

Questions for a science of moral responsibility

Marcelo Fischborn

This paper is published in the *Review of Philosophy and Psychology* (9.2, 2018, pp. 381–394) and is available at Springer via <http://dx.doi.org/10.1007/s13164-017-0360-5>

Abstract: In the last few decades, the literature on moral responsibility has been increasingly populated by scientific studies. Studies in neuroscience and psychology, in particular, have been claimed to be relevant for discussions about moral responsibility in a number of ways. And at the same time, there is not yet a systematic understanding of the sort of questions a science of moral responsibility is supposed to answer. This paper is an attempt to move toward of such an understanding. I discuss three models for framing scientific questions relevant to an investigation of moral responsibility. The favored model—the Enhancement model—proposes that a science of moral responsibility has two descriptive tasks. First, science can describe the causes and effects of the many sorts of responses that constitute the human practices of moral responsibility, such as praise, blame, and punishment. And, second, science can describe how modifications aiming at the improvement of such practices can be achieved. Relatively to the other models to be considered, the Enhancement model is broader in scope and less tied to the traditional philosophical agenda on moral responsibility.

Keywords: moral responsibility; free will; science; psychology; neuroscience

1 Introduction

Among the questions that science can in principle answer, which ones matter for an investigation of moral responsibility? Despite the increase in scientific studies related to moral responsibility and free will, little effort has been made to systematically define the questions for such an investigation. This paper discusses three general models that can be used to frame answers to the question posed, one of which is favored in the end. And while these answers are mainly a matter for philosophical debate, I hope the reflection can help those interested in the scientific study of moral responsibility think more explicitly about the questions they are—or could be—trying to answer.

The paper is structured as follows. Section 2 focuses on what I call the Minimal model, which gives science a relatively small set of questions about moral responsibility. Sec-

tion 3 discusses a model that arises in the context of the experimental philosophy movement, the Folk intuitions model. Section 4 presents my favored model, namely, the Enhancement model. Relatively to the previous models, the Enhancement model innovates by integrating a broader set of scientific questions about moral responsibility, and by giving science not just purely descriptive questions (e.g., what does causally affect responsibility ascriptions?) but also questions that pertain to a more revisionist enterprise (e.g., how can morally objectionable aspects of responsibility ascriptions be modified?).

2 The Minimal model

Among the models for framing the scientific questions that matter for an investigation of moral responsibility to be considered here, the Minimal model is (although tacitly) the most widely assumed in contemporary philosophical discussion. According to this model, science is supposed to contribute with the assessment of the empirical conditions postulated by different views about the nature of moral responsibility. By telling whether such conditions are ever satisfied, science helps to tell whether there are any morally responsible human beings, i.e., beings that satisfy conditions that are necessary (or sufficient) for moral responsibility. I call this model ‘minimal’ because it denies science any further role in the investigation of moral responsibility. In particular, it denies that science can help to determine what the conditions of moral responsibility are.

Robert Kane’s (1996) work on free will and moral responsibility is a clear example of the Minimal model at work. Kane is an incompatibilist about free will and determinism, and as such he assumes that the falsity of determinism is a necessary condition for the existence of free and morally responsible agents.¹ Kane defends incompatibilism on the basis of traditional philosophical arguments, but he turns for science when the question is whether determinism is actually false:

There are empirical aspects of the free will issue that mere philosophical speculation cannot co-opt. If free will of a nondeterminist kind should exist in nature, then the atoms must somewhere ‘swerve’ to make room for it, and they must swerve in places where it matters—in the brain, for example. (Kane, 1996, p. 17)

The investigation about the existence of such undetermined events, according to Kane, turns on questions about the physical world in general—and our brains, in particular. Kane’s project, therefore, assumes a conception about the role of science in the investigation of moral

1 Determinism can be understood as the thesis that the complete state of the world at a given time and the natural laws that are true in this world fix the state of the world at any subsequent time (Hofer, 2016).

responsibility that is in accordance with the Minimal model as previously defined.

Compatibilist theories—those that take free will and moral responsibility to be possible even if determinism is true—may also have empirically assessable commitments. John Martin Fischer and Mark Ravizza (1998), for example, argue that responsiveness to reasons (a capacity to use reasons to guide one’s behavior) is a necessary condition of moral responsibility. Like in Kane’s work, the condition is justified using philosophical arguments. But the claim that human beings are responsive to reasons as required is itself amenable to scientific scrutiny. For example, some other philosophers have examined whether the situationist literature—empirical results showing that minor aspects of the environment can influence our behavior in usually unnoticed ways—undermines reasons responsiveness (see, e.g., Nelkin, 2005; Schlosser, 2013; Vargas, 2013; Shepherd, 2015). This is another instance of the Minimal model at work.

Further examples arise in discussions about the impact of the so-called Libet-style experiments on a number of proposed conditions for free will and moral responsibility. Libet et al. (1983) found a pattern of neural activity called ‘readiness potential’ (RP) to precede simple spontaneous movements such as flexing a finger or a wrist, and also the moment subjects reported to have consciously felt the urge to execute those movements. Recent studies in this paradigm were able to use patterns of neural activity to predict which button (right or left) a subject would press approximately 7 seconds before actual movement, although with a relatively low accuracy (Soon et al., 2008; Haynes, 2011). It has been widely discussed whether and how these results do or could impact the satisfaction of a number of proposed conditions for the existence of free will and moral responsibility. For example, do the results show that decisions are deterministically caused by neural events (Roskies, 2006, 2011, 2014; Balaguer, 2009; Haynes, 2011; Misirlisoy & Haggard, 2014; Nahmias, 2014; Fischborn, 2016, 2017; Roskies & Nahmias, 2017)? Do they show that decisions, intentions, or conscious mental states in general do not play the role they are supposed to do in the generation of actions (Nahmias, 2002; Wegner, 2002; Pockett, Banks & Gallagher, 2006; Gomes, 2007; Mele, 2009; Baumeister, Masicampo & Vohs, 2011; Schlosser, 2012; Marques, 2015; Asma, forthcoming)? These questions, once again, share the assumption that science can help to tell whether the conditions of moral responsibility are ever satisfied.

