Following one's true self and the sacredness of cultural values

Jinhyung Kim, Andrew G. Christy, Grace N. Rivera, Rebecca J. Schlegel, Joshua A. Hicks*

Texas A&M University, Department of Psychology, 4235 TAMU Psychology Building, College Station, TX 77843, United States

A B S T R A C T

People seem to share a widespread lay belief that *true selves* are morally good entities. This lay belief has downstream consequences for a variety of domains such as person perception and perceived self-knowledge. The current work examines whether it also has consequences for moral decision-making. We hypothesized that people would make more moral decisions when they were focused on being authentic as opposed to being focused on other decision-making strategies. This hypothesis rests on the idea that if people believe their true selves are morally good, then attempts to follow that true self will make them less willing to behave immorally. Consistent with this hypothesis, four within-subjects studies (total $N = 817$) found that participants reported that they and others would need more money to violate a moral norm if they were focused on trying to be authentic relative to if they were focused on being rational, intuitive, or realistic.

People frequently assume that they (and others) possess a *true* self that represents who they really are inside, regardless of how they outwardly behave (Schlegel & Hicks, 2011). Emerging research suggests that people believe *true* selves not only exist, but are morally good (De Freitas, Cikara, Grossmann, & Schlegel, 2017; Strohminger, Knobe, & Newman, 2017). Despite striking differences across individuals and cultures in other beliefs about the nature and structure of the self (e.g., Markus & Kitayama, 1991), the idea that true selves are good seems pervasive. For example, even people who hold extreme negative views of others (i.e., self-identified misanthropes) agree that true selves are morally good (De Freitas et al., 2017), further suggesting people share a common lay belief that one's authentic self is inherently good.

This lay belief influences how we perceive ourselves (Christy, Seto, Schlegel, Vess, & Hicks, 2016) and others (Christy, Kim, Schlegel, Vess, & Hicks, 2017; De Freitas & Cikara, 2018; Strohminger & