

MORE WORK FOR NORMATIVE BUSINESS ETHICS
BEHAVIORAL ETHICS – A CASE FOR REFORMATION IN NORMATIVE BUSINESS ETHICS
RESEARCH CONFERENCE ON THE OCCASION OF THE 500 YEARS REFORMATION ANNIVERSARY
Wittenberg, Germany, 28-30 June 2017

Nien-hê Hsieh
nhsieh@hbs.edu

This is a rough draft. Comments are warmly welcomed. Thank you.

Behavioral ethics, it has been said, “is arguably the ‘next big thing’ in ethics teaching and research” (Prentice 2014, 326). Behavioral ethics aims to describe and explain ethical decision-making, motivation, and behavior, and how individuals may differ and how context can make a difference (Treviño et al. 2006, 953).¹ It promises to improve ethical decision-making and behavior in many spheres, including business. By highlighting limitations on moral cognition, for example, behavioral ethics cautions individuals against being overly confident in their moral judgments. In illustrating the relevance of context for behavior, the field suggests ways for managers to reduce unethical conduct in organizations. Moreover, it has been pointed out that developments in behavioral ethics may lead some to conclude that because “people cannot generally conform to the normative ethical principles that business ethics theorists prescribe ... these principles are therefore practically irrelevant” (Kim, Rosemarie, and Strudler 2015, 341).² It is to behavioral ethics, as opposed to normative business ethics, that we should look going forward.

¹ Treviño et al (2006) define behavioral ethics as “primarily concerned with explaining individual behavior that occurs in the context of larger social prescriptions” (952).

² The authors term this the “Radical Behavioral Challenge” and respond to it. To be certain, there is a great deal of scholarship on holding people morally responsible given how much empirical research suggests a lack of control for our actions and attitudes. In this particular case, the authors argue that the “Radical Behavioral Challenge” relies upon an overly narrow interpretation of the principle that “ought implies can.” That is to say, the limitations on our capacity to conform to certain principles of norms need not imply that we Another way in which to respond is to acknowledge “making judgments about reasons, or making evaluative judgments, as things with respect to which we are active and over which we might therefore be said to have a kind of control.” That is to say, “these are exercises of rational capacity, and not things that merely happen to us” (Clarke, McKenna, and Smith 2015). As Angela Smith points out, “we regularly do hold ourselves and others responsible for things that do not appear to reflect a conscious choice or decision” (2005: 236). Smith defends this practice on the grounds that “what makes an attitude ‘ours’ in the sense relevant to questions of responsibility and moral assessment is not that we have voluntarily chosen it or that we have voluntary control over it, but that it reflects our own evaluative judgments or appraisals” (2005: 237). For discussion on responsibility for implicit bias, see the essays in Brownstein and Saul 2016 as well as Holroyd, Scaife, and Stafford 2017. In this context, one way

In this paper, I ask how far behavioral ethics can take us on its own without continued work in the field of normative business ethics. I raise three questions. First, to what extent does the interpretation of results in behavioral ethics require making explicit and adjudicating among theoretical normative commitments? For example, one area of research highlights the influence of situation or context on people's behavior, but to what extent can one develop a plausible interpretation of the nature and significance of this influence without clarifying underlying normative commitments? The second question I raise concerns the permissibility of applying insights from behavioral ethics to the realm of business activity. For example, insofar as certain interventions take advantage of unconscious or subconscious behaviors, should they be considered manipulative in ways that are morally impermissible? Third, I ask just how radical a reform of normative ethics is required in order to incorporate behavioral ethics. Might normative ethical theory already incorporate many of the insights of behavioral ethics? What I suggest based on the brief exploration of these questions in this paper is the need for normative business ethics to continue its line of work.

1. Interpreting Results

In this section, I ask to what extent the interpretation of results in behavioral ethics requires making explicit and adjudicating among theoretical normative commitments. The point I raise is not that behavioral ethics requires at a minimum some commitment as to what kinds of decisions and activity fall within the realm of ethical behavior. For example, whether investigating the factors that affect the choice of the color of one's car qualifies as a topic of behavioral ethics depends in part on one's normative commitments. If morality requires maximizing the overall good and choice of one's

in which individuals are said to be responsible is that "they fail to act properly on the knowledge that they are likely (Saul 2013). Then, of course, there is the more general debate about free will and determinism. In this paper, I do not take up these debates.

car color affects the overall good, then it would qualify. However, if morality admits of a space for personal choices, the choice may not. Instead, the question I raise concerns the extent to which normative commitments affect the interpretation of observed behavior within the field of behavioral ethics. I consider two classic experiments to examine to what extent we have come to take for granted the interpretation of their results.

First, consider the well-known example of framing effects as first described by Amos Tversky and Daniel Kahneman (1981). In the experiment, the authors ask one set of respondents to choose between two programs to combat “an unusual Asian disease, which is expected to kill 600 people” (policies A and B) and another set of respondents to choose between two other programs to combat the same disease (policies C and D). The programs are described to the respondents as follows:

- If Program A is adopted, 200 people will be saved.
- If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.
- If Program C is adopted, 400 people will die.
- If Program D is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.

In terms of outcomes, the choice between Program A and Program B is identical to the choice between Program C and Program D. In each pair, respondents are choosing between 200 people being alive with certainty (Policy A and Policy C) and an expected value of 200 people being alive (Policy B and Policy D). However, what the authors find is that the preference in choosing between A and B is reversed in choosing between C and D. 72% of respondents chose A over B, whereas only 22% of respondents chose C over D. The authors attribute this reversal to “Prospect Theory”

according to which people are risk aversion when considering gains (A over B) and risk seeking when considering losses (D over C).

A great deal of scholarship has been devoted to Prospect Theory and framing effects. The point I highlight relates to the distinction made in ethical theory between doing and allowing. From a consequentialist perspective, there is no relevant distinction. Whether one kills 400 people or allows 400 people to die is irrelevant. 400 people are dead. In contrast, for some nonconsequentialists there is a distinction. Killing 400 people is not the same as allowing 400 people to die. To return to the experiment with this distinction in mind, we can apply a somewhat different frame to the choices, one that is not in terms of gains and losses, but rather in terms of the role of the Program in how deaths are brought about. In the choice between Programs C and D, it is clear both programs will result in deaths. Even if the language is not in terms of killing, insofar as respondents view either Program as playing an active role in the resulting deaths and frame the relevant distinction in terms of doing and allowing, then the choice between Programs C and D looks different from the choice between Programs A and B in which the role of the Programs in deaths is not mentioned at all. Rather relying on Prospect Theory, the interpretation is one that connects to a perceived difference on the part of respondents in terms of doing and allowing. That is, the interpretation of observed behavior and what inferences should be drawn for decision-making depend in part on our underlying normative ethical commitments.

Consider now experiments in which situational factors are said to affect people's behavior in ways they may not realize. One such factor is time pressure as made famous by an experiment involving seminary students asked to deliver a lecture on the parable of the Good Samaritan. As described by Robert Prentice (2014, 338-339):

Consider a very simple situational factor – time pressure. In a very interesting study, psychologists told seminary students that they needed to go across campus to give a

talk to a group of visitors, perhaps about the parable of the Good Samaritan. As they crossed campus to give the talk, the students happened upon a fellow lying by the sidewalk in obvious distress—in need of a Good Samaritan. If they were not under time pressure, almost all the seminary students stopped to help this fellow (who had, of course, been placed there by the experimenters). If students were placed in a “low-hurry” condition, only 63 percent offered help. If they were put in a “medium-hurry” condition, only 45 percent helped. And if asked to really hurry and put in a “high-hurry” condition, only 10 percent stopped to help. Certainly the students involved in the study did not consciously realize how the time pressure they were under dramatically affected their ethical conduct, but it clearly did.

One thing that this and other experiments have been said to show is that because behavior can be affected by seemingly small changes in our situation, there is reason to doubt any robust conception of character of the sort that features prominently in virtue ethics (e.g., Doris 2002, Harman 1999).³ The issue I examine is somewhat more general, which concerns the extent to which what we take away from experiments such as this varies with our underlying normative assumptions.

In their discussion of the original experiment, John Darley and C. Daniel Batson distinguish between two experiences reported by subjects who did not stop to help the person on the sidewalk (1973). For one group, although the subjects mentioned afterwards on reflection that the person possibly needed help, “it seems that they often had not worked this out when they were near the victim” (1973, 107). The authors continue, “either the interpretation of their visual picture as a person in distress or the empathic reactions usually associated with that interpretation had been deferred because they were hurrying” (1973, 108). The experience of this first group is consistent

³ There is a good deal of literature on the character and situationism debate. For a recent review that discusses new direction of research, see Miller 2017. For one discussion specifically in the field of business ethics, see Alzola 2012.

with the interpretation that Prentice describes. The second group of subjects that Darley and Batson describe “appeared aroused and anxious after the encounter in the alley” (108). For these subjects “it seems more accurate to conclude that they decided not to stop” (108). The authors write, “because the experimenter, *whom the subject was helping*, was depending on him to get to a particular place quickly ... he was in conflict between stopping to help the victim and continuing on his way to help the experimenter” (108).⁴ According to Darley and Baston, “conflict, rather than callousness, can explain their failure to stop” (108).

