White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/200834/

Version: Accepted Version

Article:
Friedland, Julian, Myrseth, Kristian Ove Richter and Balkin, David B. (2023) Beyond the brave new nudge: Activating ethical reflection over behavioral reaction. Academy of Management Perspectives. 162. ISSN: 1943-4529

https://doi.org/10.5465/amp.2022.0162

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Beyond the Brave New Nudge:

Activating Ethical Reflection over Behavioral Reaction

Julian Friedland, Kristian Ove R. Myrseth, David B. Balkin

[Forthcoming in *Academy of Management Perspectives.*]

Abstract

Behavioral intervention techniques leveraging reactive responses have gained popularity as tools for promoting ethical behavior. Choice architects, for example, design and present default opt-out options to nudge individuals into accepting preselected choices deemed beneficial to both the decision-maker and society. Such interventions can also employ mild financial incentives or affective triggers including joy, fear, empathy, social pressure, and reputational rewards. We argue, however, that ethical competence is achieved via reflection, and that heavy reliance on reactive behavioral interventions can undermine the development of ethical competence over the long term. Specifically, drawbacks may occur through motivational displacement, dependency, moral crowding out, loss of personal autonomy, and reactance. We introduce complementary cognitive boosting techniques designed to stimulate reflective cognition, as a more promising long-term strategy for instilling ethical behavior. One such approach is the Moral Self-Awareness (MSA) motivational construct, which incrementally leads agents to increasing levels of ethical reflection. We explain why ethical boosting approaches present more edifying and durable alternatives to reactive behavioral interventions and offer suggestions for social and organizational policy.

Societies will continue to be controlled postnatally – by punishment, as in the past, and to an ever-increasing extent by the more effective methods of reward and scientific manipulation.

- Aldous Huxley, *Brave New World Revisited*

Behavioral intervention techniques leveraging mild financial incentives or affective triggers – such as joy, fear, empathy, social pressure, and reputational rewards – are fast becoming an organizational and marketing panacea for promoting ethical and socially-responsible behavior. Such approaches have been shown to augment charitable giving (Shearman & Yoo, 2005), reduce smoking (Noar et al., 2016), increase environmental conservation (Goldstein, Cialdini, & Griskevicius, 2008), promote inclusiveness (Nielsen & Capinski, 2013), and lower health insurance premiums (Miller, 2020). However, the underlying conceptual framing relies on the managerial assumption that a morally correct way of doing things is evident, and behavioral tactics are then used to stimulate reactive cognitive responses to induce people to do what they should arguably choose to do upon ethical reflection.

There is little doubt that behavioral interventions can induce good behavior by engaging many who tend not to otherwise or by preventing predictable mistakes (Zhang, Gino, & Bazerman, 2014). However, they can also incur unintended consequences and potentially distract us from the grander goal of building human competence. When implemented in the form of mild extrinsic incentives, they can become their own targets, thereby supplanting and attenuating the ethical motives they are intended to encourage (Bowles, 2017; Sandel, 2012; Aldred, 2009). They may also lead people to support well-meaning fads that might turn out to be ill-considered in retrospect (Gibson & Tesone, 2001), may distort and instrumentalize reasons for acting (Goodhart, 1975) which can reduce authenticity and self-determination (Ryan & Ryan, 2019), and may cause those perceiving interventions as coercive to recoil from preselected goals which they may not agree with philosophically (Nilsson et al., 2020).

Ultimately, such techniques can undermine healthy habits of critical reflection that are essential to living an authentically virtuous life.

A particularly popular intervention is the nudge concept, advanced by Thaler and Sunstein (2008), which tends to exploit reactive responses or "thinking fast", as opposed to reflective thinking or "thinking slow" (Kahneman, 2011). This approach can be particularly helpful when reflective thinking is not required to accomplish a task effectively, as in the evocative case of the etched fly on the urinal – which impels improved aiming and, thereby, reduced spillage. Such nudging triggers a perfectly innocuous, but helpful reflexive reaction. However, the choice architect may feel tempted to leverage similar reactive responses to influence *ethical* decision-making, and this is when potentially serious drawbacks may arise. This is because ethical competence – as opposed to splashless urination – is achieved via reflective thinking. There is a danger, therefore, that heavy use of nudges can undermine the development of ethical competence over the long term.

In this paper, we argue that cognitive boosting techniques (Hertwig & Grüne-Yanoff, 2017) designed to stimulate reflective – as opposed to reactive – cognition offer a more promising strategy for managers to instill ethical behavior in organizations. One way to pursue ethical boosting is by leveraging the Moral Self-Awareness (MSA) motivational construct, which incrementally leads agents to increasing levels of ethical reflection (Friedland & Cole, 2019; Friedland, 2019; Friedland & Balkin, 2023). We begin with a brief historical analysis of how the field of moral psychology gradually shifted its focus from moral development to behavioral restructuring. We then proceed to explain how heavy reliance on restructuring can compromise personal autonomy and crowd out good habits of ethical reflection through motivational displacement, dependency, and reactance. Finally, we show why boosting in general – and MSA specifically – are likely to offer a more edifying and durable alternative to behavioral structuring, and proceed to recommend social and organizational policy solutions based on ethical boosting techniques. In closing, we invite readers to consider a thought experiment regarding which society they would rather live in –

one where ethics is conditioned via behavioral reaction or one guided by moral self-reflection.

FROM MORAL DEVELOPMENT TO MORAL NUDGING

For the full span of the 20th Century, moral psychology was firmly rooted in moral development theory. The aim of the field was not only to understand moral behavior and reasoning empirically but – more importantly – to progressively realize its potential within the individual and society. It was the extension of a shared humanistic commitment to self-actualization inherited from the French Enlightenment. John Dewey, with his *Moral Principles in Education* (1909), was its earliest and greatest exponent during the last century, further buttressed by Piagetian moral stage theory (Piaget, 1932). Within this wider intellectual atmosphere, Aldous Huxley was inspired to write *Brave New World*, his prescient masterpiece of social science dystopia, published ninety-one years ago in 1932. As the novel's centenary approaches, its core depiction of a future in which individuals are behaviorally conditioned to act as the societal designers deem they should, seems thrust upon us. For much of contemporary moral psychology has turned away from the lofty aims of developing human *potential* in favor of the diminished horizons of manipulating human *behavior*. How did we get here?

High-minded moral development theory culminated in the 1970's and 80's with Lawrence Kohlberg and his student Carol Gilligan's respective multilevel approaches; the former focusing on justice-based conceptual stages and the latter on care-based awareness stages that she argued were more characteristic of female psychology (Kohlberg & Mayer, 1972; Gilligan, 1982). Moral development according to them consisted mainly in leading agents up the chain of increasingly reflective self -and societal awareness, but after dominating the field for over a generation, their approaches gradually fell out of favor. This seems to have occurred largely as a result of the declining influence of their Piagetian-inspired paradigms, particularly the insistence that stages be defined as structured wholes or that developmental

sequence be invariant, which proved difficult to sustain empirically (Lapsley & Carlo, 2014). Furthermore, numerous sobering studies on confirmation bias and attitude polarization revealed how rarely adults will change their social policy opinions in the face of disconfirming evidence (Lord, Ross, & Lepper, 1979). Another contributing factor may also have been the perceived failure of moral education in the wider culture. Kohlberg (1981) himself estimated before his untimely death that roughly a mere 20% of adults ever reached the higher stages of moral development.

Compounding this, was the ideological moral foundations paradigm established by Jonathan Haidt (2013). This approach offers a synthesis of psychology, neuroscience, and evolutionary psychology to suggest that moral and aesthetic judgments consist mainly of innate affect-laden intuitions and that moral reasoning itself is but post-hoc rationalization (Haidt, 2001; 2007). Critics argue that the account is overly relativistic, thereby neglecting wider moral progress (Gibbs, 2013), and others emphasize that its "strong grounding in nativism and focused research on adults gives little credence to developmental mechanisms and lacks details on how intuitive processes become differentiated across individuals, contexts, and time" (Lapsey & Carlo, 2014: 2). From this perspective, developing moral judgment in the wider populace via cognitive reflection seems like a quixotic enterprise – noble but misguided.

Contributing to this deflationary atmosphere was the influential cognitive dual-process construct, popularized by Kahneman (2011), which describes human decision-making as dominated by reflexive reactions or "thinking fast". Though the more reflective side of cognition does exist in this view, it is far more unwieldy. Accordingly, he and other scholars – most notably Thaler and Sunstein (2008) – advocate the use of nudging techniques triggering the reactive mind to influence actors to make the best decisions for themselves and society. So, for example, retirement savings options are couched in happiness-optimizing defaults which may also employ psychological designs such as green–yellow–red iconography to encourage certain choices over others (Bhargava et al, 2021). Similarly,

government initiatives to promote energy conservation might use happy and sad faces indicating positive or negative social comparisons (Allcott, 2011). Ultimately, instead of inspiring ethical or prosocial behavior via shared moral principles to say, lower carbon emissions or donate time, money, or possessions, targeted choices are instigated by employing the behavioral tools of financial incentives, emotional triggers, and public pressure.

This behavioral approach to engaging human capabilities and dispositions has been critiqued by a moral crowding-out literature, highlighting the extent to which capitalism has fostered an atmosphere of conditioning in which economic incentives have become ubiquitous (Sandel, 2012). In such an environment, "incentivizing" is often resorted to by default, even when seeking to inspire ethical and prosocial behavior. However, instrumental incentives have a tendency to become their own targets, thereby weakening moral motives further (Aldred, 2009). On the other hand, much of this literature also contains an implicit optimism; since these homo-economic excesses are learned, they may also be countered – if not unlearned – by corrective approaches. Indeed, a growing moral agency literature, in which this article is embedded, attempts to show a way out of this predicament by appeal to civic virtue (Bowles, 2016). Furthermore, an increasing educational emphasis on character and virtues aligns with a renewed interest in moral identity (Lapsey & Carlo, 2014; Hardy & Carlo, 2011; Aquino & Reed, 2002), developing growth mindsets (Dweck, 2006), and evolving moral self-awareness (Friedland & Cole, 2019). Cognitive boosting is also a rapidly growing field that holds promise as a positive counterinfluence to the brave new nudge (Hertwig & Grüne-Yanoff, 2017), as is ethical leadership (Brown, Treviño, & Harrison, 2005; Lawton & Páez, 2015). We provide specific social and organizational policy solutions along such lines in a later section of this paper, whereby boosting can complement nudging to lead actors to progressively higher levels of reflective moral agency. But first, we will describe in further detail how commonly employed nudging tactics can undermine personal autonomy and moral competence.

NUDGES AND THEIR EFFECTS ON PERSONAL AUTONOMY

In organizational life, individuals are often faced with myriad choices with limited time to make decisions that can affect their well-being positively or negatively depending on the choice made. As such, choice architects are specialists who simplify the process for individuals to make decisions, often by designing nudges to alter behavior in a foreseeable way without denying any options, in the spirit of "libertarian paternalism" (Thaler & Sunstein, 2008: 74). The nudged option accomplishes this by requiring the least energy, thereby coaxing individuals toward the most beneficial outcome as intended by the architect (Sunstein, 2014).

A well-known type of nudge is to construct the desired outcome as the default option. The default option is chosen and implemented when the person given a choice simply does nothing rather than opt out by deciding to do something else. By doing nothing, individuals are most likely to attain the option thought best for them, without devoting time to decide whether an alternative – the opt-out option – is better. One of the advantages of the default nudge is that the desired behavior nudged can be reversed easily, so that individuals can choose the alternative later (Hertwig & Grüne-Yanoff, 2017).

A default nudge is the least intrusive type of choice architecture to interfere with employees' *autonomy*, a condition which allows individuals to self-determine their preferred choices when presented with personal decisions. However, even the default nudge is set up by an expert who decides in advance that one choice should be easier to make than the alternative, which will require more cognitive effort or emotional energy. Social scientists who promote nudges justify favoring the default choice by arguing that many individuals prefer to focus their time and effort in other directions and appreciate choices that are simplified for them in advance (Edwards, 1983).

An example of a popular default nudge is automatic enrollment of new employees in a company retirement savings plan (*The Economist*, 2017). The alternative option requires the active decision to opt out and thereby indicate that they wish not to enroll in the savings plan.

The advantage of designing a default-option nudge to enroll employees into the plan, is that they benefit from greater savings from the favorable tax treatment given to retirement contributions, and that most plans have an arrangement where the employer matches with additional money all or a significant portion of the employee contributions. At this level, people tend not to feel that the default option infringes on their autonomy, and we would agree.

Other examples of default option nudges, however, may begin to subtly constrain autonomy. For instance, at the Department of Motor Vehicles in California, users are automatically registered to vote in future elections unless they decide to opt out (McGhee, Mehlotra, & Romero, 2021). This choice architecture benefits individuals who prior to an election may forget to register as a voter, and additionally it increases the number of voters eligible to participate in a democratic election. Similarly, Australia has made posthumous organ donation the default choice, requiring people to opt out if they prefer not to. Doing so increases participation to 90 percent from a paltry 15 percent in opt-in countries, such as the U.S. and Germany (Davidai, Gilovich, & Ross, 2012). Clearly, people may have personal or religious reasons for not wanting their bodies transferred upon death for organ extraction, and such defaults may thus be perceived by some as a disturbing form of negative social pressure (Gillman, 1999). Voter registration defaults would seem rather innocuous by comparison. However, agents may wish not to be registered to vote for personal or even religious reasons, so such nudges may also be perceived as mildly coercive.

Interventions become more invasive and problematic when implemented in the form of mild monetary incentives that guide individuals to select the option deemed the "best" by the choice architect. Many U.S. companies, for example, try to reduce health insurance benefit costs by offering modest monetary rewards to entice employees to join wellness programs that instill healthy behaviors. As such, the intervention coaxes people to choose to stay healthy while reducing company expenditures on health insurance. Furthermore, a wellness program encourages practicing healthy behaviors such as ceasing to smoke, eating nutritious

foods, taking steps to lose excess weight, and doing strenuous physical exercise on a regular basis. In addition, wellness participants must monitor and report their biometrics related to good health at regular intervals (Schaefer, 2015). Sometimes this involves wearing a biometric monitoring device 24-hours a day (Possamai et al., 2020).

A 2019 survey of large companies by Fidelity Investments reported that 56 percent of companies used a wellness program, with average insurance premium incentives of $762 provided to employees to reward them for staying healthy (Miller, 2020). The purpose, of course, is to reduce the incidence of costly medical procedures. However, employees who opted out are required to pay an average additional annual amount of $762, deducted from their pay, for their coverage. The net effect is that employees who chose not to join the program are required to pay significantly higher premiums than those who agreed to participate. In other words, the choice architecture applied to this wellness program makes it more complicated and costly to opt out of the preferred choice. Presenting an opt-in nudge for significant personal lifestyle changes via monetary incentives, together with promising unpleasant outcomes for opting out, may be perceived as invasive, paternalistic, and coercive by those holding liberty-oriented moral foundational beliefs (Nilsson et al., 2020). Thus, such nudges are likely to be seen by some as infringing on employee autonomy, potentially instigating *reactance*, namely, a recoiling from anger, irritation, or aggravation from a perceived threat to one's freedom (Steindl et al., 2015).

While these examples may begin to subtly infringe on autonomy, we do not necessarily see them as morally problematic. What we are concerned with is how such behavioral interventions can gradually become ubiquitous via a process that Frischmann and Selinger (2018) and Frishmann (2022) describe as "nudge creep". They discuss biometric monitoring devices and cellphones as examples, which nudge users by leaving geolocation-tracking on as defaults, which is far more convenient. This in turn allows various apps to collect further data and function without any involvement, activating a web of location-based functionalities and marketing interventions (Berry, 2023). The alternative of turning geolocation tracking on and

off for each app entails higher transaction costs and could lead to a debilitating overload of choices (Schwartz, 2016). Such technological tools, including generative AI, can greatly simplify – or even offload – our more mundane tasks (Miller, 2023). However, reliance on such labor-saving instruments outsources reflective "slow" thinking to reactive "fast" thinking (Kahneman, 2011). This presents a risk that many of our activities and interactions might become increasingly reactive and less reflective, leading to a host of potential cognitive and affective disengagements (Friedland, 2019). Nudge creep can be thereby viewed as an example of the *too much of a good thing* (TMGT) effect that occurs in organizations. The TMGT effect occurs when a practice known to produce positive outcomes is used beyond its inflection point, where outcomes become undesirable (Pierce & Aguinis, 2013). We discuss further examples in the following section.

MORAL NUDGING OUT

Incentivizing

The TMGT effect is particularly salient in the virtue ethics literature, in which our analysis is embedded. This stems from the fact that temperance (moderation) on this moral frame is the most valuable ethical disposition, as any virtue can become a vice in either excess or deficiency (Aristotle, 2012: 26-41). Immoral acts are thereby ultimately the result of placing excessive – or deficient – importance on certain desires, which eventually leads to moral ignorance and a distortion of values. A conspicuous example known as Goodhart's Law, is that when a measure becomes a target, it ceases to become a good measure because it instrumentalizes the good, which agents will then work around to game the system (Goodhart, 1975; Aldred, 2009: 179-220). Carbon offsets, for example, would seem to provide a convenient and efficient solution for consumers to measure – and nudge – a voluntary reduction in their greenhouse gas emissions. Unfortunately, they also offer a way out of enacting the broader behavioral changes that those costs are intended to incentivize (Skopek, 2010). Such fees can be embraced as a mere matter of course, thereby absolving the

wealthy from shouldering the same civic and moral burdens as ordinary citizens to limit waste and energy use – the same way they no longer need wait in line at the airport, amusement park, stadium, concert, or congressional hearing (Sandel, 2012: 17-42). This is what is generally referred to as the wider phenomenon of *moral crowding out*, namely, the gradual erosion of non-financial motivations toward ethical and civic behaviors via the default logic of financial incentives (Sandel, 2012; Bowles, 2016). In a widely cited study of Israeli daycare centers, for example, a marginal fine was instituted for lateness of 10 minutes or more, to incentivize timely child pickups (Gneezy & Rustichini, 2000). Characteristically, parents soon began to see the fines as fees, and late pickups nearly doubled. This is because the daycare center inadvertently instrumentalized tardiness, which no longer meant inconveniencing staff and disappointing one's children so much as paying for the privilege of being late. What's more, the increased rate of tardiness persisted even after fines were removed, suggesting that financial incentive-nudging can continue to inhibit moral mindfulness over the long term by reinforcing self-interested attitudes, which then become harder to dispel. On the other hand, if the daycares had first underscored the ethical obligations against late pickups so that the fines became signals of having flouted them, they may well have worked. Indeed, there is evidence that clear ethical framing can make mild financial nudges lead to longer term positive behavioral changes, for example in recycling and carbon emissions reduction (Bowles, 2016: 203-205; Thaler & Sunstein, 2021: 297).

Absent such framing and when taken to excess, however, the utility of wealth to absolve oneself from consideration of others can encourage callous or self-centered behavior. What is crowded out in such cases is the dimension of conscientiousness, more specifically that of moral self-awareness (Friedland & Cole, 2019), which we discuss below. Much of this motivational attitude is shaped by pride, which can often be a more effective lever – or "boost" – than resorting to financial incentives. For example, jury duty stipends are paltry compensation for the inconvenience of missing work to serve on a jury, and deliberations might thus be flawed if jurors are angered by the experience. So to help alleviate possible

resentment, direct appeals to civic mindedness are commonly employed (State of New York, 2012). A substantial empirical literature reveals the power of such appeals, which we discuss as "boosting" in the next section.

In certain cases, financial incentives can even backfire by causing individuals to recoil from what they perceive as a kind of bribery degrading their civic or moral duty. In two studies on the willingness of residents to accept toxic waste repositories in their vicinity, willingness dropped dramatically when substantial material compensation was offered (Kunreuther & Easterling, 1996; Dunlap et al., 1993). Conversely, appeals to values of civic responsibility worked far better, thereby also reinforcing habits of virtuous behavior going forward. Financial incentives may also be seen as attempts to compel those of lesser means into accepting a high-risk transaction, leading to a reactant feeling of repugnance against coercion (Leuker et al., 2020), and a large Swedish study comprising 27 million individual consumer-recycling decisions over 42 months, found a clear moral crowding-out effect where participation dropped off markedly as payments increased (Wollbrant, Knutsson, & Martinsson, 2022). Similar results have obtained regarding financial incentives for blood donation, where they either led to lower donation rates (Mellström & Johannesson, 2008) or failed to generate any overall increase (Niza, Tung, & Marteau, 2013). Intriguingly, an international study of organ donation compared countries with opt-in and opt-out deceased organ-donation defaults. While subtly nudging people to donate *posthumously* by requiring them to opt out if they chose *not* to donate increased participation rates from 15 to 90%, such countries also had 2/3 fewer *living* donors with no significant difference overall in donations by both the living and deceased (Arshad & Sharif, 2019). This suggests that posthumous opt-out defaults may induce moral crowding out effects similar to those of financial rewards by reducing the perceived civic or moral need for personal sacrifice.

Social Pressure and Affect

Interventions that leverage alternative extrinsic incentives to motivate individuals to act more responsibly have also become popular. Such approaches may leverage joy, fear, empathy, gamification, social pressure, or reputational rewards. The trouble is that heavy usage of such incentives is just as prone to the TMGT effect. This is because appeals to social pressure and other affective dispositions may – just like financial incentives – reactively shift attention toward the extrinsic reward, thereby risking motivational displacement. Moral maturity and autonomy, however, are achieved through instilling good habits aimed at intrinsic rewards. The virtue ethics literature thereby contrasts *exercising* reason to merely *following* reason by reminding us that we learn by doing (Aristotle, 2012: 1-25). In the early stages of moral development, we simply follow the example of those we trust as moral experts, namely, parents and teachers who may choose to give us mild extrinsic rewards or punishments in the process. Eventually, as those habits get instilled, we acquire the ability to exercise our own judgments in order to navigate new situational dilemmas with which we are confronted. This happens through the development of practical wisdom or *phronesis,* which is the self-actualizing harmony of mind and body governed by prudent moral judgment, with no more need of extrinsic rewards (Aristotle, 2012: 115-134). At this point, virtuous agents act deliberately, guided by critically reflective understanding of their ethical reasons for action. Conversely, in the Brave New World-scenario, there is no longer any need for developing practical wisdom. This is because behavior is perpetually conditioned, via automatic cognitive reactions, to do what is lauded by the social engineers. As a result, critical reflection becomes superfluous, and moral autonomy is virtually nonexistent (Huxley, 1958: 58).

If such interventions are employed as a panacea, they can stifle moral reflection and autonomy, which can then be more difficult to restore. As with financial incentives, reputational rewards can gradually crowd out intrinsic virtuous motivations. This might occur for example via status privileges for purchasing carbon offsets. Eventually, in the absence of any financial or reputational carrot or stick, agents may have no reason to behave morally or

civically and lack the capacity to think for themselves in new and unforeseen circumstances. Such interventions can also inspire *virtue-signaling*, a form of self-aggrandizing vanity. When taken to the extreme, this is known as "moral grandstanding" in which agents seek admiration via "an impressive commitment to justice, a highly tuned moral sensibility, or unparalleled powers of empathy. Essentially, this would turn one's contribution to public discourse into a vanity project" (Tosi & Warmke, 2016: 199). Complimentary research has shown that observers will question the authenticity of moral actors perceived as pursuing status incentives, thereby undermining status conferral (Bai, Ho, & Wu, 2020).

The heavy use of social pressure can also backfire by causing reactance (Nilsson et al., 2020; Steindl et al., 2015). For example, if an online consumer or employee platform leverages social pressure and reputation by informing users that their moral choices will be made public, this can inflame resentment. In fact, even milder forms of pressure might trigger misgivings if people feel coerced and especially if they believe they have good reason to reject the purported aims of the cause itself. In such instances, agents may feel their personal freedom is subtly under threat, leading to reactance (Steindl et al., 2015). If repeatedly felt, reactant emotions can fuel resentment or *ressentiment* – a persistent politicized grievance (Nietzsche, 1998: 20-25). This phenomenon appears at work in the increasingly divisive politics witnessed today, engendering a toxic "inversion of values" where leaders convince followers that what is considered good by those they resent is bad, and vice versa (Ciulla, 2020).

