

Nothing in Their Heads

Debunking the Doyen et al. claims regarding the elderly-priming study

Published on March 5, 2012 by [John A. Bargh, Ph.D.](#) in [The Natural Unconscious](#)

Welcome back to "The Natural [Unconscious](#)," which returns after a hiatus of a couple years. Why we went underground for awhile is a topic for a later post; we'll get to that soon.

Scientific integrity in the era of pay-as-you-go publications and superficial online science

journalism. What prompts the return of the blog is a recent article titled "[Behavioral Priming: It's All in the Mind, but Whose Mind?](#)" by Stéphane Doyen, Olivier Klein, Cora-Lise Pichon, and Axel Cleeremans. The researchers reported that they could not replicate [our lab's 1996 finding](#) that priming (subtly activating in the minds of our college-age experimental participants, without their awareness) the [stereotype](#) of the elderly caused participants to walk more slowly when leaving the experiment. We had predicted this effect based on emerging theory and evidence that perceptual mental representations were intimately linked with behavioral representations, a finding that is very well established now in the field (see below). Following their failure to replicate, Doyen et al. went on to show that if the experimenter knew the hypothesis of the study, they were able to then find the effect. Their conclusion was that experimenter expectancies or awareness of the research hypotheses had therefore produced the effect in our original 1996 study as well—in other words, that there was no actual unconscious stereotype effect on the participants' behavior.

The Doyen et al. article appeared in an online journal, PLoS ONE, which quite obviously does not receive the usual high scientific journal standards of peer-review scrutiny (keep reading for the evidence of this); instead, the journal follows a "business model" in which authors pay to have their articles published (at a hefty \$1,350 per article). The journal [promises](#) a "rigorous peer review" for technical soundness but not as to the importance of the finding. On their website PLoS dismisses the use of knowledgeable editors to oversee what gets published and what does not, claiming this adds only a subjective element to the acceptance decision that can be biased against new research directions. But knowledgeable editors also can prevent articles from being published based on faulty peer reviews, such as by inexperienced, lazy, or biased reviewers. Expert editors also know the relevant theory and past research in a given domain, and also know of common methodological pitfalls that inexperienced researchers in the domain—such as, apparently, Doyen et al. (keep reading)—can fall prey to.

The lack of rigorous expert editorial scrutiny by PLoS in the Doyen et al. case means that I must supply it here, only after it has been published. If I'd been asked to review it (oddly for an article that purported to fail to replicate one of my past studies, I wasn't) I could have pointed out at that time the technical flaws, though these might not have mattered to PLoS ONE—as a for-profit enterprise, PLoS published 14,000 articles in the year 2011 alone. Fourteen *thousand*. Something tells me they don't turn down many \$1,350 checks...

Although the essentially self-published nature of the Doyen et al. article is bad enough, the misleading conclusions it drew were made even worse by the publicity given to them by some online science-journalism blogs, one of which posted about the Doyen et al. failure with the title "[Primed by expectations: Why a classic psychology experiment isn't what it seemed](#)". So as you have read this far, and are therefore a reader interested in psychological science, I'd like (and apparently need) to set the record straight.

There are two main reasons why the Doyen et al. conclusions, and those in science journalism blogs that swallowed their conclusions whole, are misleading and false. The first is that in our original 1996 research the experimenter was entirely blind both to experimental hypotheses and the experimental condition of the participant and did not even collect the main dependent variable of walking time down

the hall. Thus there is no way in our original 1996 studies that the effects could have been produced by experimenter expectations. The second is that Doyen et al. did not follow our original procedure (which is what one must do when attempting to replicate a study) but instead made critical changes that are known (in social psychology, which is not Doyen et al.'s field of specialization) to eliminate the stereotype-behavior effect. I'll take these points one at a time.

1. The experimenter was entirely blind to hypotheses and participants' experimental conditions in the 1996 elderly-priming study. As recounted [elsewhere](#), when we designed and ran the elderly stereotype priming studies (actually, back in 1991), we had solid theoretical reasons to hypothesize the effects we did obtain. But we were very careful to make sure that experimenter bias could not cause the effect. The experimenter who ran the participants in the study was blind to the study hypotheses, and all he did was to greet the participants and give them an envelope. He told the participants to go into the experimental room, open the envelope and follow the instructions inside. Moreover, he told the participants to leave the material inside the room. Thus the person who had actual contact with the participants in the elderly priming study never saw the priming manipulation (the scrambled sentence test), and certainly did not know whether the participant was in the elderly priming or the control condition.

