

## **What is a habit?**

Authors: *"In short: it is not meaningful to identify "habits" by particular empirical demonstrations of a "goal-independency"*

And yet, that is the definition of a "habit", across behavioral, neural, and computational literatures. If you are talking about something else, then use a different term.

Authors: *"Rather, habits should be defined procedurally (e.g., as cue-dependent action tendencies in PIT tasks)."*

Just because some researchers have, apparently, made that argument, that does not warrant the adoption of a definition that is inconsistent with a vast literature, particularly in a paper that does not provide any empirical or theoretical evidence regarding the nature of specific PIT or habits.

## **What is this paper about?**

Authors: *"The reviewer is treating our hypothesis of increased dACC activation as if it was a commonplace assumption shared with other influential theories of PIT"*

Your hypothesis, as far as I can tell, has nothing to do with PIT, or with outcome devaluation, per se. It is simply using the paradigm to test whether dACC implements cognitive control when it is costly to not do so. In other words, the use of specific PIT and outcome devaluation is incidental. As you note:

*"Our research question is whether cognitive control processes (or more precisely, their neural implementation as hypothesized by EVC theory) become engaged when the expected payoff of engaging in cognitive control is high."*

And herein lies the crux: While the question is *not* about PIT or habits, the bulk of the introduction is devoted to describing those paradigms; conversely, very little is said about highly relevant previous demonstrations of the role of dACC in cognitive and motivational control (e.g., Kounieher et al., 2009).

## **What is the appropriate control?**

Authors: *“Our research design includes several controls (see Statistical Hypotheses 1-3 in our response to R1 above and in the revised manuscript pp. 21-22)”*

The fact that some responses are associated with non-devalued outcomes does not control for the possibility that the dACC responds to 1) cues that signal recently changed outcomes, or any increase in the variability of the cue-outcome relationship, irrespective of instrumental processes, or 2) instrumental responses with changed outcomes, such that the mere presentation of the relevant response option (R1) would elicit post-devaluation dACC activity, without any Pavlovian phases at all.