Increasing resolution in the mechanisms of resolve

Article *in* Behavioral and Brain Sciences · April 2021

DOI: 10.1017/S0140525X20000801

2 authors:

Adam Bulley
Harvard University
38 PUBLICATIONS **631** CITATIONS
SEE PROFILE

Daniel L Schacter
Harvard University
591 PUBLICATIONS **88,528** CITATIONS
SEE PROFILE

Commentary on Ainslie, G., (2021) "Willpower With and Without Effort":

Increasing resolution in the mechanisms of resolve

Adam Bulley[1*] & Daniel L. Schacter[2]

[1] Department of Psychology, Harvard University, Cambridge, MA 02138, USA; The University of Sydney, School of Psychology and Brain and Mind Centre, NSW 2050, Australia. Email: adam_bulley@fas.harvard.edu; Web: http://adambulley.org/

[2] Department of Psychology, Harvard University, Cambridge, MA 02138, USA. Email: dls@wjh.harvard.edu; Web: https://scholar.harvard.edu/schacterlab/home

*Corresponding author

Abstract: 53 words
Main Text: 997 words
References: 301 words
Entire Text: 1482 words

Abstract

Ainslie offers an encompassing and compelling account of willpower, though his big-picture view comes occasionally at the cost of low resolution. We comment on ambiguity in the metacognitive and prospective mechanisms of resolve implicated in recursive self-prediction. We hope to show both the necessity and promise of specifying testable cognitive mechanisms of willpower.

Main text

While Ainslie frames resolve in terms of game-theoretic intertemporal bargaining, he leaves the cognitive and neural instantiation of resolve at times underspecified. In part, this is because the empirical evidence is wanting – as he acknowledges – but it is also because, by design, game-theoretical accounts remain agnostic about underlying mechanisms. In a prisoner's dilemma, the rules of the game and its payoff matrix are similar whether the agents involved happen to be bacteria or bankers. Nonetheless, we think there are costs associated with low resolution in the proximate mechanisms of resolve, as well as promising routes forward if proposals concerning the nature of these mechanisms can be sharpened up. We attempt to demonstrate these points of constructive clarification in the context of the *metacognitive* and *prospective* mechanisms implicated in "recursive self-prediction" that Ainslie argues forms the basis of resolve.

As a starting point, we take it as a given that humans don't consistently think through their intertemporal trade-offs with the kind of game-theoretic bargaining logic that observers can attribute to them. Ainslie acknowledges that the intertemporal bargaining of resolve could indeed happen below the level of self-awareness, or without any explicit representation at all. In fact, he suggests that the recursive self-prediction underpinning resolve might operate through "explicit self-enforcing contracts", via "vague awareness", perhaps "displaced away from any explicit self-knowledge", or even as purely "implicit contracts". It is therefore unclear how much "self" we should expect to find in "self-prediction."

One cost of this low specificity in the metacognitive mechanisms of resolve is that it leaves Ainslie's model resistant to disconfirmation in the face of new evidence. For instance, any failures to find recursive self-prediction in the implementation of resolve could be explained away by shuttling the relevant level of explanation around inside the mind of the resolver. Suppose that, upon a careful experimental investigation, we find that participants report resolving to delay their gratification for a later payoff simply because they foresee the long-term benefits of doing so, absent any anticipation of their own future behavior. In such a case, the enforcement mechanism that maintains an intention against lapses could be the anticipated negative costs of the smaller, sooner reward option. For instance, to answer Ainslie's question, "*Why not eat this piece of chocolate – it will barely show?*" a non-self-predictive resolver might answer, "*because I foresee even the small damage of a single piece as sufficiently costly, however tempting*". Under Ainslie's view, could we not explain away this finding by arguing that the underlying logical structure of the participant's decision-making is nonetheless one of

game-theoretical self-predictive bargaining, even if the participants themselves are not aware of it and would opt to explain their own decision-making differently?

The "prediction" portion of "self-prediction" is likewise somewhat ambiguous. Ainslie argues that because resolve is "a matter of framing and monitoring choices", it "might not be accompanied by measurable brain activity any more than *other semantic content is*" [our emphasis]. Elsewhere, though, Ainslie suggests instead that "scenarios created in episodic memory might also serve this function [of formulating and monitoring the intertemporal bargains that form resolve]."

These alternatives lead to various questions that could be productively reformulated as testable hypotheses. Does one need to actually imagine oneself failing in the future to adhere to a "no alcohol on weeknights" rule in order to implement the resolve to put down the Shiraz, as an episodic simulation account would entail? Is it enough to simply "know", in semantic terms, that one is more likely to fail in the future if one fails now? Situating resolve amidst existing frameworks of prospective cognition and deliberation could carve out a space for empirical steps forwards (see Bulley & Schacter, 2020; Szpunar et al., 2014).