The nature and significance of the impact time pressure on decision-making and behavior is rather different between these two groups of subjects. For the first group of subjects, time pressure can be seen as somehow causing them to deviate from behaving in the way they would ideally want – that is, to help the person on the ground. This, I take it, is one of the underlying aims of behavioral ethics – that is, to expose or highlight gaps we may not even notice between our motivations, behavior or cognition and what we would ideally want them to be. However, to infer that there is this gap assumes that helping the person on the ground *independently of the time pressure* is the morally correct course of action. Consider a slight variation on the experiment in which the subject is asked to cross campus to take the experimenter’s mother to the hospital because she is ill. It is not unreasonable to imagine a similar distribution of those who stop and those who don’t whiles varying time pressure conditions – that is, more will stop under the “low-hurry” condition and many fewer under the “high-hurry” condition. In this variation, however, would we interpret the result of time pressure as we did for the first group in the original experiment? I am not so certain. In this case, it seems just as plausible to interpret the impact of time pressure as a constraint on one’s ability to help others. If the subject believes there is time to help the person on the ground

⁴ Emphasis theirs.

without preventing her from helping another (who may be in greater need), then she is more likely to do so than when she believes she does not have the time.

None of this is to deny the relevance of the impact of time pressure on our behavior. Instead, the question I raise is the extent to which our interpretation of the nature and significance of that impact depends on making explicit and adjudicating among theoretical normative commitments. Insofar as this does, then we find ourselves back in the realm of normative ethical theory. Taken together with the discussion of Prospect Theory, what this discussion suggests is the need for more work in normative theory to help inform our interpretation of the results from behavioral ethics. Indeed, behavioral ethics may want to look to debates in normative ethical theory to determine just what sorts of experiments and studies are of greater interest and relevance than others. This then brings us to the second question which concerns the role of normative ethical theory in providing guidance.

2. Applying Findings

The second question I raise concerns the permissibility of applying insights from behavioral ethics to the realm of business activity. For example, insofar as certain interventions take advantage of unconscious or subconscious behaviors, should they be considered manipulative in ways that are morally impermissible? In what follows, I summarize two applications: one involving customers and another involving employees. What emerges from this discussion, I argue, is that our assessment of the desirability or permissibility of various interventions depends in part on our assessment of managerial responsibilities to customers and employees, which remain contested and a question for normative business ethics to address.

In *Phishing for Phools*, George Akerloff and Robert Shiller (2015) explore the ways in which sellers manipulate and deceive customers, in part through taking advantage of the sorts of tendencies

and cognitive processes described by behavioral ethics. I take it that on most views, such behavior on the part of sellers – at least in its the most extreme forms – is objectionable. Part of what makes such behavior objectionable, I take it, is that customers are not informed about these tactics. Were customers to be informed, the tactics are likely no longer to be as successful for the buyers.

Consider now an experiment designed by an insurance company with the help of insights from behavioral ethics (Shu et al. 2012). After asking new car insurance policy holders to report the mileage on their cars, the company would periodically ask policy holders to report the current odometer reading in order to update the premiums customers would pay. The form used by the company asked customers to sign the statement, “I promise that the information I am providing is true,” at the bottom after providing information on the odometer reading. For the experiment, the company randomly assigned customers into one of two groups. One group received the original form. The second group received a form asking customers to sign the statement, “I promise that the information I am providing is true,” at the top, presumably before they would go on to provide the information on the odometer reading. As reported by the authors of the study, “asking customers to sign at the beginning of the form led to a 10.25% increase in the calculated miles driven over the current practice of asking for a signature at the end” (Shu et al. 2012, 15200). From the perspective of the insurance company, customers reporting higher odometer readings could be charged higher premiums. As in the case of sort of deception and manipulation described by Akerlof and Shiller, the actions of the insurer were not publicized to customers and the insurer was open to the possibility that their actions could influence the behavior of their customers. If one were to find the actions of the insurance company less objectionable than other forms of manipulation on the part of sellers, one reason may be that the company is trying to dissuade customers from engaging in otherwise morally objectionable behavior – namely, lying.