The experimenter also did not collect the dependent measure of the study, which was how long the participant took to walk down the hallway leaving the experimental room. This measure was collected by a different person, a graduate student, who was posing as the next participant and sat waiting with her coat in her lap in a chair outside the experimental room. She had a stopwatch under the coat and surreptitiously started it when the participant came out of the experimental room, and stopped it when the participant disappeared around the corner at the end of the hallway.

Thus there is no possible way that experimenter expectations or any other form of experimenter bias could have produced our results in the original 1996 elderly priming study. We were fully aware of this potential problem at the time and took every precaution to make sure it did not contaminate or spuriously produce our results.

2. Doyen et al.'s replication of our 1996 study had many important differences from our procedure, all of which worked to eliminate the effect. When researchers attempt (in good [faith](#)) to replicate another lab's findings, they are supposed to follow the original procedure as closely as possible. But Doyen et al.'s procedure differed in many important respects from our original study, and three of their departures from our procedures were especially likely to mask or eliminate the stereotype priming effect, and so cause them not to find the effect.

First and most problematically, their experimenter called the participant's attention to walking down the hallway after the experiment, by giving them instructions to "go straight down the hall when leaving." We did no such thing, and just let the participant leave in the most natural way. It has long been known, and shown in many studies, that calling the participant's conscious attention to an automatically operating process can reduce or eliminate that automatic effect. The best known example of this in social psychology was a classic 1983 study by Norbert Schwarz and Gerald Clore (Note 1) in which participants were called on the phone and asked questions about how satisfied they were with their lives to this point. As the researchers hypothesized, the day's weather influenced these overall life satisfaction ratings—if it was a sunny day people rated their life as having gone better than if it was a cloudy day. But this effect was eliminated if the researchers, purportedly calling from a distant city, first asked the participant "how's the weather down there?" By calling the person's attention to the day's weather, making that factor salient to them, the automatic or [unconscious](#) influence of the weather on their life satisfaction ratings was eliminated. In the same way, Doyen et al.'s faulty procedure of calling their participants' attention to how they should walk down the hall could very well have eliminated the natural effect of the elderly [stereotype](#) activation.

The second gross departure from our procedure in the Doyen et al. study was the apparent over-use of elderly-related items in their priming manipulation, the standard "scrambled sentence test" in which participants construct a series of grammatically correct sentences out of sets of words presented out of order. As we took pains to point out in our [priming and automaticity methods chapter](#), using too many primes of the same category or theme can cause the participant to consciously notice that theme and, for the same reasons as above, eliminate or even reverse the natural priming effect. Again, this has been known for some time; in 1984 Paul Herr, Jim Sherman, and Russ Fazio (Note 2) showed that when using very salient, extreme primes of hostility (for example, Adolf Hitler and Dracula), participants considered the semi-hostile person they next read about to be less, not more, hostile than did participants in the control group, who were not exposed to those extreme primes. For this reason, in our methods chapter we suggested (and this is the practice we've always followed) using only around 10-12 critical primes out of the 30 total items. Unlike our original study as well as our methods recommendations, Doyen et al. apparently (it is not clear from their methods section but this seems to have been the case) included an elderly-related word in *all 30* of their priming items. This could well have been a second reason why they did not find the elderly-priming effect while we did.

The third major problem with the Doyen et al. study (and their conclusions) is that priming only works if there is something in the participant's head to prime in the first place. In other words, did their participants (ours did) have a mental stereotype of the elderly at all, and if they did, what was its content? The specific content of stereotypes of social groups vary from culture to culture; for example stereotypic beliefs about the elderly are quite different (more positive) in East Asia than they are in North America. Doyen et al. apparently did not check to make sure their participants possessed the same stereotype of the elderly as our participants did. Without taking the appropriate methodological steps to make sure participants hold these stereotypic beliefs in the first place, one can't assume the external primes have anything in the participants' minds to activate.

There were other deviations from our original procedure as well; in the Doyen et al. study but not in ours the experimenter was also the timer of how slowly the participant walked down the hallway. We instead had been very careful to keep the experimenter separate from any data collection. And try as I might, I could not find any mention of a control condition in their Experiment 2 (having a control condition is Methods 101, folks). While these are relatively minor problems compared to the Big 3 above, they serve as further evidence of Doyen et al.'s lack of methodological rigor, quite ironic, actually, given that they are accusing our lab (as they have accused other labs; see Note 3) of suffering from "methodological pitfalls."

