Supplementary Materials

for

On the belief that beliefs should change according to evidence: Implications for conspiratorial, moral, paranormal, political, religious, and science beliefs

Gordon Pennycook[1*], James Allan Cheyne[2], Derek J. Koehler[2], & Jonathan A. Fugelsang[2]

[1] Hill/Levene Schools of Business, University of Regina, Canada
[2] Department of Psychology, University of Waterloo, Canada

*corresponding author: gordon.pennycook@uregina.ca

Study S1 – Scale Validation

Our goal in Study S1 was to create a scale to assess one's belief that beliefs ought to change according to evidence (AOT-E). Fortunately, some items from common AOT measures do seem to be particularly relevant for belief formation. Namely, we used Stanovich and West's (2007) 41-item Actively Open-minded Thinking (AOT) scale as our starting point. The AOT consists of 6 subscales: 1) the Actively Open-minded Thinking subscale (10 items; Note: The full AOT scale has an AOT subscale), 2) Belief Identification (9 items), 3) Categorical Thinking (3 items), 4) Counterfactual thinking (2 items), 5) Dogmatism (9 items), and 6) Openness-Values (9 items). We isolated 8 items that have a clear connection to the concept of belief mutability (see Table 1). The items were drawn from 4 of the 6 full AOT subscales. To establish the resulting Actively Open-minded Thinking about Evidence (AOT-E) scale, we re-analyzed data from Pennycook, Cheyne, Koehler, and Fugelsang (2016) in which the full AOT scale was included alongside Pacini and Epstein's (1999) conceptually similar Rational-Experiential Inventory and the Cognitive Reflection Test (CRT, Frederick, 2005). We also obtained a measure of religious belief from a mass testing survey that was administered prior to the study session reported in Pennycook, Cheyne, et al. Neither AOT nor religious belief were reported by Pennycook, Cheyne, et al., as they were not relevant to the hypothesis (which had to do with CRT scoring techniques).

Method

As mentioned, components of this data set were reported by Pennycook et al. (2016). However, Pennycook et al. focused on the CRT and, as a consequence, restricted their sample only to participants who reported no prior exposure to the CRT. Here we will use the full data set. Components of this data set were also reported in Pennycook, Ross, Koehler, and Fugelsang, (2016), although that investigation focused on the association between CRT performance and religious disbelief. Here we will focus on the psychometric properties of the AOT-E scale and its performance relative to the full AOT scale.

Participants

Undergraduate students at the University of Waterloo completed an online study that included the CRT and various other reasoning measures. Only participants who had completed an earlier mass testing survey that included a religious belief measure (along with various other measures) were permitted to sign up (see Pennycook, Ross, et al., 2016). Participants who had completed both of these surveys were then permitted to sign up for a final online study that consisted of the self-report thinking disposition questionnaires. This study was run over three semesters (Winter and Fall 2013, Winter 2014). The original data set consisted of 498 participants; however, 63 participants had at least one missing response to an item in the AOT scale. The resulting sample (N = 436, *Mean* age = 20.6) consisted of 138 males and 297 females (1 individual did not indicate gender).

Materials

The Actively Open-minded Thinking scale (AOT) consists of 41 items with questions like "If I think longer about a problem I will be more likely to solve it" (AOT subscale), "Someone who attacks my beliefs is not insulting me personally" (Belief Identification subscale), "There are basically two kinds of people in this world, good and bad" (rev, Categorical Thinking subscale), "My beliefs would not have been very different if I had been raised by a different set of parents" (rev, Counterfactual Thinking subscale), "Often, when people criticize me, they don't have their facts straight" (rev, Dogmatism subscale), and "I consider myself broad-minded and tolerant of other people's lifestyles" (Openness-Values subscale).

The Rational-Experiential Inventory consists of two separate scales. The Need for Cognition (NFC) scale is intended to assess one's enjoyment of mental activities and includes items like "I enjoy intellectual challenges" and "Thinking hard and for a long time about something gives me little satisfaction" (rev). The Faith in Intuition (FI) scale is intended to assess the degree to which one relies on their intuitions and gut feelings. It includes items like "Intuition can be a very useful way to solve problems" and "I generally don't depend on my feelings to help me make decisions" (rev).

