Judgment and Decision Making, Vol. 10, No. 3, May 2015, pp. 280–282

Judging competing theoretical accounts by their empirical content and parsimony: Reply to Myrseth and Wollbrant (2015)

Pascal J. Kieslich* Benjamin E. Hilbig†

Abstract

Myrseth and Wollbrant (2015) offer an alternative theoretical explanation for our finding that defection entails more cognitive conflict than cooperation (Kieslich & Hilbig, 2014). Although we completely agree that different theoretical explanations for a result are possible, we maintain that the theoretical approach we tested (Rand et al., 2014) is parsimonious and falsifiable, excluding certain plausible results a priori. By comparison, the alternative framework proposed by Myrseth and Wollbrant requires several debatable assumptions to account for our findings, rendering it the more complex theory. Besides, their framework as a whole could have accounted for any possible finding in our experiment, making it impossible to falsify it with our data. We thus conclude that the notion by Rand et al.—that there is a spontaneous disposition to cooperate—has more empirical content while requiring fewer assumptions.

Keywords: social dilemma, cooperation, cognitive conflict, intuition, self-control, theory testing, empirical content.

In their comment on our original article (Kieslich & Hilbig, 2014), Myrseth and Wollbrant (2015, M&W in what follows) raise an important general issue concerning the conclusiveness of empirical findings. Whereas their arguments do not question that the methodology and analyses presented in our article indeed test (and support) the hypothesis that defection entails more cognitive conflict than cooperation, M&W offer an alternative explanation for the empirical pattern based on their own theoretical framework. They argue that different assumptions within their framework—assumptions that run contrary to the theoretical assumptions underlying our hypothesis—could predict the same empirical pattern. Here we will very briefly reiterate our line of reasoning and then discuss M&W's alternative theoretical framework.

From recent theoretical arguments by Rand and colleagues that cooperation is the effortless and spontaneous response in social dilemmas (Rand, Greene & Nowak, 2012; Rand et al., 2014), we derived the prediction that defection should entail more cognitive conflict—as measured by the curvature of response trajectories—than cooperation. Our findings mirrored the hypothesis, showing that the attraction of the non-chosen option was stronger when people defected—which held particularly for individuals with a stronger dispositional tendency to make cooperative choices. More importantly, the theoretical framework by Rand et al. could not have accounted for the opposite pattern (more conflict for cooperative choices) or a null effect (equivalent conflict for defection and cooperation). As such, the hypothesis derived from Rand et al.'s framework is specific and was strictly tested in the sense that it could have been falsified.

However, as is well known from the philosophy of science, the fact that empirical data are in line with a theory (even though it could have failed) is merely necessary but not sufficient for the "truth" of said theory (Popper, 2005)—simply because it is usually possible to generate various theoretical explanations for the same finding. Note that we meant to imply this in the discussion of our original article in stating that "although the different curvatures of response trajectories for cooperation and defection decisions support the idea that cooperation is the spontaneous and less conflicting response in social dilemmas, they do not necessarily imply that there are actually two distinct systems (intuition and reflection)—as assumed by Rand et al. (2012)—that interact to produce decisions in social dilemmas" (Kieslich & Hilbig, 2014, p. 519). Whereas M&W's selective citation of part of this sentence lets our reasoning appear to be inductive, it was clearly not our intention to make conclusive inferences regarding the "truth" of the theoretical framework by Rand et al. We merely sought to test a specific prediction derived from it and found support (in that the empirical data was in line with—and hence did not falsify—the prediction). Nonetheless, other theoretical models may just as well allow for the observed empirical pattern to occur and thus be viable alternatives.

Consequently, there is nothing inherently worrisome (or

*Department of Psychology, School of Social Sciences, University of Mannheim, Schloss Ehrenhof Ost, 68131 Mannheim, Germany. Email: kieslich@psychologie.uni-mannheim.de.

†Cognitive Psychology Lab, Department of Psychology, University of Koblenz-Landau, Germany.

indeed surprising) about M&W's assertion that our findings could also be accounted for by a set of assumptions that are, in essence, incompatible with the theory we intended to test. So, it seems worthwhile to consider the alternative account offered. As an alternative theory, M&W put forward a framework that conceptualizes decisions in social dilemmas as a self-control problem (Martinsson, Myrseth & Wollbrant, 2014). Specifically, decisions are assumed to be the outcome of a two-stage process (see Figure 1 in their comment): First, people may not identify a self-control conflict and follow their impulse (in M&W's framework referred to as "reflexive, default behavior"). If people do identify a self-control conflict, they may next try to resist their impulse. They may either fail to do so, thus following their impulse (i.e., choosing the default), or exercise successful self-control, thus following their deliberative goal instead. Which behavior constitutes the default (cooperation vs. defection) is assumed to be context-dependent (Martinsson et al., 2014), but it remains entirely open which context would apply to our experiment.

