

Free will eliminativism: reference, error, and phenomenology

Gregg D. Caruso

Philosophical Studies

An International Journal for Philosophy
in the Analytic Tradition

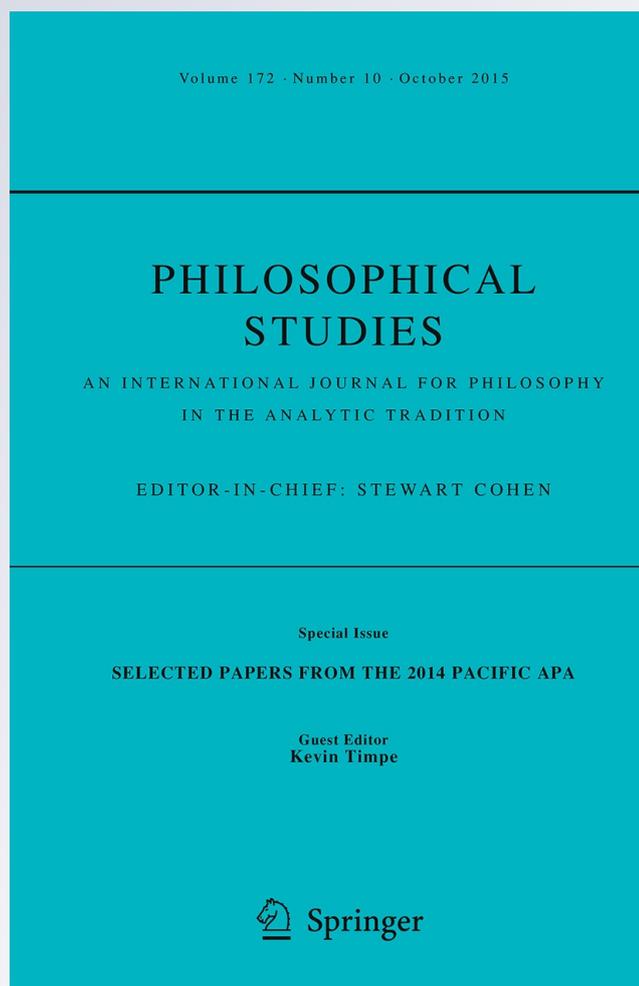
ISSN 0031-8116

Volume 172

Number 10

Philos Stud (2015) 172:2823-2833

DOI 10.1007/s11098-015-0453-x



Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Free will eliminativism: reference, error, and phenomenology

Gregg D. Caruso

Published online: 30 January 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Shaun Nichols has recently argued that while the folk notion of free will is associated with error, a question still remains whether the concept of free will should be *eliminated* or *preserved*. He maintains that like other eliminativist arguments in philosophy, arguments that free will is an illusion seem to depend on substantive assumptions about reference. According to free will eliminativists, people have deeply mistaken beliefs about free will and this entails that free will does not exist. However, an alternative reaction is that free will does exist, we just have some deeply mistaken beliefs about it. According to Nichols, all such debates boil down to whether or not the erroneous folk term in question successfully refers or not. Since Nichols adopts the view that reference is systematically ambiguous, he maintains that in some contexts it's appropriate to take a restrictivist view about whether a term embedded in a false theory refers, while in other contexts it's appropriate to take a liberal view about whether a token of the very same term refers. This, according to Nichols, affords the possibility of saying that the sentence “free will exists” is false in some contexts and true in others. In this paper I argue that even if we grant Nichols his pluralistic approach to reference, there is still good reason to prefer eliminativism to preservationism with regard to free will. My argument focuses on one important difference between the concept of “free will” and other theoretical terms embedded in false theories—i.e., the role that the *phenomenology of free agency* plays in reference fixing.

