

Running Head: Behavioral rationalizations

Quantifying the prevalence and adaptiveness of behavioral rationalizations

Warren Tierney

University of Limerick

Eric Luis Uhlmann

INSEAD

(Commentary on F. Cushman, “Rationalization is Rational”)

ABSTRACT: 60 words

MAIN TEXT: 998 words

REFERENCES: 808 words

ENTIRE TEXT: 1968 words

Authors' Note: This research was supported by an R&D grant from INSEAD. Please address correspondence to Warren Tierney, [warrentierney@hotmail.com](mailto:warrentierney@hotmail.com), or Eric Uhlmann, [eric.luis.uhlmann@gmail.com](mailto:eric.luis.uhlmann@gmail.com)

CONTACT:

Warren Tierney

Kemmy Business School

University of Limerick

Castletroy, Limerick

Ireland

Phone: +35387329150

E-mail: warrentierney@hotmail.com

Eric Luis Uhlmann

Organisational Behaviour Area

INSEAD

1 Ayer Rajah Avenue 138676

Singapore

Phone: 65 8468 5671

E-mail: eric.luis.uhlmann@gmail.com

*Abstract:* Critical aspects of the “rationality of rationalizations” thesis are open empirical questions. These include the frequency with which past behavior determines attitudes (as opposed to attitudes causing future behaviors), the extent to which post-hoc justifications take on a life of their own and shape future actions, and whether rationalizers experience benefits in well-being, social influence, performance, or other desirable outcomes.

Cushman (in press) posits that rationalizing past behaviors “extracts information from non-rational psychological processes (e.g., instinct, habit and norm compliance) and makes it available for subsequent reasoning,” and in doing so “improves subsequent reasoning.”

We suggest that key aspects of the rationality of rationalizations thesis are open empirical questions, among these the prevalence of behavioral rationalizations, the extent to which rationalizations are carried over to future judgments, and whether rationalizations lead to desirable outcomes for the person engaging in them. Such empirical questions can be addressed through studies capturing dynamic interactions between self-reported attitudes and behaviors over time, as well as the correlates and downstream consequences of behavioral rationalizations.

*How prevalent a phenomenon are behavioral rationalizations, in other words cases in which past behaviors determine future explicit attitudes?* The available longitudinal evidence suggests that Time 1 explicit attitudes predict Time 2 behaviors far better than past behaviors predict future self-reported attitudes, calling into question the prevalence of post-hoc rationalizations for past actions (Bentler & Speckart, 1981; Fredricks & Dossett, 1983; Kahle & Berman, 1979). Popular perspectives on attitude-behavioral relations may be “surprise-hacked” (Felin, Felin, Krueger, & Koenderink, 2019), overemphasizing instances in which behaviors cause explicit preferences (Bem, 1972; Festinger, 1962), and automatic and

unintentional processes determine human behavior outside of conscious awareness (Caruso, Shapira, & Landy, 2017; Forscher et al., in press; Lodder, Ong, Grasman, & Wicherts, in press; McCarthy et al., 2018; Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2015).

Although further longitudinal and meta-analytic investigations are needed, the “boring” narrative that conscious preferences and intentions typically direct future actions may capture a far greater share of the variance (Armitage & Conner, 2001; Ajzen, 1985; Fishbein & Ajzen, 1975; Randall & Wolff, 1994; Sheppard, Hartwick, & Warshaw, 1988; Webb & Sheeran, 2006), relegating the “rationalizations are rational thesis” to address only a small portion of the attitude-behavior relationship.

*Once formed, are rationalizations carried over to future judgments?* In other words, do explicit preferences formed to justify past acts play a causal role in directing future actions, or are such conscious rationalizations brief coping mechanisms, or a mere residue of behaviors determined by implicit processes (Gazzaniga, 1985)? One relevant experiment on moral judgments manipulated victim race, finding that whether the individuals sacrificed are White Americans or Black Americans impacts if consequentialist vs. deontological values are endorsed as general principles (Uhlmann, Pizarro, Tannenbaum, & Ditto, 2009). Further, once formed, such motivated moral principles impact downstream judgments. For example, if deontological morality is endorsed in a motivated fashion because the victims are Black Americans in the first moral dilemma, the same principle is then applied to a second moral dilemma in which victims are White Americans. Although further studies testing for such carryover effects are needed, this provides initial evidence that rationalizations can play a causal role in future judgments, a key aspect of Cushman’s (in press) thesis.