Fortunately however, we believe that ethics interventions that trigger the reactive mind can still be leveraged without negative habitual consequences so long as they are framed appropriately and used in moderation with the aim of developing reflective capacities. This is the motivational model we suggest, whereby agents are led to progressively higher levels of moral self-awareness.

SOCIAL AND ORGANIZATIONAL IMPLICATIONS

Ethical Boosting

Where the nudge movement emphasizes the limitations to human cognition, Hertwig and Grüne-Yanoff (2017) approach the question of behavioral intervention from a competing tradition in cognitive psychology, more focused on human potential (Gigerenzer, Hertwig, & Pachur, 2011). They suggest interventions aimed at leveraging human learning and growth, and coin their concept "boosting", which they distinguish from nudging along seven dimensions. These pertain to programmatic ambition and purpose, empirical assumptions, and normative implications. Essentially, boosting is a *reflective* approach with long-term developmental objectives, whereas nudging is *reactive* and oriented towards relatively short-term behavioral results. Thus, boosting contains a programmatic aim to build competence – domain-specific or generalizable – with effects lasting well beyond the period of intervention. This contrasts with nudging, which aims to remedy a specific, localized behavior – deemed a mistake by the choice architect. The behavioral improvement however is transitory or contingent on the intervention remaining in place. Otherwise, it tends to revert to the pre-intervention state or potentially backfires via reactance and moral crowding-out, as discussed previously.

Although boosting and nudging can coexist as complementary intervention concepts, their respective proponents have emphasized different aspects of human cognitive architecture (Hertwig & Grüne-Yanoff, 2017; Grüne-Yanoff and Hertwig, 2016). The nudge is grounded in the dual-process view of human cognition (Kahneman, 2003; 2011), and tends to seek remedies to mindless or reflexive mistakes, a prototypical example of which is the targeted default option discussed earlier. Boosting does not preclude this perspective on cognition – or human limitations for that matter – but adopts an outlook more focused on human potential and capacity for competency-building. Whereas nudging is aimed at correcting mindless mistakes, boosting addresses the other side of the coin, building competence through

improvement of mindful behavior. With boosting, the policy maker is less a choice architect and more a coach or teacher.

Normatively, nudging and boosting differ in their approach to transparency and autonomy. Nudging, often acting as a subtle intervention, may violate both, whereas boosting by design requires transparency and willful participation on the part of the subject (Hertwig & Grüne-Yanoff, 2017). This quality, together with its orientation towards competence building and enduring results, makes the boosting concept especially suitable for the domain of ethics, where a "correct" solution may not be evident and competent agency is necessary. Unfortunately, the current trend toward heavy reliance on reactive behavioral interventions may have distracted from this perhaps less obvious alternative.

Self-determination theory (SDT) provides further insight into how boosting differs from nudging in the context of employee motivation (Deci & Ryan, 2000). It posits that intrinsic motivation, referred to as autonomous motivation, stems from the individual's need for autonomy, competence, and relatedness, which are integral for intrinsic motivational boosting to take place (Gagné & Deci, 2005). In contrast, SDT explains that extrinsic motivation, referred to as controlled motivation, is a different type of work motivation where the individual responds to external factors such as rewards directly linked to performance – similar to nudging and other reactive behavioral interventions – and also may include rules, regulations, and direct supervisory oversight. As such, individuals who experience intrinsic motivation are likely to have more freedom to engage meaningfully and creatively than those performing on the basis of extrinsic motivation (Deci & Ryan, 1985). Moreover, student feelings of self-determination correlate with first-year college retention, leading researchers to propose SDT-based pedagogical techniques such as using informational rather than controlling language, offering choices, and providing meaningful feedback (Graham & Vaughan, 2022).

Hertwig and Grüne-Yanoff (2017) introduce a brief taxonomy of boosting to illustrate applications across domains, namely, risk literacy, uncertainty management, and motivational

boosts. We would propose adding a fourth category: ethics boosts, which could draw on similar cognitive and motivational tools. Ethics boosts could involve heuristics that may help individuals identify ethical dilemmas in the workplace, marketplace, or personal life – answering Bazerman and Sezer's (2015) call to mitigate bounded ethical awareness resulting from cognitive constraints. Sometimes, this could be as simple as rules to remind ourselves of the standards to which we aspire, akin to the Golden Rule of *do onto others as you would have them do onto you*. Such rules of thumb could assist individuals in identifying situations that call for special consideration, to avoid the pitfall of mindless mistakes. Consumers, for example, may learn to double-check sources before sharing on social media material pertaining to sensitive topics; they may evaluate the sourcing of certain categories of meat, such as poultry, known for problematic industry standards; or evaluate the emissions implications of long-distance travel (Friedland, 2019). Aside from helping identify dilemmas, boosts in the form of heuristics can also provide guidance for ethical conduct. Agents may learn to respond less reactively in disagreements, to compromise and "split the difference" in difficult negotiations, or to tip more conscientiously. Indeed, tipping norms have been mindfully reinforced and extended during and beyond pandemics or extreme weather events, when tipped workers are seen as essential. Such contexts provide ethical boosting opportunities via online interfaces encouraging customers to acknowledge and reciprocate for their newfound reliance on such workers they once considered non-essential (Friedland & Balkin, 2023).

The distinction between nudging and boosting is sometimes a matter of conceptual framing. Whereas a nudge to promote tipping or employee charitable giving might leverage reactive behavioral responses based on social pressure, boosts could frame the choice as an invitation to reflect on moral identity via apt ethical heuristics, consistent with Zhang, Gino, and Bazerman's s (2014) values-oriented approach to reducing dishonesty. So, for instance, a ride-sharing company might show customers how well they have been rated by past drivers, or offer special gifting options for workers who exceed expectations (Friedland & Balkin, 2023). Such communications could be framed in terms of the Golden Rule or an expression of the unifying

social purpose the organization stands for and strives to achieve. The organization could also offer to match customer or employee donations to relevant charities serving adjacent or associated localities. Here, it may be good to avoid social pressure interventions, which can backfire via reactance if customers or employees feel they have good ethical reasons for not participating – especially if the company itself isn't sacrificing anything. Such conceptual framing is thereby more likely to become internalized and durable over the longer term via the activation of moral self-awareness, as discussed in the following section.

Many mindset interventions, such as those introduced by Carol Dweck (2006), can also be understood as boosts. A growth mindset, for example, refers to an individual's belief that intelligence is dynamic, developing over time – in contrast to a fixed mindset, which takes intelligence as immutable (Rattan et al., 2015) – and relatively brief interventions have proven effective. Yeager et al. (2019: 364) demonstrated that an online growth-mindset intervention, lasting less than one hour, "improved grades among lower-achieving students and increased overall enrolment to advanced mathematics courses in a nationally representative sample of students in secondary education in the United States." By analogy, we would expect that mindset interventions in ethics domains hold promise. One could imagine, for example, a growth-mindset intervention implemented to boost pro-social behavior in corporate social dilemmas, say in the context of teamwork, where collegiality requires effort and creativity.

Related in spirit to the growth mindset, but arising from a more clinical orientation, *motivational interviewing* (Miller, 1983) aims to engender commitment to positive change by eliciting the "client's desire, ability, reasons, and need for change" (Hettema, Steele, & Miller, 2005: 92), to which the interviewer responds by reflective listening and regularly summarizing the client's self-motivational statements (Miller & Rollnick, 2002). The concept thereby leverages the principle from self-perception theory (Bem, 1972) that people become more committed to actions that they themselves defend. With its emphasis on autonomy, intrinsic resources, and lasting impact (Hall, Gibbie, & Lubman, 2012), it can thus be considered a boosting technique, and one especially directed at avoiding the pitfalls of reactance to

perceived coercion, a major challenge in counseling consultations (Harakas, 2013). Motivational interviewing was originally used by general practitioners to encourage patients to adopt healthier lifestyles – especially in the context of addiction – but has since expanded out of the clinic and into other organizational settings, such as health coaching of employees (Butterworth et al., 2006). It may therefore present a powerful complement or alternative to financial incentives in employee wellness programs previously discussed, by more effectively inspiring individuals to take ownership of their own health.