Keywords Free will · Eliminativism · Reference · Phenomenology

Shaun Nichols (2013) has recently argued that while the folk notion of free will is associated with error, a question still remains whether the concept of free will should be *eliminated* or *preserved*. He maintains that like other eliminativist arguments in

G. D. Caruso (✉)
Humanities Department, Corning Community College (SUNY), 1 Academic Drive,
Corning, NY 14830, USA
e-mail: gcaruso@corning-cc.edu

philosophy, arguments that free will is an illusion seem to depend on substantive assumptions about reference. According to free will eliminativists, people have deeply mistaken beliefs about free will and this entails that free will does not exist (e.g., Pereboom 2001; Strawson 1986). However, an alternative reaction is that free will does exist, we just have some deeply mistaken beliefs about it (e.g., Vargas 2005, 2007, 2013). According to Nichols, all such debates boil down to whether or not the erroneous folk term in question successfully refers or not. Since Nichols adopts the view that reference is systematically ambiguous, he maintains that in some contexts it's appropriate to take a restrictivist view about whether a term embedded in a false theory refers, while in other contexts it's appropriate to take a liberal view about whether a token of the very same term refers. This, according to Nichols, affords the possibility of saying that the sentence "free will exists" is false in some contexts and true in others. Hence, according to Nichols, *even if we acknowledge that our everyday conception of free will is significantly in error*, there is flexibility in whether we embrace the eliminativist claim.

In this paper I argue that even if we grant Nichols his pluralistic approach to reference, there is still good reason to prefer eliminativism to preservationism with regard to free will. My argument focuses on one important difference between the concept of "free will" and other theoretical terms embedded in false theories—i.e., the role that the *phenomenology of free agency* plays in reference fixing. I argue that once we acknowledge the importance of phenomenology, eliminativism appears to follow on whichever account of reference one adopts.

In section 1, I briefly outline Nichols's account of free will and error. Then in Sect. 2, I argue that on *either* a descriptive account of reference *or* a causal-historical one—the two main accounts of reference Nichols discusses—eliminativism appears to follow. On a descriptive account of reference, eliminativism follows since, as Nichols himself argues,¹ the folk concept of free will contains significant error, hence nothing satisfies the description. Yet, on a causal-historical account of reference, despite Nichols's claims to the contrary, eliminativism *also* seems to follow since it is prima facie plausible to think that the concept of "free will" was originally baptized in a causal-historical story which appealed, at least in part, to our first-person experience of free agency. And since Nichols and I have both argued elsewhere that the *phenomenology of free agency* is best viewed in a libertarian and incompatibilist sense (see Caruso 2012; Deery et al. 2013), one that is likely in error, eliminativism awaits us down this path too.

1 Nichols on free will and error

Nichols's argument begins with the assumption that the ordinary notion of free will is in error because it falsely presupposes that indeterminism is true (2013, 203–204).² He maintains that:

¹ See, for example, Nichols and Knobe (2007), Sarkissian et al. (2010), Roskies and Nichols (2008).

² As Nichols notes, his main argument can also be made if we assume instead that (1) in addition to the presupposition of indeterminism, people also have a presupposition that choice isn't random and that (2) all events are either determined or random (cf. Russell 1995, 14; Kane 1996, 11) (Nichols 2013, fn.1).

A growing body of empirical evidence indicates that people think that their choices aren't determined (Nichols and Knobe 2007; Sarkissian et al. 2010). It is not just that they don't have the belief that their choices are determined. Rather, they positively think that their choices are not determined. And this belief is implicated in their thoughts about free will. For instance, when presented with a description of a deterministic universe, most participants say that in that universe, people don't have free will (Roskies and Nichols 2008; also Deery et al. 2013). This provides reason to think that the everyday conception of free will is not compatible with determinism (see also Rose and Nichols 2013). (2013, 203)

If we combine these empirical findings with the assumption that determinism is true and that we do not make indeterminist (nonrandom) choices, we get the conclusion that the ordinary notion of free will is in error.

While these assumptions are, of course, highly contested,³ I will grant them for the sake of this paper, since the issue that Nichols and I are both interested in concerns the consequences of *folk error* on free will. Since “the issue can only be joined if we assume that there is something interestingly mistaken about the folk view” (2013, 203), I will join Nichols in assuming that the ordinary notion of free will is in error. Now, one may think that the deck has been so thoroughly stacked against free will that it trivially follows that it does not exist, but Nichols disagrees. It is here that he introduces his reference pluralism and his argument against free will eliminativism.

Nichols correctly points out that it is not always clear that we should adopt eliminativism for theoretical terms embedded in false theories. If contemporary philosophical debates over the last several decades have taught us anything—including debates in ethics over error theory (Mackie 1997; Blackburn 1985), in social-political philosophy over the existence of race (Appiah 1995; Andreasen 2000; Mallon 2006), in philosophy of science over scientific realism (Feyerabend 1962; Laudan 1984; Boyd 2002), and in philosophy of mind over eliminative materialism (e.g., Stich 1983; Lycan 1988)—they have taught us that wherever you find one group of philosophers advocating *eliminativism*, you will usually find another group of philosophers advocating *revisionism* or *preservationism*. “In each of these debates,” Nichols notes, “eliminativists maintain that *K doesn't exist* (where *K* might be morality, race, belief, etc.). Typically shortly after an eliminativist claim of this sort is made another group of philosophers adopt a *preservationist* position. In effect, they say, *Ks aren't what we thought they were*” (Nichols 2013, 204). According to Nichols, then, the fact that the ordinary notion of free will has a false presupposition does not immediately entail that there is no free will.

It's at this point that Nichols discusses the role that reference plays in such disputes. The key distinction is between theories of reference that are *liberal* and those that are conservative. For brevity and ease of discussion, I will follow Nichols

³ For an alternative account of our folk-psychological intuitions about free will, see (Nahmias et al. 2005, 2006), Nahmias and Murray (2010).

in focusing on *causal-historical accounts* as the key example of a liberal theory, and a *restrictive descriptivism* as the key example of a conservative theory.⁴ To see how these different theories of reference can play out in such disputes, consider, for example, the disagreement between Stephen Stich and William Lycan over eliminative materialism (see Stich 1983; Lycan 1988).

Stich famously defends a version of eliminative materialism, maintaining that concepts like “belief” are part of a folk psychological theory that is massively mistaken. As Nichols describes Stich’s argument:

[Stich] begins by setting out David Lewis’ (1972) descriptivist account of how theoretical terms get their meaning and reference from the theory in which they are embedded (Stich 1983, 17–21). A theoretical term refers to whatever object or class of objects satisfies some critical set of claims in the theory. Stich then argues that the folk psychological theory that gives meaning and reference to “belief” is deeply erroneous; from this he concludes that beliefs don’t exist. (2013, 204).

Lycan, on the other hand, rejects restrictive descriptivism (and with it eliminativism about belief), adopting instead a “liberal” view of reference fixing. Lycan writes:

I am at pains to advocate a very liberal view... I am entirely willing to give up fairly large chunks of our commonsensical or platitudinous theory of belief or of desire (or of almost anything else) and decide that we were just wrong about a lot of things, without drawing the inference that we are no longer talking about belief or desire. To put the matter crudely, I incline away from Lewis’s Carnapian and/or Rylean cluster theory of the reference of theoretical terms, and toward Putnam’s... causal-historical theory. (1988, 31–32; as quoted by Nichols 2013, 205).

On Lycan’s approach to reference fixing, the reference of a theoretical term can succeed even if the theory is radically mistaken. On this liberal/causal-historical approach, the reference of a theoretical term is the entity or kind that was “baptized” when the term was introduced. Since the term continues to refer to the entity or kind that was baptized at the end of the causal-historical chain of transmission, “people can have massive misconceptions about the objects that their terms (or concepts) refer to” (Nichols 2013, 205).

Other examples of disputes between liberal and conservative approaches to reference can be found with regard to the concepts of “whales,” “witches,” and “phlogiston.” Each of these concepts was originally enmeshed in significant error, yet some were preserved (e.g., whales) while others were eliminated (e.g., witches and phlogiston). The folksy concept “whale,” for example, falsely presupposed that whales were fish. Instead of eliminating the concept, however, a liberal approach to reference was employed and “taxonomy kept the label and adjusted the concept, moving WHALE onto the mammal branch” (Nichols 2013, 208). On the other hand,

⁴ Nichols, however, notes that a descriptivist can be more or less liberal about reference fixing, depending “on how much of the theory has to be true in order for the term to refer” (2013, fn.2).

“witches” and “phlogiston” were eliminated because a more restrictive descriptivism was adopted.

Nichols argues that we can view the free will debate, especially the debate over eliminativism, as similarly a debate over which reference-fixing convention we should employ. Free will eliminativists often employ a descriptivist convention for reference, according to which a natural kind term refers by means of an associated description. *Preservationists*, on the other hand, maintain a causal-historical convention for reference, according to which an initial baptism sets the referent of a term, even if the nature of the object referred to be misunderstood at the time of baptism. Hence, the eliminativist maintains that since the descriptive nature of the term yields a failure of reference, there is *no such thing* as free will. Preservationists (including revisionists)⁵ argue instead that there is free will it's just that “everyone has merely been under some misapprehension about its nature” (Nichols 2013, 209). How, then, should we resolve this dispute?

It is here that Nichols argues for the systematic ambiguity of reference and a pluralistic approach to free will. He introduces recent experimental evidence that suggests that speakers regularly use different referential conventions for different kinds of sentences in which natural kind terms appear (see Pinillos, Mallon, and Nichols; as discussed in Nichols 2013, 209–213). Since Nichols concludes that reference is systematically ambiguous, he maintains that when it comes to free will it is sometimes correct to say, “Free will does not exist,” since people employ the descriptivist convention in sentences like this one. Nevertheless, it is also correct to say, “Free will isn't what we thought,” since a causal-historical convention of reference can be employed for sentences of this type. With all the pieces in place, Nichols concludes that both sides can agree that libertarian free will does not exist and the everyday conception of free will is significantly in error, *without* precluding the possibility of preservationism—e.g., there is a flexibility in whether we embrace the eliminativist claim.

2 Both paths lead to eliminativism

My reply to Nichols will concede four of his main assumptions. I am willing to grant for the sake of argument that (1) the concept of “free will” is enmeshed in significant error, (2) the free will debate depends on substantive assumptions about reference, (3) not all theoretical terms embedded in false theories should be eliminated, and (4) reference is systematically ambiguous. I maintain that even granting these assumptions a strong case can be made for free will eliminativism. My argument will focus on the causal-historical approach to reference fixing since this is the one typically employed by preservationists. But before discussing this liberal approach to reference fixing, let me just say a quick word about descriptivist accounts.

⁵ On Nichols's taxonomy of positions, revisionism (e.g., Vargas 2005, 2007) is just one form of preservationism. Preservationism is a broader category, including other possible positions as well (see Nichols 2013).

It's pretty clear that on Nichols's own assumptions the path of restrictive descriptivism leads to eliminativism. Given assumption (1) above, the assumption that the concept of "free will" is enmeshed in significant error, it follows rather directly that if we adopt a descriptivist approach to theoretical terms (according to which the meaning and reference of a theoretical term is fixed by some critical set of claims within a given theory), free will eliminativism follows since the term yields a failure of reference. Some philosophers will no doubt resist this conclusion by questioning assumption (1), but seeing as though Nichols appears to concede that this path leads to eliminativism, and since I am granting assumption (1) for the sake of argument, I will focus my attention on the second main path.

Can we avoid eliminativism if we adopt a causal-historical approach? I think it depends on how one conceives of the original baptism. It's rather easy to imagine how concepts like "whale" and "water" were initially baptized. On a causal-historical account, the term "whale" can continue to refer despite the fact that a key component of the folk description of whales (i.e., that they were fish) turned out to be wrong, since the initial baptism most likely took place in a demonstrative way. But free will seems significantly different than whales. The same reference-fixing properties are not available for free will. How, then, are we to conceive of the initial baptism that set the referent of "free will"? Nichols is unfortunately silent on this point, but I believe there are several leading candidates. It is possible, for example, that the initial baptism was to whatever power or ability is required to justify ascriptions of desert-based moral responsibility, or to that feature of choice and action that justifies our reactive attitudes, or to a set of compatibilist-friendly capacities (e.g., reasons-responsiveness). While I cannot adequately address all these possibilities here (although I will say something about them below), my proposal is that we should look elsewhere, i.e. to the *phenomenology of free agency*.