At the same time, the Uhlmann et al. (2009) results and related findings on intergroup attitudes (e.g., Brescoll, Uhlmann, & Newman, 2013; Hodson, Dovidio, & Gaertner, 2002; Norton, Vandello, & Darley, 2004; Tannenbaum, Valasek, Knowles, & Ditto, 2013) seriously question whether rationalizations improve subsequent reasoning. For example, individuals who exhibit negative automatic associations with the overweight on indirect measures are also more likely to explicitly favor increased insurance premiums for overweight employees. Yet they justify such punitive policy preferences in terms of cost effectiveness, rather than personal beliefs about body weight (Tannenbaum et al., 2013). Given that target ethnicity and obesity are not defensible inputs into moral judgments in the first place, how does rationalizing group-based biases and then carrying forward such justifications improve subsequent reasoning in any way? Even assuming for a moment that implicit preferences are somehow “truer” or more authentic than explicit preferences (a highly debatable characterization), the rationalization process has obscured, rather than revealed, this deeper attitude. Applying the criterion of subjective rationality (Pizarro & Uhlmann, 2005) it seems doubtful that decision makers themselves would, if made aware of it, welcome the influence of implicit overweight bias on their recommended company insurance policies. More likely, we think, they would seek to correct for and remove such unwanted prejudices (Fazio, 1990), and perceive them as in conflict with their ideal self (Monteith, Devine, & Zuwerink, 1993). This leads us to the broader issue of whether post-hoc justifications are “good” for the rationalizer in some measurable way.

*Do rationalizations lead to positive objective or subjective outcomes for the agent?* If the “ultimate purpose of reasoning” is “fitness maximization”, and rationalizations improve reasoning (Cushman, in press), then individuals who engage in rationalizations should score higher on measures of adjustment, effectiveness, and performance. For instance, rationalizers

may display higher levels of psychological well-being, enjoy better social reputations, have an easier time influencing their peers, and exhibit superior job performance. Conversely, rationalizers could tend to be unhappy, socially unpopular underperformers, rejected and ineffective due to their self-serving arguments and lack of insight into their own actions. This is analogous to the debate between Taylor and Brown (1988) and Colvin, Block, and Funder (1995) on the adaptiveness of positive illusions about the self, and is an empirical question to be addressed in future studies. Some relevant evidence is provided by Uhlmann and Cohen (2005), who find that individuals who rationalize their hiring decisions engage in greater gender discrimination, and yet perceive themselves as more objective and unbiased. This suggests rationalizations may be associated with favorable subjective self-assessments (see also Dunning, Leuenberger, & Sherman, 1995), but with suboptimal objective outcomes (see Hunt, Layton, & Prince, 2015; Inglehart & Norris, 2003; Woolley, Chabris, Pentland, Hashmi, & Malone, 2010; for evidence that gender inclusiveness improves group performance). That rationalizers are more likely to make sexist decisions and suffer from an illusion of objectivity would seemingly count as initial evidence against the putative rationality of rationalizations.

Ultimately, the rationalizations are rational thesis (Cushman, in press) is important, insightful, and likely to prove generative of further empirical research on attitude-behavior relations, reasoning processes, and human adaptability and performance.

### References

- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckman (Eds.), *Action control: From cognition to behavior* (pp. 11-39). Berlin: Springer-Verlag.
- Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behavior: A meta-analytic review. *British Journal of Social Psychology, 40*, 471-499.
- Bem, D. J. (1972). Self-perception theory. In *Advances in Experimental Social Psychology*, Volume 6, pp. 1–62. Elsevier.
- Bentler, P.M., & Speckart, G. (1981). Attitudes “cause” behaviors: A structural equation analysis. *Journal of Personality and Social Psychology, 40*, 226-238.
- Brescoll, V.L., Uhlmann, E.L., & Newman, G.N. (2013). The effects of system-justifying motives on endorsement of essentialist explanations for gender differences. *Journal of Personality and Social Psychology, 105*, 891-908.
- Caruso, E. M., Shapira, O., & Landy, J. F. (2017). Show me the money: A systematic exploration of manipulations, moderators, and mechanisms of priming effects. *Psychological Science, 28*, 1148-1159.
- Colvin, C.R., Block, J., & Funder, D.C. (1995). Overly-positive self evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology, 68*, 1152-1162.
- Dunning, D., Leuenberger, A., & Sherman, D. A. (1995). A new look at motivated inference: Are self-serving theories of success a product of motivational forces? *Journal of Personality and Social Psychology, 59*, 58-68.
- Fazio, R.H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. (In M.P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol 23, pp. 75-109). San Diego, CA: Academic Press.)