The Cognitive Reflection Test (CRT) consists of three word problems that reliably cue an incorrect intuitive response. For example: "A bat and ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?" The response that comes to mind for most people is 10 cents; however, this is not the correct answer (which is 5 cents). According to dual-process theorists, the CRT is a measure of analytic thinking disposition because it requires a willingness or propensity to think analytically to stop and question the initial intuitive response (Pennycook & Ross, 2016; Travers, Rolison, & Feeney, 2016).

Participants also completed a religious belief scale that consisted of 8 questions about eight conventional supernatural religious beliefs (Pennycook, Ross, et al., 2016): heaven, hell, miracles, afterlife, angels, demons, soul, and the devil/Satan. Participants indicated their degree of belief in religious concepts on a 5-point scale: 1) I strongly disagree, 2) I disagree, 3) I don't know, 4) I agree, 5) I strongly agree.

All scales were converted to POMP scores (ranging from 0-100) prior to analysis.

Results

Descriptive statistics can be found in Table S1. Participants rated their beliefs as more mutable than immutable and, indeed, the mean score on the AOT-E scale was well above scale midpoint, $t(435) = 22.79$, $SE = .69$, $p < .001$. Given how difficult is it to change people's beliefs in an experimental setting, this indicates a likely overestimation of belief mutability. This is paralleled by mean scores for the other thinking disposition scales, which are also all above scale midpoint (all t's > 7.77, p's < .001). All scales had acceptable reliability (Table S2).

Table S1. Mean score and associated standard deviations and skewness for variables in Study S1.

Scale	*Mean*	*SD*	Skew[1]	Kurtosis[2]

Actively Open-minded Thinking (AOT-1)	66.1	10.2	-0.28	0.03
AOT about Evidence (AOT-E)	65.7	14.4	-0.04	-0.07
AOT with AOT-E items removed (AOT-2)	66.1	10.0	-0.32	0.08
Need for Cognition (NFC)	65.0	13.7	-0.01	-0.27
Faith in Intuition (FI)	55.0	13.5	-0.15	0.24
Cognitive Reflection Test (CRT)	40.4	36.8	0.33	-1.26
Religious Belief	48.9	26.4	-0.20	-0.46

1. S. E. = .12
2. S. E. = .23

Table S2. Correlations (Pearson *r*) among variables in Study S1. Cronbach's Alpha for each scale is listed in brackets along major diagonal. $N = 436$.

	1	2	3	4	5	6	7
1. AOT-E	(.77)						
2. AOT-1	.83***	(.89)					
3. AOT-2	.70***	.98***	(.85)				
4. NFC	.37***	.35***	.31***	(.87)			
5. FI	-.22***	-.11*	-.06	-.07	(.89)		
6. CRT	.35***	.34***	.31***	.35***	-.19***	(.63)	
7. Religious Belief	-.54***	-.45***	-.39***	-.22***	.18***	-.27***	(.94)

***indicates $p < .001$, **indicates $p < .01$, *indicates $p < .05$.

As is evident from Table S2, there were significant correlations among all of the variables. The only exception is that the FI scale did not correlate with the AOT-2 (i.e., AOT with AOT-E items removed) and NFC scales. Somewhat surprisingly, the 8-item AOT-E scale was a nominally (but not significantly) stronger predictor of NFC, FI, and CRT scores than the 33-item AOT scale (i.e., with the AOT-E items excluded). More importantly, the AOT-E scale was more strongly predictive of religious belief ($r = -.54$) than was the AOT-2 scale ($r = -.39$). This difference was significant according to a William's test, $t(429) = 4.97$, $p < .001$. Consistent with previous research (e.g., Svedholm & Lindeman, 2013), the NFC and FI scales only moderately predicted religious belief (r's = -.22 and .18 for NFC and FI scales, respectively).

Study S2 – Discriminant Validity

Here we demonstrate discriminant validity by showing predictive validity of the AOT-E scale over openness to experience.

Method

Participants

The original sample included 2222 undergraduate students at the University of Waterloo. However, 15 participants had no data for the AOT-E and 33 participants had no data for the religious belief scale. A further 11 participants had missing data in one or more items for either AOT-E or religious belief scales. The resulting sample (N = 2163) consisted of 635 males, 1520 females, and 8 individuals who did not indicate their gender.

Materials

Participants were administered the same AOT-E and religious belief scale as in Study 1. However, unlike in Study S1, the 8-item AOT-E scale was not imbedded in the larger AOT scale (i.e., participants only responded to the 8 AOT-E items). The mass testing survey included a large number of scales presented in a random order, but the AOT-E scale was anchored to the second position (following only a boredom proneness scale).