I maintain that it is prima facie plausible to think that the concept of "free will" was originally baptized in a causal-historical story that appealed, at least in some significant way, to our first-person experience of free agency.⁶ This does not mean that cultural attitudes about moral responsibility played no role in reference fixing, nor does it deny that certain philosophical and religious doctrines were important in shaping our beliefs and attitudes about free will. My suggestion is simply that our first-person experience is more primitive and basic, and hence prima facie plausible as the primary candidate for reference fixing. Imagine, for example, a possible world in which we retain reasons-responsiveness and other compatibilist capacities but do not experience ourselves as acting freely—or worse still, actively experience ourselves as *lacking* free agency, as (say) patients with anarchic hand syndrome, Tourette's syndrome, and schizophrenic thought insertion often do. In such a world, it is unlikely that the concept of free will would have ever been introduced, despite the fact that other candidates for reference fixing are still present.

Now one may be tempted to think that this has more to do with the voluntary versus involuntary distinction—e.g., that free will is baptized in terms of its role in

⁶ I should note that there is some precedent of natural kind terms getting their reference fixed by phenomenology by causal-historical theorists Kripke (1980), for example, famously does this for the concept of *pain*.

distinguishing voluntary from involuntary behavior, and that my example has more to do with involuntariness than it does with phenomenology—but I do not believe that this is the case. Agents can engage in voluntary actions (actions that are clearly distinct from involuntary actions) while not experiencing a sense of agency or free will. Consider, for example, the long distance truck driver who finds himself at his destination but does not recall how he got there. Or consider Daniel Wegner's examples of *automatisms*—i.e. cases where we experience no sense of free agency while performing voluntary actions (2002, ch.4). Such automatisms not only involve a “lack of the feeling of doing an action but may even go beyond this to include a distinct feeling that we are *not* doing” (Wegner 2002: 99). Common examples include automatic writing, Ouija board spelling, the Chevreul pendulum, dowsing, and the phenomenon of ideomotor action (Wegner 2002: ch.4).

The fact that the phenomenology of free agency is absent in such cases is so profound that during an automatism “the person may vehemently resist describing the action as consciously or personally caused. It seems to come from somewhere else or at least not from oneself” (2002: 99). Yet despite this phenomenological fact, what's remarkable about such cases is that they are clearly caused by the agent, are driven by goals and intentions, and involve sophisticated actions and movements (unlike the herky-jerky movements of many involuntary actions). I maintain that if we experienced *all of our actions* in this way, even if we retained the voluntary/involuntary distinction as a candidate for reference fixing, it is still unlikely that the concept of free will would have ever been introduced.

An additional reason for thinking that the *phenomenology of free agency* is important comes from the role it has played historically in libertarian arguments. Libertarians have long emphasized our *feeling of freedom* and our introspective abilities. In fact, many libertarians have suggested that our introspection of the decision-making process, along with our strong feeling of freedom, provides some kind of evidence for the existence of free will. As Ledger Wood describes this common form of reasoning:

Most advocates of the free will doctrine believe that the mind is directly aware of its freedom in the very act of making a decision, and thus that freedom is an immediate datum of our introspective awareness. “I feel myself free, *therefore*, I am free,” runs the simplest and perhaps the most compelling of the arguments for freedom. (1941: 387)

While I have elsewhere argued against this *introspective argument* for free will (Caruso 2008, 2012), there is no denying its intuitive appeal. It captures the fact that for many philosophers and ordinary folk, the phenomenology of free agency is *the* thing to be explained and preserved, often at all costs! In fact, many agent-causal libertarians are willing to bend over backwards to preserve it, for it is our feeling of freedom that is the *raison d'être* of libertarianism (see, e.g., O'Connor 1995, 196; Taylor 1992, 51; Campbell 1957, 169). Conceiving of the initial baptism in terms of phenomenology therefore has the virtue of acknowledging this long historical tradition as well as its wider appeal among ordinary folk.