- Felin, T., Felin, M., Krueger, J. I., & Koenderink, J. (2019). On surprise-hacking. *Perception*, 48, 109-114.
- Festinger, L. (1962). *A theory of cognitive dissonance, Volume 2*. Stanford university press.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B.A. (in press). A meta-analysis of procedures to change implicit measures. *Journal of Personality & Social Psychology*.
- Fredricks, A.J., & Dossett, D.L. (1983). Attitude-behavior relations: A comparison of the Fishbein-Ajzen and the Bentler-Speckart models. *Journal of Personality and Social Psychology*, 45, 501-512.
- Gazzaniga, M. (1985). *The social brain*. New York: Free Press
- Hodson, G., Dovidio, J.F., & Gaertner, S.L. (2002). Processes in racial discrimination: Differential weighting of conflicting information. *Personality and Social Psychology Bulletin*, 28, 460-471.
- Hunt, V., Layton, D., & Prince, S. (2015). *Diversity matters*. McKinsey.
- Inglehart, R., & Norris, P. (2003). *Rising tide: Gender equality and cultural change around the world*. Cambridge University Press.
- Kahle, L.R., & Berman, J.J. (1979). Attitudes cause behaviors: A cross-lagged panel analysis. *Journal of Personality and Social Psychology*, 37, 315-321.
- Lodder, P., Ong, H. H., Grasman, R. P. P. P., & Wicherts, J. (in press). A comprehensive metaanalysis of money priming. *Journal of Experimental Psychology: General*.
- McCarthy, R. J., Skowronski, J. J., Verschuere, B., Meijer, E. H., Jim, A., Hoogesteyn, K., Orthey, R.,..... Yildiz, E. (2018). Registered Replication Report: Srull & Wyer (1979). *Advances in Methods and Practices in Psychological Science*, 1, 321-336.

- Monteith, M. J., Devine, P. G., & Zuwerink, J. R. (1993). Self-directed versus other-directed affect as a consequence of prejudice-related discrepancies. *Journal of Personality and Social Psychology, 64*, 198-210.
- Norton, M.I., Vandello, J.A., & Darley, J.M. (2004). Casuistry and social category bias. *Journal of Personality and Social Psychology, 87*, 817–831.
- Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2015). Using the IAT to predict ethnic and racial discrimination: Small effect sizes of unknown societal significance. *Journal of Personality and Social Psychology, 108*(4), 562-571.
- Pizarro, D.A., & Uhlmann, E.L. (2005). Do normative standards advance our understanding of moral judgment? *Behavioral and Brain Sciences, 28*, 558-559.
- Randall, D. M., & Wolff, J. A. (1994). The time interval in the intention-behaviour relationship: Meta-analysis. *British Journal of Social Psychology, 33*, 405-418.
- Sheppard, B. H., Hartwick, J., & Warshaw, P. R. (1988). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research, 15*, 325-343.
- Tannenbaum, D., Valasek, C.J., Knowles, E.D., & Ditto, P.H. (2013). Incentivizing wellness in the workplace: Sticks (not carrots) send stigmatizing signals. *Psychological Science, 24*, 1512-1522.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.
- Uhlmann, E.L., Pizarro, D.A., Tannenbaum, D., & Ditto, P.H. (2009). The motivated use of moral principles. *Judgment and Decision Making, 4*, 476–491.
- Uhlmann, E.L., & Cohen, G.L. (2005). Constructed criteria: Redefining merit to justify discrimination. *Psychological Science, 16*, 474-480.
- Webb, T. L., & Sheeran, P. (2006). Does changing behavioural intentions engender behavior

change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, 132, 249-268.

Woolley, A.W., Chabris, C.F., Pentland, A., Hashmi, N., & Malone, T.W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330, 686–688.