For instance, we might test the evidence accumulation process by which people generate whatever predictions are central to resolve. Ainslie describes the act of reneging on a rule as if it constitutes a piece of empirical evidence that people use to anticipate their own future behaviors. But how so? One possibility is that episodic memories of reneging serve as raw material in the constructive episodic simulation of one's behavior in facing future willpower challenges. Convergent lines of evidence support the proposal that episodic future simulation operates via the recombination of episodic details from memory (Schacter et al., 2007; Suddendorf & Corballis, 2007), with a common *core network* of brain activity supporting remembering the past and imagining the future (Benoit & Schacter, 2015). Accordingly, if Ainslie's "recursive self-prediction" is a constructive process that samples episodic memories to inform anticipated behaviors, we should hypothesize that resolve will be associated with activity in this core network, similar to when participants directly retrieve episodic memories of willpower failures.

Research on prospection may also help to accommodate the idea that both semantic and episodic processes are sufficient for resolve in different contexts. The development of "good habits" that Ainslie equates to the successful operation of resolve may involve shifting contributions along a gradient of semantic and episodic processes (Irish & Vatansever, 2020; Szpunar et al., 2014). For instance, episodic simulation might be required to get resolve "off the ground", but after repeated (successful) instances, resolve could be eventually implemented in entirely semantic terms (for a similar suggestion about external precommitment see Bulley & Schacter, 2020). In this case, we should hypothesize that people with hippocampal damage who have deficits in the ability to imagine the future (Schacter et al., 2017) would be less capable of *initiating* intertemporal resolve in Ainslie's terms – but perhaps less impaired when it comes to maintaining "good habits" once these have been established (see Bakkour et al., 2019; Kwan et al., 2012; Palombo et al., 2015).

In the foregoing, we have pointed out some costs associated with ambiguities in Ainslie's otherwise encompassing big-picture account of willpower. We have provided some examples where pinning down specific mechanisms leads to testable predictions, focusing on the nature of the metacognitive and prospective mechanisms involved in recursive self-prediction where increased clarity would be perhaps most instructive.

Conflicts of interest: None

Funding statement: AB is supported by an Australian National Health and Medical Research Council CJ Martin Biomedical Fellowship APP1162811 (GNT1162811). DLS is supported by by National Institute of Mental Health grant R01 MH060941 and National Institute on Aging grant R01 AG008441.

References

Bakkour, A., Palombo, D. J., Zylberberg, A., Kang, Y. H., Reid, A., Verfaellie, M., Shadlen, M. N., & Shohamy, D. (2019). The hippocampus supports deliberation during value-based decisions. *ELife*, *8*, 1–28. https://doi.org/10.7554/elife.46080

Benoit, R. G., & Schacter, D. L. (2015). Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. *Neuropsychologia*, *75*, 450–457. https://doi.org/10.1016/j.neuropsychologia.2015.06.034

Bulley, A., & Schacter, D. L. (2020). Deliberating trade-offs with the future. *Nature Human Behaviour*, *4*, 238–247. https://doi.org/https://doi.org/10.1038/s41562-020-0834-9

Irish, M., & Vatansever, D. (2020). Rethinking the episodic-semantic distinction from a gradient perspective. *Current Opinion in Behavioral Sciences*, *32*, 43–49. https://doi.org/10.1016/j.cobeha.2020.01.016

Kwan, D., Craver, C. F., Green, L., Myerson, J., Boyer, P., & Rosenbaum, R. S. (2012). Future decision-making without episodic mental time travel. *Hippocampus*, *22*(6), 1215–1219. https://doi.org/10.1002/hipo.20981

Palombo, D. J., Keane, M. M., & Verfaellie, M. (2015). How do lesion studies elucidate the role of the hippocampus in intertemporal choice? *Hippocampus*, *25*(4), 407–408.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*, *8*(9), 657–661. https://doi.org/10.1080/08995600802554748

Schacter, D. L., Addis, D. R., & Szpunar, K. K. (2017). Escaping the Past: Contributions of the Hippocampus to Future Thinking and Imagination. In D. E. Hannula & M. C. Duff (Eds.), *The Hippocampus from Cells to Systems: Structure, Connectivity, and Functional Contributions to Memory and Flexible Cognition* (pp. 439–465). Springer. https://doi.org/10.1007/978-3-319-50406-3

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, *30*(3), 299–351. https://doi.org/10.1017/S0140525X07001975

Szpunar, K. K., Spreng, R. N., & Schacter, D. L. (2014). A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition. *Proceedings of the National Academy of Sciences*, *111*(52), 18414–18421. https://doi.org/10.1073/pnas.1417144111