The examples considered above suggest that the Minimal model has been very influential in guiding the incorporation of scientific results into the philosophical literature on free

will and moral responsibility. One of the model's main strengths is that it transfers to science those questions that cannot be answered by philosophy alone.² But as I will argue in the following sections, the Minimal model fails to include a number of scientific questions that may also be relevant in an investigation of moral responsibility. In other words, even if it is granted that the Minimal model covers relevant and legitimate scientific *questions*, it may still fail as a *framework* for the investigation of moral responsibility as a whole.³ The next section considers a second model that covers some additional questions.

3 The Folk intuitions model

According to the Minimal model, the only goal of a science of moral responsibility is to help to determine whether the conditions of moral responsibility are ever satisfied. A different model arises within the experimental philosophy movement. Because the term 'experimental philosophy' itself is associated with a number of different projects (Alfano & Loeb, 2016, section 2), I reserve the term 'Folk intuitions model' to refer to one of them in particular. In this section I describe the envisaged model and argue that it also fails to capture the full range of questions that matter for a science of moral responsibility.

The methodology of experimental philosophy was first applied to questions related to moral responsibility in an attempt to advance the traditional dispute among compatibilists and incompatibilists (Nahmias et al., 2005). As I said earlier, this is a dispute about what the conditions of moral responsibility are—incompatibilists affirm, while compatibilists deny, that the falsity of determinism is among such conditions. Nahmias et al. (2005) noted that "so many philosophers [compatibilists and incompatibilists alike] claim that their own position has the most intuitive appeal and best fits our ordinary conception of free will and our practices of responsibility attribution" (p. 563). Because compatibilists and incompatibilists cannot be both right in their claims, Nahmias et al. proposed that surveying non-experts' actual ascriptions of responsibility when exposed to relevant thought experiments could help to better assess the

2 It should be noted that the Minimal model is consistent with the possibility that *none* of the conditions for the existence of moral responsibility can be assessed scientifically. A defense of this thesis would involve showing that science cannot help to assess the truth of any of the suggested conditions of moral responsibility. Although this position has been defended regarding *some* of the conditions discussed—for example, observations have been said to be insufficient to justify a choice between stochastic and deterministic models in some cases (Suppes, 1993, but see Werndl, 2013, 2016)—an argument would still be needed to show that the same holds for all of them.

3 Similar remarks apply to the other models to be discussed below. Accordingly, while there is no conflict among the scientific questions that arise within each of the models, there may be conflict among the integral views about the investigation of moral responsibility the models assume.

intuitive appeal of each view. The results—the authors suggested—would help not to solve the old philosophical disputes, but to better assign the burden of proof, in such a way that more intuitive views could start off with an advantage or “squatters’ rights” (p. 564). In addition, the results could help to assess how revisionary different views about free will and moral responsibility are (pp. 564-565).

Nahmias et al.’s (2005) own results suggested that most people are willing to attribute free will and moral responsibility to agents in deterministic scenarios. Nichols and Knobe (2007) presented the first evidence that most people deny that agents in deterministic scenarios are morally responsible, at least if the scenario is described in more abstract terms. The debate has remained inconclusive since then. While some results and interpretations support the claim that most people are naturally compatibilists (Nahmias, Coates & Kvaran, 2007; Murray & Nahmias, 2014; Andow & Cova, 2016; Monroe, Brady & Malle, 2017), others support the claim of a natural incompatibilism (Rose & Nichols, 2013; Feltz, 2015; Bear & Knobe, 2016), and some other results are mixed (Roskies & Nichols, 2008; Nadelhoffer et al., 2014). As things stand now, the question remains open (Schooler et al., 2015) and some have suggested that it may not be even worth looking for a unified ordinary view in this domain (Feltz, Cokely & Nelson, 2016).

I take the research program just reviewed to instantiate what I call the Folk intuitions model. The key assumption in this model is that science can help to determine the conditions for the existence of morally responsible agents. The Folk intuitions model, therefore, rejects the Minimal model’s assumption that science can only help to tell which of the conditions for moral responsibility are actually satisfied. One of the strengths of the Folk intuitions model is that it helps to keep philosophical theories close to those human activities they are supposed to be about. Alfred Mele claimed that a risk for the philosophical discussion of free will is to have “nothing more than a philosophical fiction as its subject matter” (Mele, 2001, p. 27). The Folk intuitions model can help to avoid the risk by providing a more accurate portrait of folk concepts and judgments. The Folk intuitions model, moreover, can also incorporate some of the scientific questions assigned to science within the Minimal model, even though the models themselves diverge. In this enlarged version, science contributes with both the specification of the conditions of moral responsibility and the assessment of their satisfaction. However, the Minimal model and the Folk intuitions model (even in its enlarged version) still fail to give science the full role it can have in an investigation of moral responsibility. Both mod-

els arise out of a philosophical agenda which does not (and does not have to) represent all of the scientific questions about moral responsibility we may be interested in answering. In the next section, I describe some of the questions the Minimal model and the Folk intuitions model leave behind. These further questions set the stage for developing an alternative view about the science of moral responsibility.

4 The Enhancement model

Although the models so far considered propose legitimate scientific questions about moral responsibility, they are far from complete. At best, they give a complete picture of the scientific questions that are relevant for the kind of philosophical work on moral responsibility that has been predominant. In this section I describe an alternative model—the Enhancement model—which gives a unified shape to an investigation aiming at the understanding and improvement of the social practices associated with moral responsibility. I begin by describing the general features of the Enhancement model, and then I illustrate how it integrates a broader variety of results, some of which have not been recognized so far as pertaining to a broader and shared enterprise on moral responsibility.

It was implicit in the discussion thus far that a central objective in an investigation of moral responsibility is to specify under what conditions it is acceptable or appropriate to respond to an agent in certain ways. Of course it is open to discussion what exactly ‘appropriate’ means in this context, as well as what the intended responses are (see, e.g., Rosen, 2015; Zimmerman, 2015), but the same general objective could be set for various candidates. The Enhancement model proposes that pursuing the general objective involves at least three broad types of tasks—namely, conceptual, descriptive, and normative tasks—and that science’s primary role is to execute the descriptive tasks. Within the Enhancement model, however, this descriptive task goes beyond the description of responsibility practices as they currently stand. The model leaves open the possibility that a normative assessment can find current practices problematic and in need of revision. In such cases, science has the further task of describing the available routes for change.

An example of a conceptual task is the definition of terms used in the investigation of moral responsibility. This is a traditional business of philosophers, and key questions include what it is to be an agent, what responsibility involves, what it is to blame, punish, or praise someone, and so on. Conceptual assumptions are involved in the very definition of the central

objective of the investigation of moral responsibility mentioned in the previous paragraph. Once this initial conceptualization of key notions is available (even if provisional), the Enhancement model gives science its first task, which is to describe how responsibility practices actually operate. This description focuses on what I will call ‘responsibility responses’, i.e., the kinds of responses or reactions that are characteristic of moral responsibility, which include praise, blame, and punishment as paradigmatic cases. And the description should aim at providing us with a comprehensive understanding of the causes and effects of those responses. I take this descriptive enterprise to be a task for a plurality of scientific disciplines and fields—certainly including many fields of psychology, neuroscience, and the social sciences—and for a plurality of methodological approaches within those disciplines and fields.

A second task science gets within the Enhancement model is conditional on a normative assessment of the workings of ordinary responsibility practices as revealed in the first task. The Enhancement model assumes that any aspect of actual responsibility responses may be found faulty upon normative scrutiny. Like the descriptive tasks assigned to science, this normative assessment is not supposed to be restricted to a single field or approach—potential candidates include, but are not limited to ethical, social, political, and legal normative theories. For the purposes of this paper, the most important consequence of this normative assessment is that it generates a further potential role for science. If ordinary responsibility practices turn out to have any problematic aspects, then it is desirable to find ways for improvement. Therefore, the second task science gets within the Enhancement model is to describe the possible routes for effectively changing those problematic aspects. And, because interventions are likely to also involve potential risks or costs, the choice to pursue change in any of the available ways is likely to involve further normative assessments.⁴

In sum, the Enhancement model gives the science of moral responsibility two main tasks. The first task is to describe the causes and effects of responsibility responses. And the second task is to describe possible ways for changing problematic aspects associated with the workings of those responses. In the remainder of this section I illustrate how these tasks can be and have already been carried on. But I emphasize from the start that although the philo-

4 The influence of values is often considered in discussions about the objectivity of science. The Enhancement model allows for research questions and applications of scientific results to be affected by those values assumed in the normative assessment of responsibility practices. Most philosophers, however, do not take this sort of interference to be a threat to scientific objectivity. As Fischer et al. (2007, sec. 3.1) note, “[t]he real debate is about whether or not the ‘core’ of scientific reasoning—the gathering of evidence and the assessment and acceptance of scientific theories—is, and should be, value-free.”.

sophical and scientific tasks (and the two scientific tasks themselves) can be clearly distinguished for the purposes of reflection, they are often mixed in actual investigation. In illustrating how the proposed tasks can be realized, I also hope to show how the Enhancement model integrates within the science of moral responsibility a number of studies that so far have been largely ignored in more philosophically-oriented work on the topic.

The Path model of blame developed by Malle, Guglielmo and Monroe (2014) exemplifies what the first descriptive task can look like. This model hypothesizes that ordinary attributions of blame are guided by a process that involves the evaluation of certain conditions. First, blame only emerges if an agent is judged to have caused an event or outcome that violates a norm (2014, p. 151). If this condition is met, then a second step involves judging “whether the agent brought about the event intentionally” (p. 151). Once intentionality is assessed, two routes may follow. If the event was intentionally brought about, then the amount of blame depends on whether the agent had good reasons for the action, resulting in “minimal blame if the agent was justified in acting this way; maximal blame if the agent was not justified” (p. 151). If no intentionality is detected, then the amount of blame is guided by an assessment of the agent’s obligations and capacity to prevent the undesirable event, in such a way that the existence of such obligation and capacity results in more blame, and their absence results in low or no blame (see 2014, p. 151, figure 2).

The Path model of blame contrasts with the kind of scientific result that is emphasized within the Minimal and the Folk intuitions models. Granted, *philosophical* theories of moral responsibility propose conditions under which blame is appropriate, many of which are roughly in accordance with the Path model of blame. But the scientific questions within the Minimal model concern whether or not those conditions are satisfied in most or in special cases, regardless of whether considerations about those conditions are causally operative or not in ordinary blame ascriptions. Similarly, even though the Folk intuitions model assumes that proposed conditions of moral responsibility should approximate ordinary ascriptions of blame, the questions assigned to science have been largely concerned with solving controversial philosophical disputes about what is intuitive (as in the compatibilism versus incompatibilism issue), and not with providing a comprehensive explanation of blame ascriptions. The Enhancement model emphasizes the goal of describing the causes and effects of ordinary responsibility reactions in a comprehensive and accurate way whether or not such causes and effects capture what has been emphasized in more traditional philosophical theories.

Similar empirical questions can be raised about any other kind of response associated with moral responsibility. Consider punishment. Throughout adult life, the primary use of punishment is supposed to be regulated by criminal justice systems. Despite this fact, there is evidence that external factors—such as the gender, race, and age of punished individuals—influence punishment patterns (Steffensmeier, Ulmer & Kramer, 1998; Spohn & Holleran, 2000). Different forms of punishment are also often used by parents in the process of rearing their children. Regarding *physical* punishment, in particular, studies indicate that predictive factors of its use include the “perception of the seriousness and intent of the child misbehavior” (Ateah & Durrant, 2005). On the positive side of responsibility responses, gratitude, for example, has been shown to arise when one has benefited from someone else’s intentional behavior (McCullough, Kimeldorf & Cohen, 2008). The Enhancement model takes all these studies as contributions to the task of describing the workings of ordinary responsibility responses.

Now, as noted earlier, the causes and effects of responsibility responses may not coincide with the conditions for moral responsibility postulated by philosophical theories, nor need all of them be accepted as desirable. In more general terms, we may have normative reasons to think that the way responsibility responses operate needs modification. And this possibility sets the stage for the second task assigned to science within the Enhancement model, namely, the investigation about how to effect changes in problematic aspects of responsibility practices.

The Path model of blame predicts that blame arises when someone violates a norm intentionally. But, under certain circumstances, it seems disputable that blaming responses arising in this way are acceptable, much less recommended. Hanna Pickard (2011; 2013), for instance, has developed a notion of responsibility without blame. In certain therapeutic settings, she argues, blame is counterproductive for recovery even though capacities that are required for moral responsibility are significantly operative. But it is beneficial for treatment that patients see themselves as responsible for their actions. How then can responsibility and blame be separated? Pickard’s proposal is that therapists need to withhold the expression of negative emotions and judgments that give blame its characteristic ‘sting’. Instead, they should appeal to what she calls ‘detached blame’: “Detached blame consists in judgments of blameworthiness, and may further involve correspondingly appropriate revisions of intentions, the imposition of negative consequences, and accountability and answerability” (Pickard, 2013, p.

1146). But this is to be distinguished from an ‘affective’ form of blame, which characteristically involves the blamer feeling herself entitled to express certain negative reactive attitudes and emotions. The affective form of blame, she says, is counterproductive for treatment and active steps should be taken to prevent its manifestation in therapeutic contexts. Pickard’s proposal is in accordance with what I call a normative assessment of responsibility practices as well as with the scientific description of alternatives. Of course, it can be replied that therapeutic contexts are too peculiar to imply any general verdict about responsibility norms, and so it may be worth looking at more common contexts as well.

Research on relationships and conflict resolution provides a more ordinary case for reflection. Counselors working in schools are sometimes advised to seek conflict resolution in ways that make students stop “from blaming one another or pointing fingers at whose fault it is” (Brinson, Kottler & Fisher, 2004, p. 300). In workplace settings, blaming was found positively associated with seeking revenge and negatively associated with reconciliation (Aquino, Tripp & Bies, 2001). And marital satisfaction can be negatively affected when spouses keep blaming one another for their problems and faults (Madden & Janoff-Bulman, 1981; Kubany et al., 1995). Malle, Guglielmo and Monroe (2014, pp. 173-174) themselves also describe a “darker side” of blame, which includes blame expressed with exaggerated emotional intensity, blame followed by complaints and countercomplaints that increase the odds of conflict, and situations where the value of repairing a relationship is missing, or a powerful social group creates “scapegoats” to be blamed for the larger community’s problems. The physical punishment of children raises similar concerns. The practice has become increasingly condemned world-wide due to its long-lasting effects on children’s behavior, relationships, and mental health (Ateah & Durrant, 2005; Afifi et al., 2006; Lansford et al., 2007; Teicher, 2010; Durrant & Ensom, 2012; Initiative, 2015; Kish & Newcombe, 2015).

Rather than making specific judgments about the acceptability of blame or punishment in such a variety of contexts, my point is simply that studies like the ones just mentioned should all be seen as part of a comprehensive science of moral responsibility. Science can provide us with an accurate view of current responsibility practices and guidance about how to avoid problematic aspects. These are both essential tasks for the science of moral responsibility as defined by the Enhancement model.

While in the cases just considered the suggestion is that science can help to improve responsibility practices by finding ways to *prevent* certain responsibility responses, the prac-

tices can also be improved by having other responses *promoted*. This possibility arises from the fact that some factors can influence responsibility responses in an inhibitory way. Some studies about violence against women exemplify this situation.

Since at least the late 1980s, researchers have been aware of a serious problem of sexual violence against women in the university context. In a study in the United States by Koss, Gidycz and Wisniewski (1987), 28% of women in college reported being a victim of a crime satisfying the legal definition of rape since the age of fourteen. The problem was neither restricted to the higher education context, nor to the United States. In a population-based survey in the United States, 18% of women reported having been raped, and 45% reported having been victims of other forms of sexual violence (Black et al., 2011). And the problem has been documented world-wide (WHO, 2013).

From the standpoint of the Enhancement model, it matters to note that proposed strategies to reduce those sorts of violence have included a promotion of responsibility responses. Black et al. (2011, p. 4) say that “[a]n important part of any response [...] is to hold perpetrators accountable”. Responsibility responses are also promoted within so-called ‘bystander approaches’ to the prevention of sexual violence:

Bystander models share a literature that provides guidance on which factors increase the likelihood that a bystander will intervene to prevent violence. Briefly, the objective of bystander intervention is to involve both men and women to change the context or environment that may tacitly support violence against women. (Coker et al., 2011, p. 779)

The Green Dot Intervention Program is one such bystander program that was found effective in promoting bystander interventions (Coker et al., 2011). In assessing the program, Coker et al. used a modified version of the Bystander Behaviors Scale (Banyard, Plante & Moynihan, 2005), in which participants were asked to inform things such as how often they “spoke up if somebody said that someone deserved to be raped or to be hit by their partner” (pp. 785-786) and how often they “spoke up to someone who was bragging or making excuses for forcing someone to have sex with them” (p. 786). These ‘speaking-ups’ clearly express disapprobation and censure, and as such, they count as expressions of blame. Thus, while in the cases earlier considered responsibility practices could be enhanced by having certain responses prevented, in the case of sexual violence the practices can be enhanced by having some responses promoted.

The Enhancement model, therefore, improves on the Minimal model and the Folk intuitions model by giving the science of moral responsibility a set of questions that is broader

in scope and less tied to the traditional philosophical agenda. As the Folk intuitions model is intended to do, the Enhancement model also keeps a tight connection with real-world practices. But it takes the practices as they currently stand as just a starting point for an investigation whose ultimate goal is their improvement. After all, actual practices can be suboptimal in a number of ways, in which case we would all benefit from acknowledging the imperfections and finding ways to fix them. Hence, according to the Enhancement model science should describe the current state of the practices as well as the available ways for pursuing their improvement.

5 Some related views

The many scientific studies mentioned in the previous section instantiate different aspects of the investigation of moral responsibility described by the Enhancement model. In this respect, the Enhancement model is just a framework that integrates a variety of available studies, making explicit their underlying assumptions and relations. This section describes some views that approximate aspects of the Enhancement model as such, i.e., as an explicit framework for the investigation of moral responsibility that integrates descriptive and evaluative work with the aim of improving responsibility practices.

The approach to moral responsibility originally put forth by P.F. Strawson (1962, p. 152) and further developed by Adina Roskies and Bertram Malle (2013) has a lot in common with the Enhancement model. One of the insights in Strawson's landmark essay is the idea that we should look at how responsibility practices figure in ordinary interpersonal relations in order to establish what the conditions for moral responsibility are.⁵ Although Strawson sought to describe the practices based on armchair considerations, Roskies and Malle (2013) updated the method by turning to empirical studies, including an earlier version of the Path model of blame (Guglielmo, Monroe & Malle, 2009; Malle, Guglielmo & Monroe, 2012). In this first aspect, Roskies and Malle's approach is very similar to the first descriptive task assigned to science within the Enhancement model. Roskies and Malle also mention that ordinary responsibility practices could be "measured against deeper values such as fairness and justice (and [that] perhaps some cultures have more just practices of blame and punishment than others)" (2013, p. 147). This is close to the Enhancement model's proposed normative assessment of

5 Although it is open to debate how exactly Strawson's point is to be understood, it seems to involve the assumption that people are responsible because they are *held* responsible in interpersonal relations, and not the other way around—hence the claim that the approach involves some sort of 'reversal' (see Todd, 2016, for an in-depth discussion of the point).

responsibility practices, although Roskies and Malle do not describe the search for improvement that such an assessment can motivate neither the role science can have in it. On the other hand, when briefly discussing certain studies in psychology, Strawson did acknowledge “the possibility and desirability of redirection and modification of our human attitudes in the light of these studies” (1962, p. 170). This largely ignored aspect of Strawson’s contribution approximates the Enhancement model’s evaluative and descriptive tasks directed toward the improvement of ordinary practices, even though Strawson himself did not distinguish between the normative and scientific elements involved. Finally, Strawson as well as Roskies and Malle were very interested in assessing the relevance of determinism in ordinary practices, and this is a point where both views depart from the Enhancement model. For within the Enhancement model, determinism may become relevant only if it is explanatorily relevant—for example, if beliefs about determinism play a causal role in current responsibility practices—or if determinism is thought to be both true and normatively relevant for an assessment of the practices. If none of these conditions is met, then the topic of determinism is at best secondary within the Enhancement model.

Hagop Sarkissian’s (2010) work on the situationist literature also has elements in common with the Enhancement model. As mentioned earlier (section 2), research in social psychology provides evidence that seemingly minor environmental factors can have a major influence on one’s subsequent moral behavior. Although the situationist literature may seem troubling, Sarkissian argues that it also opens the possibility that minor changes in our actions and attitudes—such as “choice of words, emotional expressions, mannerisms, tone of voice, posture” (2010, p. 10)—can positively impact the situations we find ourselves in: “By proactively introducing signals that foster an environment amenable to cooperation, one can enhance the probabilities of positive outcomes emerging.” (p. 10). This concern for improving the moral quality of human interactions is something the Enhancement model shares with Sarkissian’s approach. But there are also differences in scope. First, Sarkissian focuses on improving the outcomes of human interactions in general, while the Enhancement model is particularly concerned with interactions that involve responsibility responses. And, second, Sarkissian is specifically concerned with the impact of “minor tweaks” in our behavior and attitudes, while the Enhancement model leaves open the type of adjustment to be considered.

Finally, it may be instructive to distinguish between the Enhancement model and Manuel Vargas’ (2007; 2013) revisionist approach to moral responsibility. Vargas assumes that

most people take responsibility practices to rely on incompatibilist commitments (2013, p. 40) but contends that libertarian theories assuming the existence of indeterministic processes at “particular places and times along the pathway to human decisions” (p. 67) fail to meet a standard of naturalistic plausibility. Vargas thinks this situation justifies not skepticism—the view that no one is morally responsible—but revisionism, a “repair to commonsense thinking, one that strips away the metaphysically demanding elements of libertarianism and preserves the justifiable core of our attitudes and practices.” (2007, p. 160). The Enhancement model shares Vargas’ idea that responsibility practices can be justified in part because they help to cultivate a valuable form of agency. But the two views have different targets when it comes to their proposed revisions. Vargas’ approach is mainly concerned with the revision of ordinary commitments or beliefs (particularly the incompatibilist ones). The Enhancement model, in contrast, focuses on possible revisions to the very responsibility responses that may or may not depend on those beliefs.

6 Conclusion

This paper has considered three models that provide general frameworks for specifying the scientific questions that matter for an investigation of moral responsibility. While the Minimal model only gives science the task of telling whether specific conditions of moral responsibility are satisfied, the Folk intuitions model gives science a role in the very specification of those conditions. Both models, however, are incomplete and leave out further relevant questions about moral responsibility. The Enhancement model gives voice to a group of such additional questions, one that arises out of an interest to understand and improve the actual workings of responsibility responses. Within this framework, science gets the tasks of explaining responsibility responses and describing how normatively problematic aspects of the responses can be effectively changed. The Enhancement model, therefore, gives the science of moral responsibility both purely descriptive tasks and tasks that pertain to a more revisionist enterprise whose ultimate goal is the improvement of responsibility practices.

Acknowledgments

I would like to thank Frank Sautter, Adam Bear, and two anonymous reviewers for this journal for their comments on earlier versions of the paper. Thanks also to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support.

References

- Afifi, T. O.; Brownridge, D. A.; Cox, B. J. and Sareen, J. (2006). "Physical punishment, child abuse and psychiatric disorders", *Child Abuse and Neglect* 30: 1093-1103.
- Alfano, M. and Loeb, D. (2016). "Experimental moral philosophy". In: Zalta, E. N. (Ed.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University.
- Andow, J. and Cova, F. (2016). "Why compatibilist intuitions are not mistaken: A reply to Feltz and Millan", *Philosophical Psychology* 29: 550-566.
- Aquino, K.; Tripp, T. M. and Bies, R. J. (2001). "How employees respond to personal offense: The effects of blame attribution, victim status, and offender status on revenge and reconciliation in the workplace", *Journal of Applied Psychology* 86: 52-59.
- Asma, L. (forthcoming). "There is no free won't: The role definitions play", *Journal of Consciousness Studies* .
- Ateah, C. A. and Durrant, J. E. (2005). "Maternal use of physical punishment in response to child misbehavior: Implications for child abuse prevention", *Child Abuse & Neglect* 29: 169-185.
- Balaguer, M. (2009). "Why there are no good arguments for any interesting version of determinism", *Synthese* 168: 1-21.
- Banyard, V. L.; Plante, E. G. and Moynihan, M. M. (2005). *Rape prevention through bystander education: Bringing a broader community perspective to sexual violence prevention*. Washington, DC: Report submitted to the U.S. Department of Justice.
- Baumeister, R. F.; Masicampo, E. J. and Vohs, K. D. (2011). "Do conscious thoughts cause behavior?", *Annual Review of Psychology* 62: 331-361.
- Bear, A. and Knobe, J. (2016). "What do people find incompatible with causal determinism?", *Cognitive Science* 40: 2025-2049.
- Black, M. C.; Basile, K. C.; Breiding, M. J.; Smith, S. G.; Walters, M. L.; Merrick, M. T.; Chen, J. and Stevens, M. R. (2011). *The National Intimate Partner and Sexual Violence Survey: 2010 summary report*. Atlanta: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.
- Brinson, J. A.; Kottler, J. A. and Fisher, T. A. (2004). "Cross-cultural conflict resolution in the schools: Some practical intervention strategies for counselors", *Journal of Counseling & Development* 82: 294-301.
- Coker, A. L.; Cook-Craig, P. G.; Williams, C. M.; Fisher, B. S.; Clear, E. R.; Garcia, L. S. and Hegge, L. M. (2011). "Evaluation of Green Dot: An active bystander intervention to reduce sexual violence on college campuses", *Violence Against Women* 17: 777-796.
- Durrant, J. and Ensom, R. (2012). "Physical punishment of children: lessons from 20 years of research", *Canadian Medical Association Journal* 184: 1373-1377.
- Feltz, A. (2015). "Experimental philosophy of actual and counterfactual free will intuitions", *Consciousness and Cognition* 36: 113-130.

- Feltz, A.; Cokely, E. T. and Nelson, B. (2016). "Experimental philosophy needs to matter: Reply to Andow and Cova", *Philosophical Psychology* 29: 567-569.
- Fischborn, M. (2016). "Libet-style experiments, neuroscience, and libertarian free will", *Philosophical Psychology* 29: 494-502.
- Fischborn, M. (2017). "Neuroscience and the possibility of locally determined choices: Reply to Adina Roskies and Eddy Nahmias", *Philosophical Psychology* 30: 198-201.
- Fischer, J. M.; Kane, R.; Pereboom, D. & Vargas, M. (Ed.) (2007). *Four views on free will*. Malden: Wiley-Blackwell.
- Fischer, J. M. and Ravizza, M. (1998). *Responsibility and control: A theory of moral responsibility*. Cambridge: Cambridge University Press.
- Gomes, G. (2007). "Free will, the self, and the brain", *Behavioral Sciences and the Law* 25: 221-234.
- Guglielmo, S.; Monroe, A. E. and Malle, B. F. (2009). "At the heart of morality lies folk psychology", *Inquiry* 52: 449-466.
- Haynes, J.-D. (2011). "Beyond Libet: Long-term prediction of free choices from neuroimaging signals". In: Sinnott-Armstrong, W. & Nadel, L. (Ed.), *Conscious will and responsibility: A tribute to Benjamin Libet*, pp. 85-96. Oxford: Oxford University Press.
- Hofer, C. (2016). "Causal determinism". In: Zalta, E. N. (Ed.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University.
- Initiative (2015). "Global initiative to end all corporal punishment of children", <http://www.endcorporalpunishment.org/>.
- Kane, R. (1996). *The significance of free will*. Oxford: Oxford University Press.
- Kish, A. M. and Newcombe, P. A. (2015). "'Smacking never hurt me!' Identifying myths surrounding the use of corporal punishment", *Personality and Individual Differences* 87: 121-129.
- Koss, M. P.; Gidycz, C. A. and Wisniewski, N. (1987). "The scope of rape: Incidence and prevalence of sexual aggression and victimization in a national sample of higher education students", *Journal of Counseling and Clinical Psychology* 55: 162-170.
- Kubany, E. S.; Bauer, C. B.; Muraoka, M. Y.; Richard, D. C. and Read, P. (1995). "Impact of labeled anger and blame in intimate relationships", *Journal of Social and Clinical Psychology* 14: 53-60.
- Lansford, J. E.; Miller-Johnson, S.; Berlin, L. J.; Dodge, K. A.; Bates, J. E. and Pettit, G. S. (2007). "Early physical abuse and later violent delinquency: A prospective longitudinal study", *Child Maltreatment* 12: 233-245.
- Libet, B.; Gleason, C. A.; Wright, E. W. and Pearl, D. K. (1983). "Time of conscious intention to act in relation to onset of cerebral activity (readiness potential): The unconscious initiation of a freely voluntary act", *Brain* 106: 623-642.
- Madden, M. E. and Janoff-Bulman, R. (1981). "Blame, control, and marital satisfaction: Wives' attributions for conflict in marriage", *Journal of Marriage and Family* 43: 663-674.

- Malle, B. F.; Guglielmo, S. and Monroe, A. E. (2012). "Moral, cognitive, and social: The nature of blame". In: Forgas, J.; Fiedler, K. & Sedikides, C. (Ed.), *Social thinking and interpersonal behaviour*, pp. 313-331. New York: Psychology Press.
- Malle, B. F.; Guglielmo, S. and Monroe, A. E. (2014). "A theory of blame", *Psychological Inquiry* 25: 147-186.
- Marques, B. S. (2015). "Different kinds of decisions and an experiment on unconscious generation of free decisions: A conceptual analysis", *Filosofia Unisinos* 16: 44-57.
- McCullough, M. E.; Kimeldorf, M. B. and Cohen, A. D. (2008). "An adaptation for altruism? The social causes, social effects, and social evolution of gratitude", *Current Directions in Psychological Science* 17: 281-285.
- Mele, A. (2001). "Acting intentionally: Probing folk notions". In: Malle, B.; Moses, L. & Baldwin, D. (Ed.), *Intentions and intentionality: Foundations of social cognition*, pp. 27-43. Cambridge, MA: MIT Press.
- Mele, A. (2009). *Effective intentions: The power of conscious will*. Oxford: Oxford University Press.
- Misirlisoy, E. and Haggard, P. (2014). "A neuroscientific account of the human will". In: Sinnott-Armstrong, W. (Ed.), *Moral psychology: Free will and moral responsibility*, pp. 37-42. Cambridge, MA: MIT Press.
- Monroe, A. E.; Brady, G. and Malle, B. F. (2017). "This isn't the free will worth looking for: General free will beliefs do not influence moral judgments; agent-specific choice ascriptions do", *Social Psychological and Personality Science* 8: 191-199.
- Murray, D. and Nahmias, E. (2014). "Explaining away incompatibilist intuitions", *Philosophy and Phenomenological Research* 88: 434-467.
- Nadelhoffer, T.; Shepard, J.; Nahmias, E.; Sripada, C. and Ross, L. T. (2014). "The free will inventory: Measuring beliefs about agency and responsibility", *Consciousness and Cognition* 25: 27-41.
- Nahmias, E. (2002). "When consciousness matters: A critical review of Daniel Wegner's *The illusion of conscious will*", *Philosophical Psychology* 15: 527-541.
- Nahmias, E. (2014). "Is free will an illusion? Confronting challenges from the modern mind sciences". In: Sinnott-Armstrong, W. (Ed.), *Moral psychology: Free will and moral responsibility*, pp. 1-25. Cambridge, MA: MIT Press.
- Nahmias, E.; Coates, D. J. and Kvaran, T. (2007). "Free will, moral responsibility, and mechanism: Experiments on folk intuitions", *Midwest Studies in Philosophy* 31: 214-242.
- Nahmias, E.; Morris, S.; Nadelhoffer, T. and Turner, J. (2005). "Surveying freedom: Folk intuitions about free will and moral responsibility", *Philosophical Psychology* 18: 561-584.
- Nelkin, D. K. (2005). "Freedom, responsibility and the challenge of situationism", *Midwest Studies in Philosophy* 29: 181-206.
- Nichols, S. and Knobe, J. (2007). "Moral responsibility and determinism: The cognitive science of folk intuitions", *Noûs* 41: 663-685.
- Pickard, H. (2011). "Responsibility without blame: Empathy and the effective treatment of