Therefore, it is the Doyen et al. findings, not ours, that are ["not what they seem"](#). The take-home lesson for science bloggers and their readers is that there are other reasons for failures to replicate other than the invalidity of the original published study. Incompetent or ill-informed researchers performing the replication (such as in research domains outside of their areas of expertise) is another reason.

The stereotype priming of behavior effect has been widely replicated already and is also consistent with other established findings in psychology and [neuroscience](#). As for the stereotype priming effect itself, there should be no doubt by this point. It is well established in a variety of research domains. Stereotype priming effects are the cause of the well-established ["stereotype threat" effect](#) (see, for example, Notes 5 & 6) in which activating a person's racial or ethnic or [gender](#) group [identity](#) has deleterious effects on that person's subsequent academic or athletic performance. Stereotype threat effects have been documented, disturbingly, even in 5 year old children; when Nalini Ambady and colleagues (Note 7) primed preschool girls (in a Harvard day care center) with their female identity they do worse on math tests than they would if that identity (which contains the

stereotypic belief that "girls can't do math") was not primed. Priming effects on behavior, as in the elderly-priming study, have now been demonstrated in infants as young as 18 months old (Note 8). That behavior can be primed by acts of social perception, as are involved in stereotype activation (by the easily-perceived, distinguishing features of social groups), is also supported by neuroscience research demonstrating that witnessing another person's behavior activates relevant areas of one's own motor cortex (Note 9). This unconscious perception-behavior link is a natural support to human imitation and mimicry, as [Chartrand and I showed in 1999](#), and thus is a hugely helpful aid to the social learning of very young children. And there are now dozens of peer-reviewed (that is, by peers who actually know about and do research in this field) psychological journal publications in which stereotype priming effects on a variety of behaviors have been obtained (see reviews in Notes 10 & 11).

Like many scientists, I take the long view and continue to have [faith](#) in science as a cumulative process, with particular studies that vary in their methods and approaches converging on deeper underlying principles and mechanisms. No single experiment, standing by itself, should be the basis for concluding anything, maybe most especially in psychological research. And when a single study does not replicate another one whose findings are solidly embedded in theories of more than one scientific field and which is consistent with dozens if not hundreds of other conceptual replications, then responsible scientists—and responsible science journalists—do not rush to judgment and make claims that the entire phenomenon in question is illusory. Thomas Kuhn famously argued in *The Structure of Scientific Revolutions* that it is the accumulation of contrary evidence over time, not single studies, which is needed to overturn established concepts and principles in any branch of science. But I am not so much worried about the impact on science of essentially self-published failures to replicate as much as I'm worried about your ability to trust supposedly reputable online media sources for accurate information on psychological science.

Notes

1. Schwarz, N., & Clore, G., "Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states". *Journal of [Personality](#) and Social Psychology*, 45, 513-523.
2. Herr, P., Sherman, S. J., & Fazio, R. (1984), "On the consequences of priming: Assimilation and contrast effects". *Journal of Experimental Social Psychology*, 19, 323-340.
3. Waroquier, L., Marchiori, D., Klein, O., & Cleeremans, A. (2009), "Methodological pitfalls of the [Unconscious](#) Thought paradigm". *Judgment and [Decision Making](#)*, 4, 601-610.
4. Rosenthal, R. (1966). *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts.
5. Steele, C., & Aronson, J. (1995). "[Stereotype](#) threat and the intellectual test performance of African Americans". *Journal of Personality and Social Psychology*, 69, 797-811.
6. Stone, J., Perry, W., & Darley, J. (1997). "'White Men Can't Jump': Evidence for the Perceptual Confirmation of Racial Stereotypes Following a Basketball Game". *Basic and Applied Social Psychology*, 19, 291-306.
7. Ambady, N., Shih, M., Kim, A., & Pittinsky, T. (2001), "Stereotype susceptibility in children: Effects of [identity](#) activation on quantitative performance". *Psychological Science*, 12, 385-390.
8. Over, H., & Carpenter, M. (2009). "Eighteen-month-old infants show increased helping following priming with affiliation". *Psychological Science*, 20, 1189-1193.
9. Fadiga, L., Craighero, L., Olivier, E. (2005). "Human motor cortex excitability during the perception of others' action". *Current Opinion in Neurobiology*, 15, 213-218.
10. Dijksterhuis, A., & Bargh, J. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 1-40). San Diego: Academic Press.

11. Dijksterhuis, A., Chartrand, T., & Aarts, H. (2007). Effects of priming and perception on social behavior and goal pursuit. In J. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes*. New York: Psychology Press.