A third reason for focusing on the role of phenomenology in the initial baptism is that many of the other candidates for reference fixing seem historically

anachronistic. Take reasons-responsiveness for example. A preservationist could argue that when the term “free will” was initially baptized, it was *not* to some libertarian phenomenology but to a more innocuous set of reasons-responsive capacities. It need not be the case that we recognized these capacities as reasons-responsiveness at the time, only that reference was primarily fixed to these capacities. While it is impossible to rule out this possibility completely, I believe there are reasons to doubt that the reference-fixing properties were so narrowly focused. The concept of free will stretches back to pre-scientific, pre-theoretical times when it was likely that *more than* just reasons-responsiveness would have been part of the initial baptism.

Reasons-responsiveness is of course essential to our phenomenology as agents, reasoners, and moral beings, but it does not capture in *totality* our pre-theoretical self-conception as agents. If Nichols’s own work is any indication, people pre-theoretically conceive of themselves as capable of exercising indeterminist free will (2012), as being able to do otherwise in a way that is not easily amendable to conditional analyses (Nichols and Knobe 2007), and as being introspectively aware of the key mental aspects of decision-formation (Kozuch and Nichols 2011). Paul Bloom’s (2004) work further reveals that a major component of our folk psychology (at least in the West) is the belief in dualism. According to Bloom, the belief in dualism comes naturally to children and this dualistic worldview persists in adulthood. Preservationists that want to restrict the reference-fixing properties of the initial baptism so as to preserve “free will” are being disingenuous, I maintain, by downplaying or dismissing the anti-scientific components of our pre-theoretical self-conception as agents.

Preservationists, of course, could shift their focus to more external reference-fixing properties, such as the cultural practices associated with punishment, reward, praise, and blame. Perhaps these practices could, in some way, ground an initial baptism. But problems belie this path too. Drawing on Tamler Sommers’s recent work in *Relative Justice* (2012), there are past and present cultures that feel completely justified in praising, blaming, and punishing people (sometimes severely) without requiring that those people satisfy the control conditions often associated with free will. Sommers, for example, describes how “honor cultures” often condone severely punishing, sometimes even killing, seemingly innocent individuals for the ‘sins’ of their relatives.⁷ Non-honor cultures, like our own, consider this fundamentally unfair and unjust. Furthermore, honor cultures shun third-party punishment, whereas third-party punishment is the norm in Western cultures.⁸ According to Sommers, these cultural variations are the product of different ways of life resulting from adaptations to local ecological conditions.

⁷ Consider the practice of collective punishment—“retaliation against people who play no part in committing the offence that is to be avenged” (2012, 48). “The anthropologist Ram Nath Sharma (1997, 377) writes that in most tribal societies, “the punishment given for murder is death, but this punishment may not be given to him who has murdered. In his place, some other member of his family, group, or clan may be killed since the group is collectively responsible for the criminal acts of each of its members” (Sommers 2012, 48).

⁸ In honor cultures, it is very important that the victim respond personally, and not necessarily to the person that harmed them. As Sommers describes it, “Third-party punishment strips the victim of his

If Sommers is correct about the cultural diversity of our moral responsibility practices, then preservationists who propose that the initial baptism was to a natural kind picked out by these practices would be forced to conclude that there are no universal reference-fixing properties for “free will,” only a divergent and culturally relative set of norms and practices.⁹ If, for example, individualist cultures like our own require that a control condition be met for ascriptions of desert-based moral responsibility, while other cultures de-emphasize or lack this condition altogether,¹⁰ then there would be no common feature to serve as the reference fixer for “free will.” This lack of a common property (or set of norms and practices) among cultures represents a serious problem for preservationists; especially those who wish to avoid a wholesale relativism about free will.