- personality disorder”, *Philosophy, Psychiatry, Psychology* 18: 209-223.
- Pickard, H. (2013). “Responsibility without blame: Philosophical reflections on clinical practice”. In: Fulford, K. W. M.; Davis, M.; Gipps, R. G. T.; Graham, G.; Sadler, J. Z.; Stanghellini, G. & Thornton, T. (Ed.), *The Oxford handbook of philosophy and psychiatry*, pp. 1134-1154. Oxford: Oxford University Press.
- Pockett, S.; Banks, W. & Gallagher, S. (Ed.) (2006). *Does consciousness cause behavior?*. Cambridge, MA: MIT Press.
- Rose, D. and Nichols, S. (2013). “The lesson of bypassing”, *Review of Philosophy and Psychology* 4: 599-619.
- Rosen, G. (2015). “The alethic conception of moral responsibility”. In: Clarke, R.; McKenna, M. & Smith, A. M. (Ed.), *The nature of moral responsibility: New essays*, pp. 65-87. New York: Oxford University Press.
- Roskies, A. (2006). “Neuroscientific challenges to free will and responsibility”, *Trends in Cognitive Science* 10: 419-423.
- Roskies, A. and Nahmias, E. (2017). ““Local determination”, even if we could find it, does not challenge free will: Commentary on Marcelo Fischborn”, *Philosophical Psychology* 30: 185-197.
- Roskies, A. L. (2011). “Why Libet's studies don't pose a threat to free will”. In: Sinnott-Armstrong, W. & Nadel, L. (Ed.), *Conscious will and responsibility: A tribute to Benjamin Libet*, pp. 11-22. Oxford: Oxford University Press.
- Roskies, A. L. (2014). “Can neuroscience resolve issues about free will?”. In: Sinnott-Armstrong, W. (Ed.), *Moral psychology: Free will and moral responsibility*, pp. 103-126. Cambridge, MA: MIT Press.
- Roskies, A. L. and Malle, B. F. (2013). “A Strawsonian look at desert”, *Philosophical Explorations* 16: 133-152.
- Roskies, A. L. and Nichols, S. (2008). “Bringing moral responsibility down to earth”, *The Journal of Philosophy* 105: 371-388.
- Sarkissian, H. (2010). “Minor tweaks, major payoffs: The problems and promise of situationism in moral philosophy”, *Philosophers' Imprint* 10: 1-15.
- Schlosser, M. E. (2012). “Free will and the unconscious precursors of choice”, *Philosophical Psychology* 25: 365-384.
- Schlosser, M. E. (2013). “Conscious will, reason-responsiveness, and moral responsibility”, *The Journal of Ethics* 17: 205-232.
- Schooler, J.; Nadelhoffer, T.; Nahmias, E. and Vohs, K. D. (2015). “Measuring and manipulating beliefs and behaviors associated with free will: The good, the bad, and the ugly”. In: Mele, A. R. (Ed.), *Surrounding free will: Philosophy, Psychology, Neuroscience*, pp. 72-94. New York: Oxford University Press.
- Shepherd, J. (2015). “Scientific challenges to free will and moral responsibility”, *Philosophy Compass* 10: 197-207.
- Soon, C. S.; Brass, M.; Heinze, H.-J. and Haynes, J.-D. (2008). “Unconscious determinants of

- free decisions in the human brain”, *Nature Neuroscience* 11: 543-545.
- Spohn, C. and Holleran, D. (2000). “The imprisonment penalty paid by the young, unemployed black and hispanic male offenders”, *Criminology* 38: 281-306.
- Steffensmeier, D.; Ulmer, J. and Kramer, J. (1998). “The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, black, and male”, *Criminology* 36: 763-798.
- Strawson, P. F. (1962). “Freedom and resentment”. In: Pereboom, D. (Ed.), *Free will*, pp. 148-171. Indianapolis: Hackett.
- Suppes, P. (1993). “The transcendental character of determinism”, *Midwest Studies in Philosophy* 18: 242-257.
- Teicher, M. H. (2010). “Commentary: Childhood abuse: New insights into its association with posttraumatic stress, suicidal ideation, and aggression”, *Journal of Pediatric Psychology* 35: 578-580.
- Todd, P. (2016). “Strawson, moral responsibility, and the “order of explanation”: An intervention”, *Ethics* 127: 208-240.
- Vargas, M. (2007). “Revisionism”. In: Fischer, J. M.; Kane, R.; Pereboom, D. & Vargas, M. (Ed.), *Four views on free will*, pp. 126-165. Malden: Wiley-Blackwell.
- Vargas, M. (2013b). *Building better beings: A theory of moral responsibility*. Oxford: Oxford University Press.
- Vargas, M. (2013a). “Situationism and moral responsibility: Free will in fragments”. In: Clark, A.; Kiverstein, J. & Vierkant, T. (Ed.), *Decomposing the will*, pp. 325-349. Oxford: Oxford University Press.
- Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.
- Werndl, C. (2013). “On choosing between deterministic and indeterministic models: Underdetermination and indirect evidence”, *Synthese* 190: 2243-2265.
- Werndl, C. (2016). “Determinism and indeterminism”. In: Humphreys, P. (Ed.), *The Oxford handbook of philosophy of science*, pp. 210-232. New York: Oxford University Press.
- WHO (2013). *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence*. : World Health Organization.
- Zimmerman, M. (2015). “Varieties of moral responsibility”. In: Clarke, R.; McKenna, M. & Smith, A. M. (Ed.), *The nature of moral responsibility: New essays*, pp. 45-64. New York: Oxford University Press.