To derive predictions for cognitive conflict, M&W further make the following assumptions: Following the default because no self-control conflict was identified leads to a "no conflict experience". Identifying a self-control conflict leads to a "conflict experience" in which case either self-control failure (sometimes also called unsuccessful resistance to impulse by M&W) or successful self-control may lead to a relatively higher degree of conflict (both cases are considered to be equally plausible in M&W's theoretical framework). M&W argue that, according to their framework, six different scenarios are possible, and that two of these are compatible with the results of our experiment. As one of these scenarios assumes that cooperation is the majority default (which is compatible with the theory by Rand et al.) but the other presumes that defection is the default, M&W conclude that our findings do not allow for any inference whether cooperation or defection is the default behavior.

However, the scenario in which defection is the default option does require a number of additional assumptions in order to fit our results. First, M&W assume that self-control failure (leading to defection) is associated with more cognitive conflict than successful self-control (leading to cooperation). However, M&W explicitly state that the opposite (more conflict for successful self-control than failed resistance) is equally plausible.

Moreover, according to M&W's framework a defection choice can be reached by two routes when assuming that defection is the default: People can either directly follow their impulse to defect (without identifying a self-control conflict) or they can identify a self-control conflict and fail to resist (and thus choose defection). Importantly, when directly following the impulse, defection should be accompanied by a "no conflict experience" (see Figure 1 in M&W). Yet, this would run contrary to the results in our study that M&W are trying to explain (more conflict for defection than cooperation). To account for this, M&W additionally assume "that the difference in conflict [between self-control failure and successful self-control] more than offsets the low levels of conflict exhibited by those individuals who, without resisting (see Figure 1), act on the[ir] impulse" (Footnote 2 as referred to in M&W's Table 1, squared brackets added). However, we believe that it is questionable to assume that the difference in conflict between self-control failure and successful self-control is so large that it offsets a "no conflict experience"—especially since M&W themselves consider a reversed pattern (more conflict for successful self-control than self-control failure) to be equally plausible. Importantly, even if one assumes that the amount of cognitive conflict for successful self-control and self-control failure is comparable, the additional impulsive choices for defection would lead to the prediction that cooperation is associated with more conflict than defection (i.e., the prediction would run contrary to the observed empirical pattern).

In other words, the proposed alternative account by M&W requires additional untested and, in our view, debatable assumptions. If, in line with Occam's razor, the "goal of model selection is to choose the simplest (i.e., least complex) model that describes the data well (i.e., descriptive adequacy)" (Myung & Pitt, 1997, p. 79), it is our impression that M&W's alternative theoretical explanation is more complex (i.e., less parsimonious) than the theory by Rand et al.

Besides, M&W do not make any predictions as to the overall relative frequency of cases in which a conflict is identified or not—although this is clearly relevant for the predicted results pattern. Only when talking about our finding that individuals high in Honesty-Humility (HH) experience more conflict when defecting vs. cooperating than individuals low in HH, M&W assume that the likelihood of identifying a self-control conflict differs: M&W assume that individuals low in HH should have a low likelihood while individuals high in HH should have a high likelihood. However, based on this assumption people low in HH should show less conflict when defecting even when compared to individuals high in HH that cooperate (since no conflict identification should be accompanied by no conflict experience). Yet, this is incompatible with the results we reported (see Figure 4; Kieslich & Hilbig, 2014).

Finally, M&W argue that by altering the assumptions within their framework one can actually explain every possible empirical pattern in our experiment (more conflict for cooperation, more conflict for defection, and no difference between the two). However, M&W do not provide any theoretical prediction as to when which of their assumptions should hold. As a consequence, there is no empirical pattern in our experiment that could falsify M&W's theoretical framework. It thus does not have sufficient empirical con-

tent[1] (see also Glöckner & Betsch, 2011) in the context of our study, at least not without further theoretical specification. (See Postscript by Jonathan Baron who points out that more specific predictions for M&W's framework can be derived, which are only partly supported by our data.)

