Evolution and Human Behavior 38 (2017) 279–280

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehbonline.org

Discussions

Validation is a Galilean enterprise

William H.B. McAuliffe, Michael E. McCullough*

University of Miami, United States

Frey (2017) used gameplay data from an online multi-player strategy game named Ikarium to test whether the behavioral strategies identified in the public goods game (PGG) also emerge in a setting where potential contributors to public goods interact repeatedly over long periods of time. In particular, Frey was interested in whether people who play Ikarium could be categorized either as "free riders" who rarely contribute to the common resource, "conditional cooperators" whose contributions are commensurate with others' contributions, or "high cooperators" who unconditionally contribute most of their endowment (Fischbacher, Gächter, & Fehr, 2001). In apparent support of the PGG's external validity, Frey found that Ikarium players manifest the same behavioral strategies as PGG players do, and that the proportion of free-riders in Ikarium is similar to their proportion in many PGG experiments. Frey also found lower proportions of both conditional cooperators and high cooperators in Ikarium than are typically found in laboratory PGGs. Taken together, these findings led him to conclude that the evidence for the PGG's external validity is mixed but encouraging.

We contend that Frey's analyses actually have little bearing on the external validity of the PGG. Evidence from recent experiments using modified versions of the PGG and stringent comprehension checks indicate that individual differences in people's tendencies to contribute to the public good are better explained by individual differences in participants' comprehension of the game's payoff structure than by individual differences in cooperativeness (Burton-Chellew, El Mouden, & West, 2016). For example, only free riders reliably understand right away that complete defection maximizes one's own payoff, regardless of how much other participants contribute. This difference in comprehension alone explains the so-called free riders' low PGG contributions. These recent results also provide a new interpretation of why conditional cooperators often contribute generously in early rounds, and then less in later rounds (Fischbacher et al., 2001). Fischbacher et al. (2001) attribute the relatively high contributions in the early rounds to cooperativeness and the subsequent decline in contributions to conditional cooperators' frustration with free riders. In reality, the decline in cooperation observed over the course of PGGs occurs because so-called conditional cooperators initially believe that their payoff-maximizing decision depends on whether others contribute, but eventually learn that contributing never benefits the contributor (Burton-Chellew, Nax, & West, 2015). Because contributions in the PGG do not actually reflect cooperativeness, there is no real-world cooperative setting to which inferences about contributions in the PGG can generalize.

We think Frey's error was in focusing on behavior rather than the motives underlying the behavior. To see why, consider Lewin's (1931) distinction between Aristotelian and Galilean social science. Lewin asserted that most of his contemporaries studied psychology in the same way that Aristotle studied physics: Their goal was to identify the essential surface features of particular phenomena shared by all instances of any particular phenomenon. For instance, until the rise of the Gestalt theory, perceptual psychologists apparently studied optical illusions as if they were all part of a common category, unified by the misperceptions they cause (Lewin, 1931). In contrast, Galileo categorized observable phenomena by the underlying laws of nature that caused them; whether the many effects of a law had similar observable consequences was irrelevant. Thus, a Galilean perceptual psychologist would study each optical illusion as a byproduct of particular functional structures in the vision system.

Frey's investigation was Aristotelian in spirit because he evaluated whether the behavioral types in Ikarium and the behavioral types in the PGG are members of the same set of categories. This approach cannot shed light on external validity, however, because external validity reflects the extent to which the psychological processes engaged in the original experimental setting produce the average outcome that would be observed across the entire set of situations that engage those psychological processes (Campbell & Stanley, 1963; Shavelson & Webb, 1991), not whether two settings that share structural features (but might activate different psychological processes) generate the same outcomes. Thus, no set of results from Frey's study could have vindicated the external validity of the PGG because individual differences in contribution behavior in Ikarium are (presumably) driven by cooperativeness, whereas individual differences in the PGG are primarily driven by comprehension of how to maximize one's own payoff.

Frey's error is by no means uncommon among researchers who seek evidence for the external validity of economic games. For example, Franzen and Pointner (2013) aimed to test whether the dictator game (DG) is an externally valid measure of fairness. They interpreted a modest positive association between DG offers and returns of misdirected letters as evidence of the DG's external validity. However, redirecting a letter to the correct recipient has less to do with fairness than with norms about lost property (West, 2003): Evidence from modified DGs in which property norms are relevant indicate that respect for what is fair and respect for property motivate different patterns of behavior (Oxoby & Spraggon, 2008). Consequently, the misdirected letter technique has little bearing on whether results from the DG generalize to real-life decisions about whether to divide resources fairly.

☆ Acknowledgments: Research was supported by a grant from the John Templeton Foundation (award no. 29165) to M.E. McCullough.

DOI of original article: http://dx.doi.org/10.1016/j.evolhumbehav.2016.11.004.

* Corresponding author.

E-mail address: mikem@miami.edu (M.E. McCullough).

http://dx.doi.org/10.1016/j.evolhumbehav.2016.11.003

1090-5138/© 2016 Elsevier Inc. All rights reserved.