3 Summarizing the problem

For the reasons outlined above, I maintain that the phenomenology of free agency remains the best candidate for the initial baptism. If correct, there is good reason to conclude that free will eliminativism follows on *both* the causal-historical path and the restrictive descriptivist path. On a descriptivist account of reference, eliminativism follows since the folk concept of free will contains significant error, hence nothing satisfies the description. On a causal-historical approach, eliminativism likewise follows if we assume, as Nichols does, that the phenomenology of free will is illusory. Nichols has elsewhere argued, as have I, that our phenomenology is libertarian and incompatibilist in nature (see Deery et al. 2013; Caruso 2012). But on the assumption that we lack libertarian free will, one of the key assumptions of Nichols’s argument, it turns out that our first-person experience of free will is illusory as well. Hence, if the initial baptism was to the phenomenology of free agency, and that phenomenology is illusory (a conclusion reached on Nichols’s own assumptions), we once again arrive at a failure of reference.

I conclude that on the set of assumptions laid out in this paper, both descriptivist and causal-historical approaches lead to free will eliminativism. This conclusion is, of course, contingent on a set of assumptions, and critics are free to challenge them.

Footnote 8 continued

opportunity to avenge himself. Since personal retaliation is crucial, one would expect the emergence of norms that discourage third-party punishment and attitudes that find it to be unsatisfying” (2012, 44).

⁹ A preservationist could, perhaps, argue that there *is* an underlying common property (or set of properties) among cultures, but that moral responsibility practices evolved differently given different ways of life resulting from adaptations to local ecological conditions (Justin Caouette, in correspondence). The challenge, however, is for the preservationist to explain exactly what this common property (or set of properties) is. The proposal under consideration, remember, is to conceive of the initial baptism in terms of an external set of reference-fixing properties, such as the cultural practices associated with punishment, reward, praise, and blame. Sommers, however, maintains that there is no common property (or set of properties) that can fit this bill, hence preservationists looking to make this move need to address his challenge head on.

¹⁰ As Sommers notes, “In many honor cultures, the condition appears to either be de-emphasized or in some cases absent; agents need not have control over an act in order to be deemed fair and morally appropriate targets of blame and retaliation” (2012, 48).

But for those who agree with Nichols that the folk notion of free will is significantly in error and that the phenomenology of free will is illusory as well, the dilemma presented above remains a serious one for anyone wishing to avoid free will eliminativism.

