



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/161892/>

Version: Accepted Version

Article:

Stafford, T. (2020) Evidence for the rationalisation phenomenon is exaggerated.
Behavioral and Brain Sciences, 43. e48. ISSN: 0140-525X

<https://doi.org/10.1017/s0140525x19002085>

This article has been published in a revised form in Behavioral and Brain Sciences
<https://doi.org/10.1017/S0140525X19002085>. This version is free to view and download
for private research and study only. Not for re-distribution, re-sale or use in derivative
works. © The Author(s), 2020. Published by Cambridge University Press.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Target article: Cushman, F. (in press). Rationalization is rational. *Behavioral and Brain Sciences*

Evidence for the rationalisation phenomenon is exaggerated

Tom Stafford
Department of Psychology
University of Sheffield, UK

t.stafford@sheffield.ac.uk
<http://tomstafford.staff.shef.ac.uk/>

Abstract

The evidence for rationalization, which motivates the target article, is exaggerated. Experimental evidence shows that rationalization effects are small, rather than gross and, I argue, largely silent on the pervasiveness and persistence of the phenomenon. At least some examples taken to show rationalization also have an interpretation compatible with deliberate, knowing, reason-responsiveness on the part of participants.

The evidence for rationalization, which motivates the target article, is exaggerated. There are two sources for this. First, it is an outcome of structural features of the experimental psychology tradition, which isolates effects using experimental control, and then uses null-hypothesis significance testing to establish their reality (i.e. their non-zero size) at the neglect of gauging their importance (Stafford, 2014). The second source of this exaggeration is the rhetoric of psychologists, who make hay out of emphasising the supposedly irrational aspects of our behaviour, de-emphasising reason-responsiveness (Stafford, 2015).

With respect to the target article, we can see this in the motivating section (1.1, “rationalization”, p7 ff). Rationalization is presented as pervasive (“people rationalize all the time”), solidly evidenced (“exhaustively documented”) and leading to “gross

errors” which are “stubbornly irrational”.

A study recruited as a key illustration of rationalization is Sharot et al (2010), but when we look at this example, we see that the largest effect reported in this paper was an average within-subject change of ~ 0.07 on a 6 point scale (experiment 1, see Figure 1, $t(20)=2.4, p<0.03$), so the rationalization manipulation produced a mean shift of $\sim 1\%$ in people’s judgements. Hardly gross or stubborn.

Other examples cited by the target article are similar - showing small movements in people’s ratings of belief, rather than flips from one belief state to another. This contrasts with the rhetorical portrayal of rationalisation. Participants in Brehm’s (1956) study showed an average change of ~ 0.9 on an 8 point scale (ie an $\sim 11\%$ between-groups shift, $p<0.01$, with $n\sim 30$ in each group). The experiment design, analysis and presentation of results presented by Vinckier et al (2019) does not make a simple estimation of effect size for rationalization obvious, but reading Figure 3 it looks like the within subject effect size of choice (i.e of rationalisation) is ~ 0.15 of a standardised (z) score. A “small effect”, as classically determined (Cohen, 1992).

Some might view this as unproblematic. A reasonable view is that rationalization is indeed common and commonly produces “gross errors”, but the methods of experimental psychology mean we can only hope to consistently capture rationalisation in the proxy-form of meagre shifts on rating scales. But it is also a reasonable view, I believe, that the extant evidence for rationalization does not support the grand claims for the power of the phenomenon. It is not enough that an idea be intuitively plausible.

The target article also invokes classic and widely known studies in psychology as evidence of rationalisation. Given space it is not possible to review all of them, but it is instructive to pick a couple of salient examples. Cognitive dissonance is the first citation of the target article (Festinger, 1962). The foundational demonstration of cognitive dissonance is Festinger and Carlsmith (1959). Taking the largest effect reported in this study, participants’ ratings of “enjoyability of task”, this showed a mean shift of 1.40 on an 11 point scale (i.e. $\sim 13\%$ change, between groups, $p<0.03$, $t(38)=2.22$, with $n=20$ in each group).

Another foundational contribution is Nisbett and Wilson (1977), for which (surely) the most discussed study is Wilson and Nisbett (1978, experiment 2), in which shoppers were asked to choose between four identical pairs of stockings. Famously, shoppers preferred the stockings on the right ($p < 0.025$, $n = 52$), but gave reasons other than position for their preference. Here my issue is not the size of the effect, but of its interpretation. While the explanations offered could be due to rationalization, failure to report reasons is not the same as their inaccessibility. Stafford (2014) argues that, where participants are ignorant of the conditions by which experimenters analyse their behaviour, failure to report those conditions is wholly compatible with rational choice.

Note my argument is not about the reproducibility, or not, of the evidence base, but rather of its interpretation.

It is telling that other motivating examples evoked by Cushman are fictional and/or psychologically exceptional (amnesia in the Bourne Identity, split-brain patients). While these are highly suggestive of the human potential of rationalization they cannot be taken as evidence that rationalization is pervasive in ordinary cognition.

I have argued that errors due to rationalization are often small, rather than gross. The interpretive slippage between experimental effects at single time points, and unusual edge cases (like split brain patients) means that the evidence for the persistence ('stubbornness') of rationalization is simply unclear. Further, while the target article notes that reasons are a factor in driving action, alongside rationalization, it is possible that some evidence presented as demonstrating rationalization could actually be showing rational behaviour (so, for example, in Wilson and Nisbett, 1978, it could be viewed as rational to give and defend an implausible answer if a psychologist asks you an impossible question such as which pair of identical stockings you prefer).

Where, then, does this leave the Cushman "representational exchange" account? I am not claiming that rationalization doesn't exist, only that the evidence psychology has produced in support of it is far less strong than is commonly supposed. There is still - potentially - something for Cushman's account to explain. My criticism highlights an opportunity: one test of the value of Cushman's account is if it can provide experimentalists with the leverage to produce better evidence of rationalization; that is,

the value of the account could be in allowing us to predict when and how rationalization will be at its strongest and most divergent from simple rationality in which actions are motivated by consciously-accessible reasons.

Perhaps, guided by better theory, experimentalists will be able to generate manipulations which show rationalisation in the lab consonant with our intuitions of its importance in our everyday lives.

REFERENCES

Brehm, J. W. (1956). Post-decision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology* 52(3), 384.

Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155-159.

Festinger, L. (1962). *A theory of cognitive dissonance, Volume 2*. Stanford university press.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The journal of abnormal and social psychology*, 58(2), 203-210.

Nisbett, R. E. and T. D. Wilson (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review* 84(3), 231

Sharot, T., Velasquez, C. M., & Dolan, R. J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological science*, 21(9), 1231-1235.

Stafford, T. (2014). [The perspectival shift: how experiments on unconscious processing don't justify the claims made for them](#). *Frontiers in Psychology*, 5, 1067. doi:10.3389/fpsyg.2014.01067

Stafford, T. (2015). [For argument's sake: evidence that reason can change minds](#).

Vinckier, F., L. Rigoux, I. T. Kurniawan, C. Hu, S. Bourgeois-Gironde, J. Daunizeau, and M. Pessiglione (2019). Sour grapes and sweet victories: How actions shape preferences. *PLoS*

Computational Biology 15(1), e1006499.

Wilson, T. D. C., & Nisbett, R. E. (1978). The Accuracy of Verbal Reports About the Effects of Stimuli on. *Social Psychology*, 41(2), 118-131.