Participants also completed an openness to experience scale, which is a component of the Big-Five Inventory (Gosling, Rentfrow, & Swann, 2003). For this, participants rated the degree to which their personality matched a series of openness-related questions. Example items include "Is original, comes up with new ideas" and "Prefers work that is routine" (rev). The openness to experience scale was only administered to a subset of the participants (N = 1053).

Results

As in Study S1, the AOT-E scale had acceptable reliability (Cronbach's Alpha = .78). The religious belief scale (Cronbach's Alpha = .95) and the openness to experience scale (Cronbach's Alpha = .77) were also reliable. There was a strong negative correlation between AOT-E and religious belief, $r(2163) = -.48$, $p < .001$. As expected, AOT-E was significantly but not strongly correlated with openness to experience, $r(1053) = .16$, $p < .001$. Crucially, openness to experience, unlike AOT-E, was not significantly correlated with religious belief, $r(1053) = -.04$, $p = .230$. Thus, AOT-E does appear to be measuring something distinct from a general openness factor. Openness does not appear to play a role in religious belief, whereas belief mutability is very strongly predictive.

Additional Analyses of Studies 1-3

Controlling for political ideology and demographics

In the main text (Study 3), we correlate AOT-E with our various DVs separately for Democrats, Republicans, and Independents. Another way to investigate the predictive power of the AOT-E is to simply take political ideology (measured continuously using our two liberal-conservative ideology questions), along with various demographics, into account via multiple regression. For Study 1, social and fiscal conservatism were entered along side age and gender. For Studies 2 and 3, demographics included age, gender (male, female), education, and income. For Study 2, we report correlations across both versions of AOT-E for simplicity. As is evident from Table S3, AOT-E is significantly correlated with every DV in all three studies even after political ideology and demographics were statistically controlled. The same was true for CRT performance (albeit to a lesser extent), except for care/fairness moral values and free market ideology.

Table S3. Correlations (β) between AOT-E/CRT and primary measures in Studies 1-3 once controlling for political ideology and demographic factors (age, gender, education, income). AOT-E = Actively Open-minded Thinking about Evidence. CRT = Cognitive Reflection Test. MV = Moral Values.

		Conspiracy Beliefs	Paranormal Beliefs	Traditional MV	Care/Fairness MV	Free Market Ideology	Conserv Opinions	Pro-Science Beliefs	Religious Beliefs	God Skepticism
AOT-E	Study 1	-.28***	-.33***	-.40***	.20***	-.24***	-.60***	.55***	-	.40***
	Study 2	-.21***	-.18**	.21***	.23***	-.16***	-.42***	.31***	-.23***	.20***
	Study 3	-.14***	-.20***	-.13***	.36***	-.16***	-.28***	.29***	-.16***	.11**
CRT	Study 1	-	-	-	-	-	-	-	-	-
	Study 2	-.10*	-.14**	-.16***	-.01	-.06	-.12*	.20***	-.16**	-.15**
	Study 3	-.09*	-.16***	-.16***	.01	-.04	-.18***	.13**	-.15***	.13***

***indicates $p < .001$, **indicates $p < .01$, *indicates $p < .05$

References

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, *76*(6), 972–987. https://doi.org/10.1037/0022-3514.76.6.972

Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2016). Is the cognitive reflection test a measure of both reflection and intuition? *Behavior Research Methods*, *48*, 341–348. https://doi.org/10.3758/s13428-015-0576-1

Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2016). Atheists and Agnostics Are More Reflective than Religious Believers: Four Empirical Studies and a Meta-Analysis. *Plos One*, *11*(4), e0153039. https://doi.org/10.1371/journal.pone.0153039

Pennycook, G., & Ross, R. M. (2016). Commentary on: Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology*, *7*, 9. https://doi.org/10.3389/fpsyg.2015.00532

Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning*, *13*(3), 225–247. https://doi.org/10.1080/13546780600780796

Svedholm, A. M., & Lindeman, M. (2013). The separate roles of the reflective mind and involuntary inhibitory control in gatekeeping paranormal beliefs and the underlying intuitive confusions. *British Journal of Psychology*, *104*(3), 303–319. https://doi.org/10.1111/j.2044-8295.2012.02118.x

Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the Cognitive Reflection Test. *Cognition*, *150*, 109–118. https://doi.org/http://dx.doi.org/10.1016/j.cognition.2016.01.015