Training interventions consistent with the cognitive boosting concept have also proven to be effective for correcting biases in decision making and judgment. A prominent example is the Sirius program of the Intelligence Advanced Research Projects Activity (2011), which employs virtual simulation games where cognitive skills are instilled through active experience-based engagements, in line with the development of practical wisdom or *phronesis* as discussed previously (Aristotle, 2012: 115-134). The program yielded promising results showing that two one-shot training interventions debiased participants both immediately following the intervention and two months later (Morewedge et al., 2015). The debiasing effects also had broad relevance across organizational contexts in professional and private life, specifically regarding confirmation bias. Sellier, Scopelliti, and Morewedge (2019) successfully implemented the training to reduce confirmation bias in business case write-ups submitted for professionally-oriented graduate programs. Given that similar cognitive biases often impair ethical decision-making, especially in cases involving conflicts of interest (Chugh, Bazerman, & Baniai, 2005), such simulation techniques might prove effective for building virtuous habits in ethical contexts.

Moral Self-Awareness

The Moral Self-Awareness (MSA) construct (Friedland & Cole, 2019; Friedland, 2019) offers an apt framework for ethical boosting, compatible with Dweck's (2006) growth mindset, mindfulness interventions promoting pro-sociality (Schindler & Friese, 2022), social

contemplation promoting ethicality (Gunia et al., 2012), and deliberation promoting honesty (Bereby-Meyer & Shalvi, 2015). While the model is relatively new, it is firmly embedded in the well-established moral identity literature (Aquino & Reed, 2002) and has been empirically confirmed, specifically showing that such awareness reduces tolerance of deceptive sales techniques (Xie, Chang, & Rank-Christman, 2022). The four-stage model is rooted in the virtue-theoretical approach, which leverages a web of affective, cognitive, and motivational responses (Waring, 2016). While its stage-based developmental account is somewhat comparable to Kohlbergian moral stage theory, it engages affective and motivational responses to a far greater degree than the strictly cognitive Kohlbergian approach (Kohlberg, 1981). It does so by encouraging moral development through reflection on moral identity when considering the impact of one's behavior, positive and negative, on others and society. In so doing, the MSA framework comprises three fundamental aspects of virtue-oriented reasoning: pride, shame, and guilt, in which pride is increasingly operative at each successive level of engagement, ultimately becoming the sole motivation at the fourth. This approach has been offered to support virtue-oriented mindsets in business education, including pedagogical strategies for ethical boosting in the business school context (Friedland & Jain, 2022).

MSA boosting is effectuated by *showing notices,* which act as cognitive speed bumps (Calo, 2013), by engaging ethical reflection, which functions within the "slow" part of cognitive dual-process theory (Kahneman, 2011). Such notices reinforce individuals' moral motivation while leading them up the chain of increasing MSA (Friedland, 2019). A key challenge for effective boosting design is to reach agents at their present level of MSA. Just as a fourth-level appeal may fall on deaf ears to someone only at the first or second level, a first-level appeal may be perceived as condescending to someone already at the third or fourth. Within managerial and consumer contexts, individuals' past behavior can be a reliable indicator of where they engage best. Below, we provide suggestions for boosting employee and consumer MSA at each of the four levels.

MSA level 1 - social reflection. At the first level, individuals rely chiefly upon negative feedback from observers in order to guilt or shame them into avoiding harming others or defiling their surroundings. Imagine notices reminding public transit riders not to litter, play loud music, or place their feet on the seats. This lowest level of motivation functions partly as a nudge as it leverages a reflexive reaction to social pressure. However, its focus on moral identity is designed to lead agents to higher levels of reflective self-awareness and autonomy. For example, visual depictions of each offensive behavior might plant seeds of reflection on the kind of person individuals want to be. Another example currently in use on ride apps, such as Uber and Lyft, are profile updates showing customers how well they have been rated by past drivers (Friedland & Balkin, 2023). Such notices have been shown to be particularly effective in organizational contexts, even undoing egoistic priming. For example, college students primed to be self-interested via a non-cooperative Tragedy of the Commons gaming experiment, gradually learned to temper their self-interest in subsequent rounds after being shamed by other subjects left with fewer resources (Sadowski et al., 2013; 2015). At a key moment in the experiment, one subject asked: "Are we bad people?" Eventually, all subjects showed preference for lowered individual returns in favor of equitable and sustainable outcomes. This suggests that the exercise plausibly yielded longer-term ethical boosting effects. Employing similar non-cooperative gaming techniques in a classroom environment may be effective in boosting future professionals' awareness of the complex tensions regarding climate change and other moral problems (Spierre et al, 2011).

MSA level 2 - self-reflection. At the second level, individuals become more self-reflective – often following another person's positive example in considering the interests of others and their surroundings. Imagine observing someone picking up litter and saying to oneself, "Hmmm, maybe I should do that, too...". Rather than relying on castigation of *negative* behavior, actors acknowledge a *positive* action and start to take pride in participating, thereby becoming their own aspirational source of feedback. For example, Cialdini, Kallgren, and Reno (1991) showed that participants were 2.5 times less likely to litter in a room containing

swept litter than in an unswept litter-strewn room. Observing the swept litter informs observers that others have taken the trouble to clean the room, thereby inviting others to join in to help keep it that way. This is where cognitive boosting begins to take fuller self-reflective form. This level of awareness might be boosted by notices of apt ethical heuristics such as the Golden Rule (deontology), the greatest good for the greatest number (utilitarianism), and especially one's idealized self (virtue theory). A level-two boosting technique might invite public transit riders to offer their seats to the elderly or disabled. Similarly, during supermarket shortages from pandemics or extreme weather events, customers could be asked to limit their purchases of household essentials such as toilet paper. Useful associated heuristics in such instances might include the image of a customer handing a package to a grateful parent with a child in tow, as a subtle reminder of the Golden Rule, or an image of a customer returning a shopping cart from the parking lot, in line with the so-called shopping-cart theory of moral character (Hauser, 2021).

MSA level 3 - anticipatory self-reflection. At the third level, individuals start to become forward looking, conceiving of potential negative impacts on others before acting. This forward-looking attitude arises from taking pride in level-two habituations and reflecting on how prior negative behavior has led to an *internal* sense of guilt or shame. Such a mindset could, for example, lead to correcting myriad unhealthy workplace patterns, such as hostile message chains. A level-3 boosting notice might remind users to wait 24 hours before replying to any ambiguous email messages they might perceive as callous or insulting, thereby allowing time for tempers to cool and the more reflective mind to engage. An image of an angry person typing with rage might be an effective visual illustration. Ultimately, such boosts could lead agents to increased self-control and to engage in more productive communication. Applied more broadly, they could appeal to individuals' revealed preferences to help improve their patterns of consumption towards longer-term reductions in their carbon footprints, lower use of toxic chemicals, or pursue healthier lifestyles. This could be accomplished via carbon-count data provided for high-footprint purchases, such as air travel

or luxury automobiles. This differs from mere calorie-counting menu nudging, for instance, by only engaging users who have already acted in ways consistent with these aims. As such, it becomes a boosting self-acknowledgment of their aspirational moral identity.

MSA level 4 - proactive self-reflection. At the fourth and highest level, individuals begin to internalize the ideal of the self as potential hero rather than potential wrongdoer or sinner (Golpadas, 2014). They no longer focus on avoiding shameful behavior, instead taking pride in realizing long-term positive outcomes for other stakeholders beyond themselves. Here, agents recognize their place within a broader system and self-actualize by working toward the greater good. This is the state of mind ethical shoppers hold, as their purchases tend to be motivated much more by positive moral identity as opposed to empathy-based nudges (Hwang & Kim 2016). At this level, ethical boosting is most effective as actors fully engage the reflective mind toward their idealized selves. Such decisions are conscientiously habit-forming, bringing persons closer to becoming whom they aspire to be (Aristotle, 2012). This level of awareness can be activated at the employee level via charitable giving or volunteering with NGOs in communities where a business operates. Employers can most effectively stimulate this attitude by demonstrating sincerity via matching gifts or providing paid time to volunteer. Such commitments confirm the unifying social purpose of the organization, thus reinforcing stakeholder solidarity (Friedland, 2011).