References

- Andreasen, R. O. (2000). Race: Biological reality or social construct? *Philosophy of Science*, *67*, S653–S666.
- Appiah, A. (1995). The uncompleted argument: Du Bois and the illusion of race. In Linda A. Bell & D. Blumenfeld (Eds.), *Overcoming racism and sexism* (pp. 59–78). Lanham, MD: Rowman and Littlefield.
- Blackburn, S. (1985). Errors and the phenomenology of value. In T. Honderich (Ed.), *Morality and objectivity* (pp. 1–22). London: Routledge & Kegan Paul.
- Bloom, P. (2004). *Descartes' baby*. New York: Basic Books.
- Boyd, R. (2002). Scientific realism. In E.N. Zalta (Ed) *The Stanford encyclopedia of philosophy* (Summer 2002 ed.), <http://plato.stanford.edu/archives/sum2002/entries/scientific-realism/>
- Campbell, J. (1957). *Of selfhood and godhood*. London: Allen & Unwin.
- Caruso, G. (2008). Consciousness and free will: A critique of the argument from introspection. *Southwest Philosophy Review*, *24*(1), 219–231.
- Caruso, G.D. (2012). *Free will and consciousness: A determinist account of the illusion of free will*. Lanham, MD: Lexington Books.
- Deery, O., Bedke, M., & Nichols S. (2013). Phenomenal abilities: Incompatibilism and the experience of agency. In D. Shoemaker (Ed) *Oxford studies in agency and responsibility*. New York: Oxford University Press.
- Feyerabend, P. K. (1962). Explanation, reduction and empiricism. In H. Feigl & G. Maxwell (Eds.), *Minnesota studies in the philosophy of science: Scientific explanation, space, and time* (Vol. 3, pp. 28–97). Minneapolis: University of Minnesota Press.
- Kane, R. (1996). *The significance of free will*. New York: Oxford University Press.
- Kozuch, B., & Nichols, S. (2011). Awareness of unawareness. Folk psychology and introspective transparency. *Journal of Consciousness Studies*, *18*(11–12), 135–160.
- Kripke, S. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.
- Laudan, L. (1984). *Science and values* (Vol. 66). Berkeley: University of California Press.
- Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, *50*(3), 249–258.
- Lycan, W. G. (1988). *Judgement and justification*. New York: Cambridge University Press.
- Mackie, J.L. (1997). *Ethics: Investigating right and wrong*. London: Penguin Books Ltd.
- Mallon, R. (2006). Race: Normative, not metaphysical or semantic. *Ethics*, *116*(3), 525–551.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005). Surveying freedom: Folk intuitions about free will and moral responsibility. *Philosophical Psychology*, *18*, 561–584.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2006). Is incompatibilism intuitive? *Philosophy and Phenomenological Research*, *73*, 28–53.
- Nahmias, E., & Murray, D. (2010). Experimental philosophy on free will: An error theory for incompatibilist intuitions. In J. Aguilar, A. Buckareff, & K. Frankish (Eds.), *New waves in philosophy of action* (pp. 189–215). New York: Palgrave-Macmillan.
- Nichols, S. (2012). The indeterminist intuition. *The Monist*, *95*(2), 290–307.
- Nichols, S. (2013). Free will and error. In Gregg D. Caruso (Ed.), *Exploring the illusion of free will and moral responsibility*. Lanham, MD: Lexington Books.
- Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous*, *41*, 663–685.
- Nichols, S., Pinillos, A., & Mallon, R. (Forthcoming). Ambiguous reference. *Mind*.
- O'Connor, T. (1995). Agent causation. In T. O'Connor (Ed.), *Agents, causes and events: Essays on free will and indeterminism* (pp. 173–200). Oxford: Oxford University Press.

- Pereboom, D. (2001). *Living without free will*. New York: Cambridge University Press.
- Rose, D., & Nichols, S. (2013). The lesson of bypassing. *Review of Philosophy and Psychology*, 4, 599–619.
- Roskies, A., & Nichols, S. (2008). Bringing moral responsibility down to Earth. *Journal of Philosophy*, 105(7), 371–388.
- Russell, P. (1995). *Freedom and moral sentiment*. New York: Oxford University Press.
- Sarkissian, H., Chatterjee, A., De Brigard, F., Knobe, J., Nichols, S., & Sirker, S. (2010). Is belief in free will a cultural universal? *Mind and Language*, 25(3), 346–358.
- Sommers, T. (2012). *Relative justice: Cultural diversity, free will, and moral responsibility*. Princeton: Princeton University Press.
- Stich, S. P. (1983). *From folk psychology to cognitive science: The case against belief*. Cambridge, MA: MIT Press.
- Strawson, G. (1986). *Freedom and belief*. Oxford: Oxford University Press.
- Taylor, R. (1992). *Metaphysics*. 4th ed. Englewood Cliffs, NJ: Prentice-Hall (First ed., 1963).
- Vargas, M. (2005). The revisionist's guide to responsibility. *Philosophical Studies*, 125, 399–429.
- Vargas, M. (2007). Revisionism. In J. Fischer, et al. (Eds.), *Four views on free will*. Malden, MA: Blackwell.
- Vargas, M. (2013). *Building better beings*. Oxford: Oxford University Press.
- Wegner, D. M. (2002). *The Illusion of conscious will*. Cambridge, MA: Bradford Books, MIT Press.
- Wood, L. (1941). Determinism: Free will is an illusion. *Philosophy*, 16, 386–389.