard Buchanan, "Wicked Problems in Design Thinking," *Design Issues* 8, no. 2 (Spring, 1992): 5–21, 21. [CrossRef]

45 Peter M. Todd and Gerd Gigerenzer, *Ecological Rationality: Intelligence in the World* (Oxford: Oxford University Press, 2012).

46 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 28–30.

47 Considered in evolutionary terms, our ability to move ourselves at the speeds reached by modern automobiles is still very new, and far outstrips the judgment of speed that our ocular systems are adapted for.

48 Ariel Guersenzvaig, *The Goods of Design: Professional Ethics for Designers* (Lanham, MD: Rowman & Littlefield,2021), 81. [CrossRef] Guersenzvaig's own example in this passage concerns the development of urban

"Superblocks" in Barcelona and other cities in Spain, and urban design intervention intended to promote livability, promote safety, increase green spaces, and reduce automobile traffic in dense urban areas.

49 Gerd Gigerenzer, "On the Supposed Evidence for Libertarian Paternalism," *Review of Philosophy and Psychology* 6 (2015): 362. [CrossRef]

50 Ralph Hertwig and Till Grüne-Yanoff, "Nudging and Boosting: Steering or Empowering Good Decisions," *Perspectives on Psychological Science* 12, no. 6 (2017): 973–986. [CrossRef]

51 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011), 28.