We provide examples of possible intervention techniques for ethical boosting in table 1, linked to each of the four levels of moral self-awareness examined in this section.

TABLE 1

Moral Self-Awareness Boosting in Four Levels

	Social Impact Focus	Feedback Focus	Motivators	Intervention Technique
Level 1 Social reflection	• Negative, via others	• Negative • Social	• Guilt • Shame	• Appeal to agents' moral self-image and idealized self. • Example: Gig economy customer profile notices - show them how well they have been rated by past gig workers.
Level 2 Self-reflection	• Negative, via reflection on outcomes achieved by others	• Negative • Social • Self	• Guilt • Shame • Pride	• Leverage ethical heuristics, i.e. Golden Rule; Greatest Good for Greatest Number; Virtue. • Example: Notices with images of considerate actors behaving helpfully toward others, i.e. offering a bus seat to the elderly or disabled; sharing essential supplies; returning shopping carts.
Level 3 Anticipatory Self-reflection	• Negative, via forward-looking consideration	• Negative • Self	• Guilt • Shame • Pride	• Draw attention to longer-term impacts of individual actions. • Examples: Supply carbon footprint data for air travel and luxury auto purchases. Appeal to agents' revealed preferences to reduce carbon footprints; lower use of toxic chemicals; lead healthier lifestyles.
Level 4 Proactive Self-reflection	• Positive, via self-actualization of internalized values	• Positive • Self	• Pride	• Highlight agents' place within a broader stakeholder picture and unifying social purpose of the organization. • Example: Offer by organization for matching employee gifts to local charities or time for employees to work with a local NGO.

Notes: This table is developed from Friedland & Cole, 2019; Friedland & Balkin, 2023.

Toward a Social and Organizational Policy of Ethics Interventions

Given the concepts outlined, the policy question arises for how managers and policymakers can determine the optimal composition and sequence of nudges and boosts. A starting point would be to convene thought leaders in choice architecture to explore, develop, and articulate best practices. The default nudge, for example, might be held up as an exemplar of an intervention that tends to preserve employee autonomy, and is unlikely to backfire and interfere with work that requires creativity and innovation. Similarly, interventions that that rely on innocuous incentives to influence employee or consumer behavior may also be prescribed as best practices. However, at certain inflection points – which remain to be determined – thought leaders might agree on red flags that alert the choice architect to potential trouble spots where nudging ceases to be beneficial and appears to generate negative externalities, such as dependency, reactance or moral crowding out. At such points, alternative practices such as ethics boosts could be recommended to circumvent the externalities while stimulating ethical growth and reflection. Once identified, best practices could be disseminated and discussed within professional societies – in fields such as management consulting, human resource management, and business ethics – so that social and organizational guidelines for ethics intervention techniques can be shaped and defined.

Directions for Future Research

Our journey through the variegated landscape of ethics interventions suggests numerous avenues for future research. First, as suggested above, studies across organizational contexts could seek to determine inflection points empirically, beyond which reactive behavioral interventions cease to be helpful in promoting ethics, and instead backfire from the TMGT effect. Future work, for example, could try to examine more closely the point at which innocuous incentives bring about moral crowding out. This might be accomplished by exploiting systematic variation in incentives – as Wollbrant et al. (2022) do in the field and Schneider et al. (2023) in survey-based randomized-controlled trials. Secondly, studies could

be conducted comparing the effectiveness of employee wellness programs employing mild monetary incentives versus motivational interviewing (Butterworth et al., 2006) or other ethics boosting approaches.

Another line of research, already begun by Xie et al. in the context of sales (2022) investigates the extent to which MSA ethics boosts may affect tolerance of unethical business practices. The MSA construct offers much fertile ground for empirical exploration, testing, and application. For example, future studies could develop tools for classifying individuals into each of the four MSA levels and test level-contingent interventions. There is also scope for examining MSA interventions across organizational levels, ranging from the rank-and-file to the C-suite, both via formal educational programs or less formal coaching practices. Future studies could explore avenues for ethical boosting through workplace training modules and continuing education. Ultimately, structured educational programs are thought well-placed for nurturing reflection (Hibbert, 2013) and, presumably, for stimulating effective ethical agency.

CONCLUSION

Behavioral intervention techniques are clearly advantageous for setting employees and consumers along the path of developing increased ethical awareness. Nudges are especially useful for reaching agents in the early stages of ethical self-reflection and thus should remain an important part of the managerial and marketing toolkit. However, their ready accessibility yielding instant short-term results to specific and well-defined problems, puts them at risk of becoming a kind of fast-food for managers and policy wonks. Thought leaders should remember that heavy reliance on reactive behavioral interventions can lead to questionable long-term consequences, including motivational displacement, dependency, moral crowding out, loss of personal autonomy, and reactance. There is also evidence suggesting that the widespread popularity of nudges may overstate their effectiveness via publication bias (Maier et al., 2022), and choice architects may fall prey to conflicts of interest by implementing

inefficient interventions with the aim of using them again in the future (Hertwig & Ryall, 2020). Unfortunately, the ubiquity of the nudge concept has potentially diverted our attention away from promising policy tools that build human capability, such as ethical boosting, which provide more long-term benefits, including solutions to problems we have yet to imagine. This stands in contrast to the nudging paradigm, which solves only *well-defined* social problems already imagined – and even then, only to a limited extent. Unlike the behaviorally nudged, ethically boosted individuals should be able to apply their heightened ethical awareness and competence to solve new and unforeseen social problems as they arise.

Ultimately, the choice between a social policy based on behavioral reaction versus one based on ethical reflection represents an existential question for each of us as to what kind of society we aspire to build and be a part of. Let us, therefore, frame this quandary as a thought experiment inspired by the philosopher John Rawls in his masterwork, *A Theory of Justice* (1971).

Picture a future in which the world's population is divided into two parts, and sent to live on different planets. On one – let's call it the Brave New World (BNW) – citizens have opted for a complete and total commitment to nudging principles. On BNW, everyone is conditioned from an early age via reactive interventions to behave as the culture and designers deem best. On the other – we'll call it the Moral Self-Awareness (MSA) society – citizens have chosen an alternate route, committing to activating ethical autonomy. On MSA, everyone is encouraged from an early age via reflective boosting techniques to act with their highest aspirations in mind. Huxley's last novel *Island* (1962) actually depicts a utopian version of such a place, which includes bird calls reminding residents to focus mindfully on the here and now.

Now suppose you're waking up from a cryogenic sleep, which began well before the split. Several generations have passed on both planets, but you know nothing more about the societal developments, logical challenges, or broader ethical issues to which time has given rise on either planet. Honoring your cryogenic agreement, your position in either society

would begin with an all-expenses covered adaptation period, after which you may be expected – age and health permitting – to take on an available occupation appropriate to your interests and abilities.

In which society would you prefer to continue living? According to Rawls (1971), rational persons would opt for the system offering the most individual justice in which they would be most likely to thrive to their fullest potential. For us, this question is not entirely hypothetical. It largely represents an existential choice we collectively face today.

REFERENCES

Aldred, J. 2009. *The skeptical economist: Revealing the ethics inside economics*. London: Routledge.

Allcott, H. 2011. Social mores and energy conservation. *Journal of Public Economics*, 95(9-10): 1082-1095.

Aquino, K., & Reed II, A. 2002. The self-importance of moral identity. *Journal of Personality and Social Psychology,* 83(6): 1423-1440.

Aristotle. 2012. *Aristotle's nicomachean ethics*. Chicago, Il: University of Chicago Press.

Arshad, A., Anderson, B., & Sharif, A. 2019. Comparison of organ donation and transplantation rates between opt-out and opt-in systems. *Kidney International*, 95(6): 1453-1460.

Bai, F., Ho, G. C. C., & Wu, L. 2020. Do status incentives undermine morality-based status attainment? Investigation the mediating role of perceived authenticity. *Organizational Behavior and Human Decision Processes*, 158: 126–138.

Bazerman, M. H., & Sezer, O. 2016. Bounded awareness: Implications for ethical decision making. *Organizational Behavior and Human Decision Processes*, 136: 95–105.