The principles of Aristotelian science also figure prominently in measures of punishment derived from economic games. Many researchers have claimed that negative reciprocity (i.e., a willingness to punish non-cooperators, even at a personal net cost) partially explains humans' propensity for cooperating in relatively large groups of unrelated individuals (Fehr, Fischbacher, & Gächter, 2002; Gintis, 2000). The experimental evidence for this claim comes from observations of negative reciprocity in economic games (Fehr & Fischbacher, 2004), which is thought to sustain group cooperation in games with multiple rounds (Fehr & Gächter, 2002). Studies of negative reciprocity are inevitably Aristotelian because researchers count all situations in which individuals punish uncooperative others at a net cost as evidence of negative reciprocity, regardless of whether the same motives undergirded the different behaviors. For instance, researchers in the negative reciprocity literature consider both rejections of low offers in the ultimatum game (UG) and punishment of stinginess in the third-party punishment game (TPPG) as examples of negative reciprocity (Guala, 2012). However, punishment in the TPPG disappears after eliminating the influence of experimental demand (Kriss, Weber, & Xiao, 2016; Pedersen, Kurzban, & McCullough, 2013), whereas participants reject low offers even in modified versions of the UG where rejection does not affect the payoff of the proposer (Yamagishi et al., 2009). These findings suggest that a desire to appear moralistic explains punishment in the TPPG, while preoccupation with respect explains rejections in the UG.

The Aristotelian nature of negative reciprocity has sown mischief in researchers' attempts to gauge the external validity of punishment games. For instance, Guala (2012) reviewed the ethnographic record to assess whether negative reciprocity might have promoted cooperation in groups of ancestral humans. Guala found no evidence that hunter-gatherers engaged in negative reciprocity on behalf of non-relatives, but he did find abundant evidence that people enacted revenge when their own interests had been harmed. However, these vengeful acts typically undermined group cooperation (Guala, 2012), which does not support the hypothesized role of negative reciprocity in explaining the evolution of human cooperation. When laboratory results do not generalize to the settings in which the hypothesized psychological processes are theorized to operate, one should wonder whether researchers misidentified the psychological processes that are operative in the original laboratory setting. Guala did not consider this possibility because focusing on negative reciprocity limited him to theorizing only about behavior, even though it is psychological processes and not behaviors per se that generalize from one situation to another. Instead, Guala explained away the lack of correspondence between laboratory observations and field observations by positing that negative reciprocity is "artefactual insofar as it is produced by the specific experimental procedures, but nevertheless real because it does take place in a limited range of (laboratory-like) conditions" (2012, p.7). Guala concluded that in ecologically valid settings negative reciprocity motivates supporting institutions that curb free-riding while keeping the costs of punishment low. In reality, assessing the external validity of games that involve punishment on behalf of unrelated others is pointless because they do not measure punitive sentiment anyhow.

Perhaps researchers who evaluate the external validity of economic games make the mistake of practicing Aristotelian science because they are primarily interested either in how cooperative behavior evolved or in how cooperative behavior relates to game theory. Game theorists often explicitly eschew references to psychological states, and ultimate explanations of behavioral traits do not involve any characterization of proximate psychological mechanisms. Fewer cooperation researchers are interested primarily in the psychological structures that have been shaped by evolutionary processes to adaptively regulate behavior. However, as Tooby and Cosmides (1992) emphasized, understanding the evolved psychological structures that generate behavior in concert with environmental input is crucial to understanding behavior and identifying the selection pressures that made the behavior possible. Only then will behavior cease to appear "endlessly variable" (Tooby & Cosmides, 1992, p. 64)—or "context-specific," as Frey says in many places— and begin to instead conform to psychological regularities which researchers can in turn elucidate using principles from evolutionary biology.

References

Burton-Chellew, M. N., El Mouden, C., & West, S. A. (2016). Conditional cooperation and confusion in public-goods experiments. *Proceedings of the National Academy of Sciences. Vol. 113(5)*. (pp. 1291–1296).

Burton-Chellew, M. N., Nax, H. H., & West, S. A. (2015). Payoff-based learning explains the decline in cooperation in public goods games. *Proceedings of the Royal Society of London B: Biological sciences. Vol. 282 (1801)*. (pp. 20142678).

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, *25*(2), 63–87.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, *13*(1), 1–25.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, *71*(3), 397–404.

Franzen, A., & Pointner, S. (2013). The external validity of giving in the dictator game. *Experimental Economics*, *16*(2), 155–169.

Frey, U. J. (2017). Cooperative strategies outside the laboratory—Evidence from a long-term large-N-study in five countries. *Evolution and Human Behavior*, *38*(1), 109–116.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, *206*(2), 169–179.

Guala, F. (2012). Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences*, *35*(1), 1–15.

Kriss, P. H., Weber, R. A., & Xiao, E. (2016). Turning a blind eye, but not the other cheek: On the robustness of costly punishment. *Journal of Economic Behavior & Organization*, *128*, 159–177.

Lewin, K. (1931). The conflict between Aristotelian and Galilean modes of thought in contemporary psychology. *The Journal of General Psychology*, *5*(2), 141–177.

Oxoby, R. J., & Spraggon, J. (2008). Mine and yours: Property rights in dictator games. *Journal of Economic Behavior & Organization*, *65*(3), 703–713.

Pedersen, E. J., Kurzban, R., & McCullough, M. E. (2013). Do humans really punish altruistically? A closer look. *Proceedings of the Royal Society of London B: Biological sciences. Vol. 280(1758)*. (pp. 20122723).

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer. Vol. 1*.Sage Publications.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind*. (pp. 19–136) New York: Oxford University Press.

West, M. D. (2003). Losers: Recovering lost property in Japan and the United States. *Law & Society Review*, *37(2)*, 369–424.

Yamagishi, T., Horita, Y., Takagishi, H., Shinada, M., Tanida, S., & Cook, K. S. (2009). The private rejection of unfair offers and emotional commitment. *Proceedings of the National Academy of Sciences. Vol. 106(28)*. (pp. 11520–11523).