In sum, we agree with M&W that other theoretical explanations for our result are possible. Nonetheless, we test a theoretical approach that is falsifiable, excluding certain plausible results a priori—as any useful theory should (Platt, 1964). Consequently, its descriptive adequacy indeed speaks for the theory. By comparison, the alternative framework proposed by M&W requires several additional (and debatable) assumptions to account for the findings (especially when assuming that defection is the default behavior). Moreover, it could have accounted for *any* finding in our experiment. We would of course welcome any suggestion for a critical experiment that can disentangle the different theoretical explanations and conclusively test them against each other. For the time being, we conclude that the approach of Rand et al. and the explanation assuming a spontaneous disposition to cooperate has more empirical content while requiring fewer assumptions, and, consequently, being more parsimonious than M&W's alternative framework.

References

Glöckner, A., & Betsch, T. (2011). The empirical content of theories in judgment and decision making: Shortcomings and remedies. *Judgment and Decision Making, 6*(8), 711–721.

Kieslich, P. J., & Hilbig, B. E. (2014). Cognitive conflict in social dilemmas: An analysis of response dynamics. *Judgment and Decision Making, 9*(6), 510–522.

Martinsson, P., Myrseth, K. O. R., & Wollbrant, C. (2014). Social dilemmas: When self-control benefits cooperation. *Journal of Economic Psychology, 45*, 213–236.

Myrseth, K. O. R., & Wollbrant, C. (2015). Less cognitive conflict does not imply choice of the default option: Commentary on Kieslich and Hilbig (2014). *Judgment and Decision Making, 10*(3), 277–279.

Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Review, 4*(1), 79–95.

Platt, J. R. (1964). Strong inference. *Science, 146*(3642), 347-353.

Popper, K. R. (2005). *The logic of scientific discovery*. London: Routledge.

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature, 489*(7416), 427–430.

[1]"[T]he empirical content of a statement increases with its degree of falsifiability: the more a statement forbids, the more it says about the world of experience" (Popper, 2005, p. 103).

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., et al. (2014). Social heuristics shape intuitive cooperation. *Nature Communications, 5*, 3677.

Postscript by Jonathan Baron

I think that one point has not been explicitly recognized by the original Kieslich/Hilbig paper, by M&W, or by this reply to M&W. In my view, the most plausible interpretation of the mouse track is that it reflects, with error, the current status of the decision process in real time. Thus, a curved track not only indicates "conflict" but also measures timing, insofar as we take the track at face value. When the track bends towards cooperation on its way to a defection response, this can indeed be understood as representing conflict at that point, but, in the end, the conflict is resolved in favor of defection. Thus, for this bending to happen, the conflict must come *before* the process that pulls the final response toward defection.

Think of the track as representing the balance of accumulated evidence in a drift diffusion model. And then suppose, contrary to my timing assumption, that the conflict is present with equal strength throughout the interval, but the conflict is greater when the subject is defecting. Then the balance will be a straight line for both cooperation and defection responses, but the straight line for defection will have a lower absolute slope. This will (other things being equal) make the defection response take longer, but curvature will be absent. Curvature can thus result only when the cooperation response is *temporarily* strong, so that it slows down the drift toward defection but then loses its power over time.

In sum, the original results indicate, within the assumptions of the method, that, when most subjects defect, they first experience some pull toward cooperation, but when they cooperate they do not usually experience a comparable pull toward defection early in the course of deciding. This means that a pull toward cooperation tends to come early in the decision process, more often or more strongly than does a pull toward defection.

By M&W's account, a cooperative response would result from an initial tendency and corresponding movement toward defection, followed by a movement toward cooperation resulting from self-control. And a defection response would result from the same initial tendency to defect, followed by a move toward cooperation resulting from self-control, then followed by a move toward defection as the self-control failed. The former sequence would result in a mouse track that is first curved toward defection, then moving to cooperation. The latter sequence would result in a tendency toward an s-shaped mouse track. For both cases, the prediction is contrary to the dominant pattern in the data, where no initial movement toward defection but instead a movement toward cooperation is observed.

Such an early pull toward cooperation could come from a self-control effort, following an initial impulse to defect. But the self-control would have to start so early that the initial impulse to defect is not seen in the mouse track when subjects choose cooperation. Such a failure to register the initial impulse to defect could happen if the track did not capture the subject's earliest impulse, contrary to the assumption of the method. Or, it could happen if the self-control process pulling toward cooperation started so soon that it would itself be intuitive and immediate. It would be "self-control" only in the sense that it was acquired from repeated practice, until it became (almost) automatic.