Berry, S. 2023, April 12. Geofencing marketing: What is it and how can you get started? *Web FX*. https://www.webfx.com/blog/marketing/geofencing-marketing/#

Bhargava, S., Conell-Price, L., Mason, R. T., & Benartzi, S. 2021. Save(d) by Design. Working paper. Available at SSRN: https://ssrn.com/abstract=3237820 or http://dx.doi.org/10.2139/ssrn.3237820

Bem, D. J. 1972. Self-perception theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, 6:1–62. New York: Academic.

Bereby-Meyer, Y. & Shalvi, S. 2015. Deliberate honesty. *Current Opinion in Psychology*, 6: 195-198.

Bowles, S. 2016. *The moral economy: Why good incentives are no substitute for good citizens.* New Haven, CT: Yale University Press.

Brown, M. E., Treviño, L. K., & Harrison, D. A. 2005. Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes*, 97: 117-134.

Butterworth, S., Linden, A., McClay, W., & Leo, M. C. 2006. Effect of motivational interviewing-based health coaching on employees' physical and mental health status. *Journal of Occupational Health Psychology, 11*(4): 358–365.

Calo, M. R. 2013. Against notice skepticism in privacy (and elsewhere). *Notre Dame Law Review,* 87(3): 1027-1072.

Chugh, D., Bazerman, M. H., & Banjai, M. R. 2005. Bounded ethicality as a psychological barrier to recognizing conflicts of interest. In D. A. Moore, D. M. Cain, G. F. Loewenstein & M. H. Bazerman (Eds.), *Conflicts of Interest: Challenges and Solutions in Business, Law, Medicine, and Public Policy*: 74-95. New York, NY: Cambridge University Press.

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. 1991. A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in Experimental Social Psychology*, 24: 201–234.

Ciulla, J. B. 2020. Leadership and the power of resentment/ressentiment. *Leadership*, 16(1): 25-38.

Davidai, S., Gilovich, T., & Ross, L. D. 2012. The meaning of default options for potential organ donation. *Proceedings of the National Academy of Sciences*, 15201-15205.

Deci, E. L., & Ryan, R. M. 1985. *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E. L., & Ryan, R. M. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1): 68-78.

Dewey, J. 1909. *Moral principles in education*. Boston, MA: Houghton Mifflin.

Dunlap, R. E., Rosa, E. A., Baxter, R., & Mitchell, R. 1993. Attitudes toward siting at a high-level nuclear waste repository at Hanford, Washington, In R .E. Dunlap, M. E. Kraft & E. A. Rosa (Eds.), *Public reactions to nuclear waste*: 136-172. Durham, NC: Duke University Press.

Dweck, C. S. 2006. *Mindset: The new psychology of success*. New York, NY: Random House.

Edwards, W. 1983. Human cognitive capabilities, representativeness, and ground rules for research. *Advances in Psychology*, 14: 507-513.

Friedland J. 2011. Wittgenstein and the Challenge of Global Ethics, In C. Dierksmeier, et al. (Eds.), *Humanistic ethics in the age of globality*: 210-222. New York, NY: Palgrave-Macmillan.

Friedland, J. 2019. AI can help us live more deliberately. *MIT Sloan Management Review*, 60(4): 45-51.

Friedland, J., & Balkin, D. B. 2023. When gig workers become essential: Leveraging customer moral self-awareness beyond COVID-19. *Business Horizons*, 66(2): 181-190

Friedland, J., & Cole, B. M. 2019. From homo-economicus to homo-virtus: A system-theoretic model for raising moral self-awareness. *Journal of Business Ethics*, 155(1): 191-205.

Friedland J., & Jain T. 2022. Reframing the purpose of business education: Crowding-in a culture of moral self-awareness. *Journal of Management Inquiry*, 31(1): 15-29.

Frischmann, B. 2022. Nudging humans. *Social Epistemology*, 36(2): 129-152.

Frischmann, B. & Selinger, E. 2018. *Re-engineering humanity*. Cambridge, UK: Cambridge University Press.

Gagné, M., & Deci, E. L. 2005. Self-determination theory and work motivation. *Journal of Organizational Behavior*. 26: 331-362.

Gibbs, G. 2013. Reflections on the changing nature of educational development. *International Journal for Academic Development*, 18(1): 4-14.

Gibson, J. W., & Tesone, D. V. 2001. Management fads: Emergence, evolution, and implications for managers. *Academy of Management Perspectives*, 15(4): 122-133.

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

Gilligan, C. 1982. *In a different voice: Psychological theory and women's development.* Boston, MA: Harvard University Press.

Gillman, J. 1999. Religious perspectives on organ donation. *Critical Care Nursing Quarterly*, 22(3): 19-29.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. 2008. A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, 35(3): 472-482.

Golpadas, A. 2014. Marketplace sentiments. *Journal of Consumer Research*. 41(4): 995–1014.

Goodhart, C. A. E. 1975. *Problems of monetary management: The UK experience. Papers in monetary economics.* Sydney: Reserve Bank of Australia.

Gneezy, U., & Rustichini, A. 2000. A fine is a price. *Journal of Legal Studies*, 29(1): 1-17.

Graham, M., & Vaughan, A. 2022. An adapted self-determination measure and college student first-year achievement. *International Journal of Teaching and Learning in Higher Education*, 33(2): 135-142.

Grüne-Yanoff, T., & Hertwig, R. 2016. Nudge versus boost: How coherent are policy and theory? *Minds and Machines*, 26: 149–183.

Gunia, B. C., Wang L., Huang, L., Wang, J., & Murnighan, J. K., 2012. Contemplation and conversation: subtle Influences on moral decision making. *Academy of Management Journal*, 55: 13-33.

Haidt, J. 2001. The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review,* 108(4): 814–834.

Haidt, J. 2007. The new synthesis in moral psychology. *Science,* 316(5827): 998–1002.

Haidt, J. 2013. Moral psychology for the twenty-first century. *Journal of Moral Education,* 42(3): 281–297.

Hall, K., Gibbie, T., & Lubman, D. I. 2012. Motivational interviewing techniques: Facilitating behaviour change in the general practice setting. *Australian Family Physician*, 41(9): 660–667.

Harakas, P. 2013. Resistance, motivational interviewing, and executive coaching. *Consulting Psychology Journal: Practice and Research, 65*(2): 108–127.

Hardy, S., & Carlo, G. 2011. Moral identity: What is it, how does it develop, and is it linked to moral action? *Child Development Perspectives*, 5: 212–218.

Hauser, C. 2021, June 13. Everyone has a theory about shopping carts. *New York Times*. https://www.nytimes.com/2021/06/08/style/shopping-cart-parking-lot.html

Hertwig, R., & Grüne-Yanoff, T. 2017. Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, 12(6): 973-86.

Hertwig, R., & Ryall, M. D. 2020. Nudge versus boost: Agency dynamics under libertarian paternalism. *The Economic Journal*, 130(629): 1384-1415.

Hettema, J., Steele, J., & Miller, W. R. 2005. Motivational interviewing. *Annual Review of Clinical Psychology*, 1: 91–111.

Hibbert, P. 2013. Approaching reflexivity through reflection: issues for critical management education. *Journal of Management Education*, 37(6), 803–827.

Hwang, K., & Kim, H. 2016. Are ethical consumers happy? Effects of ethical consumers' motivations based on empathy versus self-orientation on their happiness. *Journal of Business Ethics*, 151(2): 579–598.

Huxley, A. 1932. *Brave new world*. London: Chattow & Windus.

Huxley, A. 1958. *Brave new world revisited*: New York: Harper & Row.

Huxley, A. 1962. *Island.* New York: Harper & Row.

Intelligence Advanced Research Projects Activity. 2011. Sirius broad agency announcement, IARPA-BAA-11-03. http://www.iarpa.gov/index.php/research-programs/31irius/baa

Kahneman, D. 2003. A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58: 697–720.

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

Kohlberg. L. 1981. *The Philosophy of moral development: Moral stages and the idea of justice.* San Francisco: Harper & Row.

Kohlberg, L., & Mayer, R. 1972. Development as the aim of education. *Harvard Educational Review*, 42(4): 449-496.

Kunreuther, H., & Easterling, D. 1996. The role of compensation in siting hazardous facilities. *Journal of Policy Analysis and Management,* 15(4): 601-622.

