We offer two kinds of criticism in the spirit of support for Doris's socially scaffolded pluralism regarding agency. First: the skeptical force of potential "goofy influences" is not as straightforward as Doris argues. Second: Doris's positive theory must address more goofy influences due to social processes that appear to fall under his criteria for agency-promoting practices. Finally, we highlight "arms race" phenomena in Doris's social dynamics that invite closer attention in further development of theory.
Target article author: John Doris

Word Counts
Abstract: 78
Main: 2020
References: 303
Total: 2468

Getting by with a little help from our friends

Enoch Lambert, Center for Cognitive Studies, Tufts University,

Email: enoch.lambert@gmail.com

Daniel C. Dennett, Center for Cognitive Studies, Tufts University,

Email: daniel.dennett@tufts.edu, http://ase.tufts.edu/cogstud/dennett/

115 Miner Hall, Tufts University Medford, MA 02155 (617) 627-3297

ABSTRACT
We offer two kinds of constructive criticism in the spirit of support for Doris’s socially scaffolded pluralism regarding agency. First: the skeptical force of potential “goofy influences” is not as straightforward as Doris argues. Second: Doris’s positive theory must address more goofy influences due to social processes that appear to fall under his criteria for agency-promoting practices. Finally, we highlight “arms race” phenomena in Doris’s social dynamics that invite closer attention in further development of his theory.

Doris conducts a master class for psychologists on how to extract value from the philosophical debates, and for philosophers on how to use empirical work in psychology to inform their theorizing. In both endeavors, one has to learn how to take the declarations with more than a few grains of salt, which Doris applies judiciously. We heartily endorse what we take to be a major lesson: that what we learn from science, while sometimes shocking, need not destroy our confidence in our own practical agency. Rather, by informing our understanding of our agential strengths and weaknesses, science can guide us in discovering and strengthening those practices that foster our agential powers. Of special note is his case that self-ignorance can be crucial to our projects of building and expressing our central values, showing how accurate reflection can actually undermine agency in some situations. He has also done the study of practical reason a great service by setting up a framework for exploring its socially scaffolded nature. In our comment we aim to contribute to that ongoing project. While we believe Doris is right about the largely social nature of agency, we raise some questions about the skeptical force of the psychology he cites against the role of accurate self-knowledge in our deliberations. We also urge that his own “collaborative-negotiative-dialogical” framework faces significant threats from social psychology—more so than acknowledged.

Doris’s critique
First, we wish to question the strength of the case Doris mounts for *global* skepticism regarding the role of accurate self-knowledge in our deliberations. We are more concerned about the size of experimental effects and their implications for everyday decision making than Doris is. It is instructive to recall the reason why so much psychology focuses on surprising effects. Vast swaths of common wisdom concerning self-knowledge prevent psychologists from so much as attempting to confirm things like whether people tend to be accurate about whether they prefer $1000 to a pin prick, or social praise to ridicule. Finding a new way of generating small, surprising effects may be rewarded in psychology, but it is not clear whether or how the common lore of everyday psychology that psychologists never bother to investigate is undermined by it.

Doris dismisses the importance of statistically small sizes partly by saying that known “goofy influences” on behavior indicate an ocean of unknown ones; and partly by saying that such influences may “aggregate” in ways that medical interventions can (Doris, 2015, p. 64). Our own speculative mechanics of goofy influences suggest a different lesson. If “eyespots and pronouns are in the mix” (to use Doris’ nice phrasing), then humans are likely assailed by goofy influences *continuously* (ibid). The priming and automaticity literatures from across psychology suggest no principles for ruling out much of anything as potentially goofy influence. But, if this is so, how do we manage to hold it all together? Why are we not driven every which way by the onslaught of disparate priming stimuli? And how are we able to come by the amount of common knowledge of human psychology that we do? Why can we predict so well what others will do based on “typical” perceptions and desires (which we also attribute to ourselves)? When predicting what the drivers of other cars on the road will do, we justifiably pay no attention to which images on which billboards they recently saw, or the content of the radio advertisement they are hearing, or whether their vehicle interior is leather, or . . . . It isn’t that we are in a position to rule out such things ever having some influence on how they drive, whether at a micro-level, or, on occasion, at a life-altering level. But our attributions are sufficiently reliable enough of the time so that it makes no sense to let such influences trigger general skepticism of our usual interpretive and predictive capacities. Similar considerations apply to our own case. It would be silly, for instance, to decide to live as close as possible to the market simply because it would minimize the amount of goofy influence encountered every time we need to do our shopping.

Moreover, Doris ignores the prospect of a gradient between goofy and not-so-goofy, to go along with his valuable gradient between explicit self-reflection and the sort of automatic self-monitoring that gets us relatively gracefully through the day. The fact that pictures of watchful eyes should nudge more honest coffee transactions is striking, but not so striking or upsetting as the non-fact—we wager—that pictures of bicycles or rooftops have the same effect. Doris’s richly detailed account of actual decision-making suggests that in the real-time hasty triage involved in all but the
most portentous moral decisions, a “subliminal” hint about being observed and caught could be just enough to bias the choices made without the choosers’ noticing.

Next consider one of the roughly third of test-subjects who detected the switches in the moral choice blindness experiment Doris cites (139; see Hall et al 2012). What should such a subject conclude upon learning the results of the experiment? That she got lucky? Why would that be more reasonable than to conclude that, for whatever reason, she was more attentive (or cared more, or …)? Perhaps she should conclude that her capacity to recognize her own moral positions is more susceptible to error than she would have thought, and so she should keep watch. But it doesn’t seem reasonable to conclude that she should be an outright skeptic of her ability to recognize her own morality. And, in general, we urge that individual variation in susceptibility to goofy influences not be swept aside as so much noise. Why is it that goofy influences do not affect some subjects in any given experiment? Are some people who are less susceptible in specific experiments more generally resistant to goofy influences? If so, why? Can any pattern at all be detected in failure to succumb to goofy influence? It seems that such possibilities remain live empirical hypotheses to be ruled out (or in!) rather than assumed. Until we know more about the mechanics of goofy influences, it seems rash to let them completely undermine the role of accurate reflection in our deliberative decision making.

*Doris’s positive framework*

Given Doris’s conservatism about our everyday attributions of agency and responsibility, it is surprising that he uses psychotherapy as a model for how collaboration and dialogue can facilitate agency. In the history of agential responsibility, psychotherapy has been around for a blink of an eye, and has been employed by a sliver of agents. So it is at best a device for highlighting what aspects of our common practices actually do facilitate agency. Dialogue and “positive alliance” are the agency-facilitating aspects of beneficial psychotherapy highlighted by Doris. But both phenomena are also present in collaborative enterprises where anti-agential forces often prevail. We review some below, but we encourage Doris to say more about what lessons to take from psychotherapy, as well as suggest additional cultural models for understanding what facilitates agency.

Doris argues that “agency requires … mutual influence” (148), the kind of influence involved in negotiating our relationships via exchange of “rationalizations” (in Doris’ non-pejorative sense) (153). While he makes a nice case that, at least some of the time, our own self-ignorance plays a necessary supporting role in such influence, we are less sanguine. For such influence is subject to forces which often do not facilitate the expression of values. Some means of social influence are often simply irrelevant or arbitrary from the perspective of our values. (Robert Cialdini’s (2008) classic *Influence* reviews the evidence for many, as well as the fact that they have been exploited far longer than they have been documented by social psychology.) Other forms that may sometimes be congruent with our values, such as reciprocity
(Gialdini et al 1975) or deference to authority (Milgram 1974), are also easily exploited to influence in ways that are not at all reflective of our values.

Several types of group processes can drive us toward actions that are in conflict with our values. Processes involving groupthink (Baron 2005; Janis 1982), group polarization (Moscovici & Zavalloni 1969; Sunstein 2009), intergroup phenomena (Sherif et al 1961; Tajfel et al 1971), power dynamics (Keltner, D. et al 2008; Galinsky et al 2006), etc., can all influence people to behave in ways they wouldn’t endorse in non-group or other-group settings. In all such cases, negotiated rationalizations are a part of, or are influenced by, these very processes. Indeed, “Animal Farm phenomena”, whereby groups or even whole societies become what they usually, or used to, condemn, come about partly through processes that, as far as we can tell, have not yet been excluded from Doris’ category of negotiated rationalizations. To be sure, Doris insists that the “right” kinds of social processes are needed for agency (125). But it is not yet clear how to specify those processes or how often they occur. Take the case of Milgram’s obedience experiments. Doris points out how they demonstrate that the presence of some dissent enables more (119). Comrades can help us act out our values. But note that the primary experimental paradigm itself fails, so far as we can tell, under the rubric of a collaborative social negotiation, complete with rationalizations (there is a sense in which the experimenter and subjects formed what Doris calls a “positive alliance”—the situation was set up as one in which they team together for the sake of advancing science). How can we know which social groups and processes undermine our agency? Should classic social psychological results be taught in schools (sounds good to us)? What about reflection on the results of our myriad interactions?

Doris’s excellent demonstration of the complexity and variation found in everyday (responsible) decision-making nicely exposes the problem with the traditional philosophical isms that Doris so patiently analyzes—and abandons: they all tend to be static and absolute, laying down presumably eternal policies of self-control that may look good on paper but nobody could live with. The fact is that human decision-making has always been something of an arms race (in the evolutionary sense), with novel techniques being introduced, identified and warned against, refined, further disarmed, etc., and all memorialized in the world’s folktales and literature, from the country mouse and city mouse to Othello and “hidden persuaders.” The arms race is intensified and accelerated by human reflection itself, as insightful critics and other observers expose the various ploys and vulnerabilities, and since we all have the desire to persuade others of our own values, we take on the lifelong goal of honing our talents in the game of reason-giving, trying to hold our own as responsible agents who can protect themselves from goofy (and other baleful) influences. Among the social scaffolds invented by arms races are measures like strict liability laws, which have the effect of emphasizing “due diligence” in areas of particular risk of harm by removing in advance otherwise reasonable excuses. Laws criminalizing the technically benign (concealed weapons, drug paraphernalia) or even “intent”
(from conspiracy to hate) are also part of the ongoing arms races that reflect the power of human reflection on social scaffolding.

Sometimes we become entangled in meta-strategizing and second-guessing that can stultify us, turning us into accomplished boxers whose footwork and feints dazzle our opponents while we never land a punch. As Doris explains, too much attention to how we are doing can prevent us from being at our best. Most of the time, the task is made easier by benign social scaffolding: we treat each other with “the benefit of the doubt” and this trust is itself a wise policy, provided that the balance is favorable. But there is no foolproof way of preventing us from falling into bad company, and then no pure “individualist” policy can be endorsed. As Doris puts it, “In point of fact, very little of what a person does, be it good or bad, is entirely up to the person herself.” (170).

References