Consider now another experiment targeted at customers designed to dissuade otherwise morally objectionable behavior. In this experiment, the authors write, “we examined the effect of an image of a pair of eyes on contributions to an honesty box used to collect money for drinks in a university coffee room. People paid nearly three times as much for their drinks when eyes were displayed rather than a control image” (Bateson et al. 2006 412). Suppose now that we were to move the experiment to the workplace, pasting images of eyes in various rooms where employers were concerned about employee theft (e.g., supply room, cafeteria, cash register). Is there a difference in our assessment of the desirability or appropriateness of the intervention remain the same between customers and employees?⁵

The aim is not to settle the question of the permissibility or desirability of these or other interventions in the context of business activity. Rather, the point is that just because an intervention is successful in dissuading individuals from engaging in otherwise morally objectionable behavior, desirability or permissibility of its implementation need not be the same in all contexts. In this specific case, part of this assessment may involve underlying views about what managers owe customers and employees and vice versa, which are questions to be resolved in the realm of normative business ethics. For example, insofar as the intervention in the case of employees strikes us as overly manipulative or intrusive, one reason to consider is the difference in power that companies often have over employees and customers. The evaluation of the relevance of this difference in power on the permissibility of interventions is fundamentally a normative one.

⁵ To be clear, the issues discussed here are slightly different from the debate around nudging, liberal paternalism, and choice architecture (Thaler and Sunstein 2009). In the case of nudging by way of choice architecture (e.g., making the default on employee 401(k) plans to be having to opt out rather than opt in), the interventions are to be in the interests of the employee – hence the concern with paternalism. Of course, it may be thought that acting ethically is in one’s interest.

3. Incorporating Behavioral Ethics

In this section, I ask just how radical a reform of normative ethics is required in order to incorporate behavioral ethics. Rather than provide a comprehensive catalogue of the findings in behavioral ethics that are likely to be relevant to normative ethics, I ask instead if what much of behavioral ethics has done is to highlight and make more precise our understanding about aspects of human behavior and cognition that have been noted throughout the history of philosophy. In turn, if this is correct, then normative ethics may already be equipped to incorporate the learnings from behavioral ethics without any need for fundamentally rethinking the enterprise of normative ethics.

As an illustration, consider what has been termed “self-serving bias” – the “tendency people have to gather, process, and even remember information in such a way to serve their perceived self-interest and to support their preexisting beliefs” (Prentice 2014, 335). Behavioral ethics has done a great deal to help us understand the ways in which such a bias can arise and affect our behavior. At the same time, one may ask to what extent this knowledge represents a fundamental change in our understanding of human behavior and, in turn, for the practice of normative ethics. The bias involves two claims. The first is that people often act in their perceived self-interest or are motivated by their perceived self-interest. That individuals often act in their self-interest or at least may be tempted to act in their self-interest, however, represents an enduring theme throughout normative ethics. If “the safest general characterization of the European philosophical tradition is that it consists of a series of footnotes to Plato,” as Whitehead quipped (1979, 39), then the theme goes to the beginning of European philosophy with the story of the ring of Gyges.

The second claim embodied in the idea of self-serving bias is that self-interest manifests itself not only in conscious decisions, but also in other aspects of cognition, such as the gathering and processing of information. Again, a concern with the possibility of such self-deception runs

deep throughout the history of normative ethics. It is in direct contradiction to the Socratic maxim, “know thyself.” Adam Smith writes, for example, in *The Theory of Moral Sentiments*:

It is so disagreeable to think ill of ourselves, that we often purposely turn away our view from those circumstances which might render that judgment unfavourable. He is a bold surgeon, they say, whose hand does not tremble when he performs an operation upon his own person; and he is often equally bold who does not hesitate to pull off the mysterious veil of self-delusion which covers from his view the deformities of his own conduct (1759 [1853], 222-223).