Lapsley, D., & Carlo, G. 2014. Moral development at the crossroads: New trends and possible futures. *Developmental Psychology*, 50(1): 1–7.

Lawton, A., & Páez, I. 2015. Developing a framework for ethical leadership. *Journal of Business Ethics*, 130: 639-649.

Leuker, C., Samartzidis, L., Hertwig, R., & Pleskac, T. J. 2020. When money talks: Judging risk and coercion in high-paying clinical trials. *PloS ONE*, 15(1): e0227898.

Lord, C. G., Ross, L., & Lepper, M. R. 1979. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology,* 37(11): 2098–2109.

McGhee, E., Mehlotra, R., & Romero, M. 2021, June 2. Improving California's automatic voter registration. *Public Policy Institute of California*. https://www.ppic.org/blog/improving-californias-automatic-voter-registration/

Maier, M., Bartos, F., Stanley, T. D., Shanks, D. R., Harris, A. J. L., & Wagenmakers, E. 2022. No evidence for nudging after adjusting for publication bias. *Proceedings of the National Academy of Sciences*, 119(31) e2200300119

Mellström, C., & Johannesson. 2010. Crowding out in blood donation: Was Titmus right? *Journal of the European Economic Association*, 6(4), 845–863.

Miller, R. 2023, April 8. Generative AI could transform the way we interact with enterprise software. *TechCrunch.* https://techcrunch.com/2023/04/08/generative-ai-enterprise-software/

Miller, S. 2020, October 16. Employers fine-tune wellness incentives, wait for EEOC guidance. *Society for Human Resource Management*. https://www.shrm.org/ResourcesAndTools/hr-topics/benefits/pages/employers-fine-tune-wellness-incentives-and-wait-for-eeoc-guidance.aspx

Miller, W. 1983. Motivational interviewing with problem drinkers. *Behavioural and Cognitive Psychotherapy,* 11(2): 147–172.

Miller., W. R. & Rollnick, S. 2002. *Motivational interviewing: Preparing people for change*. Vol 2. New York: Guilford.

Möhlmann, M. 2021, April. Algorithmic nudges don't have to be unethical. *HBR.org* https://hbr.org/2021/04/algorithmic-nudges-dont-have-to-be-unethical

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. 2015. Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from the Behavioral and Brain Sciences*, 2(1): 129-140.

Nielsen, T. C., & Kepinski, L. 2013. Inclusion nudges for motivating allies: Action guide with 30 examples of behavioural designs. Independently published. ASIN: B08GV91VYP.

Nietzsche, F. 1998. *On the geneology of morality*. Indianapolis, IN: Hackett.

Nilsson, A., Erlandsson, A., Västfjäll, D., & Tinghög, G. 2021. Who are the opponents of nudging? Insights from moral foundations theory. *Comprehensive Results in Social Psychology*, 5(1-3): 64-97.

Niza C., Tung B., & Marteau T. M. 2013. Incentivizing blood donation: systematic review and meta-analysis to test Titmuss' hypotheses. *Health Psychology*, 32(9): 941-9.

Noar S. M., Hall M. G., Francis D. B., et al. 2016. Pictorial cigarette pack warnings: a meta-analysis of experimental studies. *Tobacco Control,* 25**:** 341-354.

O'Leary, J. & Murphy, T. 2017, May 16. How state and local governments can use nudge thinking to improve outcomes. *Deloitte Insights*. https://www2.deloitte.com/us/en/insights/focus/behavioral-economics/how-state-local-governments-can-apply-nudge-thinking.html

Pierce, J. R. & Aguinis, H. 2013. The too much of a good thing effect in management. *Journal of Management*, 39(2): 313-338.

Piaget, J. 1932. *The moral judgment of the child*. London: Kegan, Paul, Trench, Trubner & Co.

Possamai, C. G., Ravaud, P., Ghosn, L., & Tran, V. 2020. Use of wearable biometric monitoring devices to measure outcomes in randomized clinical trials: A methodological systematic review. *BMC Medicine*. https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-020-01773-w

Rattan, A., Savani, K., Chugh, D., & Dweck, C. S. 2015. Leveraging mindsets to promote academic achievement: Policy recommendations. *Perspectives on Psychological Science*, 10(6): 721-726.

Ryan W. S., & Ryan R. M. 2019. Toward a social psychology of authenticity: Exploring within-person variation in autonomy, congruence, and genuineness using self-determination theory. *Review of General Psychology*, 23(1): 99-112.

Rawls, J. 1971. *A theory of justice*. Cambridge, MA: Harvard University Press.

Sadowski, J., Seager, T. P., Selinger, E., Spierre, S. G., & Whyte, K. P. 2013. An experiential, game-theoretic pedagogy for sustainability ethics. *Science and Engineering Ethics,* 19: 1323–1339.

Sadowski, J., Seager, T. P., Selinger, E., Spierre, S. G., & Whyte, K. P. 2015. Intergroup cooperation in common pool resource dilemmas: The role of ethical leadership. *Science and Engineering Ethics*, 15: 1197–1215.

Sandel, M. J. 2012. *What money can't buy: The moral limits of markets*. New York: Farrar, Straus and Giroux.

Schaefer, J. 2015, February 24. The real ROI for employee wellness programs. *Society for Human Resource Management*. https://www.shrm.org/resourcesandtools/hr-topics/benefits/pages/real-roi-wellness.aspx

Schneider, F. H., Campos-Mercade, P., Meier, S., Pope, D., Wengström, E., & Meier, A. N. 2023. Financial incentives for vaccination do not have negative unintended consequences. *Nature*, 613: 526–533.

Schindler, S., & Friese, M. 2022. The relation of mindfulness and prosocial behavior: What do we (not) know? *Current Opinion in Psychology*, 44: 151–156.

Schwartz, B. 2016. *The Paradox of choice: Why more is less*, Revised Edition. United States: HarperCollins.

Sellier, A.-L., Scopelliti, I., & Morewedge, C. K. 2019. Debiasing training improves decision making in the field. *Psychological Science*, 30(9): 1371-1379.

Shearman, S. M., & Yoo, J. H. 2005. Even a penny will help!: Legitimization of paltry donation and social proof in soliciting donation to a charitable organization. *Communication Research Reports*, 24(4): 271-282.

Skopek, M. J. 2010. Uncommon goods: on environmental virtues and voluntary carbon offsets. *Harvard Law Review*, 123(8): 2065-2087.

Spierre, S. G., Seager, T. P., Selinger, E., & Sadowski, J. 2011. Using non-cooperative games to simulate ethical tensions in climate policy negotiations. *Proceedings of the 2011 IEEE International Symposium on Sustainable Systems and Technology*, Chicago, IL: 1-5.

State of NewYork. 2012. Jury duty: Now's it's your turn to make a difference. *New York State Unified Court System*. https://www.nycourts.gov/

Steindl, C., Jonas, E., Sittenthaler, S., Traut-Mattausch, E., & Greenberg, J. 2015. Understanding psychological reactance. *Zeitschrift Fur Psychologie*, 223(4): 205-214.

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

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

Thaler, R. H., & Sunstein, C. R. 2021. *Nudge: The final edition*. New York: Penguin.

The Economist. 2017, May 20. When nudge comes to shove, 51-52.

Tosi, J., & Warmke, B. 2016. Moral grandstanding. *Philosophy and Public Affairs*, 44(3): 197-217.

Waring, D. R. 2016. *The healing virtues.* Oxford: Oxford University Press.

Wollbrant, C. E., Knutsson, M., & Martinsson, P. 2022. Extrinsic rewards and crowding-out of prosocial behaviour. *Nature Human Behaviour*, 6(6): 774–781.

Xie, G. X., Chang, H., & Rank-Christman, T. 2022. Contesting dishonesty: When and why perspective-taking decreases ethical tolerance of marketplace deception. *Journal of Business Ethics,* 175(1)**:** 117–133.

Yeager, D. S., Hanselman, P., Walton, G. M., et al. 2019. A national experiment reveals where a growth mindset improves achievement. *Nature*, 573: 364–369.

Zhang, T., Gino, F., & Bazerman, M. H. 2014. Morality rebooted: Exploring simple fixes to our moral bugs. *Research in Organizational Behavior*, 34: 63–79.