None of this is to deny the advances made in our understanding of the self-serving bias. Rather, given the extent to which normative ethics has historically concerned itself with self-interest and self-deception, the question is how much rethinking of the enterprise of normative ethics is required in to incorporate these and other advances.⁶

4. Conclusion

This last point prompts a further question. If it is indeed the case that normative ethics has long entertained questions that are said now to fall within the realm of behavioral ethics, why is it that behavioral ethics is often perceived as presenting a challenge to normative ethics or such a distinct field of inquiry? One answer may be the ongoing march of greater specialization in the academy. If so, then developments in behavioral ethics call not so much for reform in normative ethics, but

⁶ Another way to approach this question is to point out that much of the work in this area (e.g., Prospect Theory) is meant to examine the extent to which individuals conform to a certain idealized form of decision-making (e.g., expected utility theory). While the idealized form may be taken as an approximation of actual decision-making, the point is that it would be a mistake to think that the idealized form is in fact meant to be a description of actual decision-making. This is quite clear if one recalls Leonard Savage’s reaction to the Allais Paradox. Having chosen in a manner that contradicted the tenets of expected utility theory, he wrote that “in general, a person who has tentatively accepted a normative theory must conscientiously study situations in which the theory seems to lead him astray; he must decide for each by reflection – deduction will typically be of little relevance – whether to retain his initial impression of the situation or to accept the implications of the theory for it” (1954, 102). For one discussion of this in the context of Socratic reflection, see Richardson 1994, 29-31.

rather more of a return to a broader conception of the field. On such a view, ethics remains firmly a normative discipline asking, “What is the best way to live?” and “What is right and wrong?” but remains incomplete without empirical investigation into what keeps us from following through.

References

- Akerlof, George and Robert Shiller. 2015. *Phishing for Phools*. Princeton: Princeton University Press.
- Alzola, Miguel. 2012. The Possibility of Virtue. *Business Ethics Quarterly*, 22(2): 377-404.
- Bateson, Melissa, Daniel Nettle, and Gilbert Roberts. 2006. Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, Sep 22; 2(3): 412–414.
- Brownstein, Michael and Jennifer Saul. 2016. *Implicit Bias and Philosophy*, vols. 1 and 2. Oxford: Oxford University Press.
- Clarke, Randolph, Michael McKenna, and Angela Smith, eds. 2015. *The Nature of Moral Responsibility*. Oxford University Press.
- Darley, John and C. Daniel Batson. 1973. “From Jerusalem to Jericho:” A Study of Situational and Dispositional Variables in Helping Behavior. *Journal of Personality and Social Psychology*, 27.1: 100-108.
- Doris, John. 2002. *Lack of Character*. Cambridge: Cambridge University Press.
- Harman, Gilbert. 1999. Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error. *Proceedings of the Aristotelian Society*, 99: 315-331.
- Holroyd Jules, Robin Scaife, and Tom Stafford. 2017. Responsibility for implicit bias. *Philosophy Compass*. 12:e12410. <http://doi.org/10.1111/phc3.12410>
- Kim, Tae Wan, Rosemarie Monge, and Alan Strudler. 2015. Bounded Ethicality and The Principles that “Ought” Implies “Can.” *Business Ethics Quarterly*, 25:3 341-361.
- Miller, Christian. 2017. Character and Situationism: New Directions. *Ethical Theory and Moral Practice*, 20: 459-471.
- Prentice, Robert. 2014. Teaching Behavioral Business Ethics. *Journal of Legal Studies Education*, 31(2): 325–365.
- Richardson, Henry. 1994. *Practical Reasoning about Final Ends*. Cambridge: Cambridge University Press.

- Saul, Jennifer. 2013. Implicit Bias, Stereotype Threat, and Women in Philosophy. In F. Jenkins and K. Hutchinson, eds. *Women in Philosophy: What Needs to Change?* New York: Oxford University Press. 39-60.
- Savage, Leonard. 1954. *The Foundations of Statistics*. John Wiley and Sons.
- Shu, Lisa, Nina Mazar, Francesca Gino, Dan Ariely and Max H. Bazerman. 2012. Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *PNAS* 109(38): 15197- 15200.
- Smith, Adam. 1759 [1853]. *The Theory of Moral Sentiments; or, An Essay towards an Analysis of the Principles by which Men naturally judge concerning the Conduct and Character, first of their Neighbours, and afterwards of themselves*. To which is added, *A Dissertation on the Origins of Languages*. New Edition. With a biographical and critical Memoir of the Author, by Dugald Stewart. London: Henry G. Bohn.
- Thaler, Richard and Cass Sunstein. 2009. *Nudge: Improving Decisions About Health, Wealth and Happiness*, 2d ed. New York: Penguin Books.
- Treviño, Linda. Gary Weaver, and Scott Reynolds. 2006. Behavioral Ethics in Organizations: A Review. *Journal of Management* 32(6): 951-990.
- Tversky, Amos and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science*, New Series, 211(4481): 453-458.
- Whitehead, Alfred North. 1979. *Process and Reality*. New York: Free Press.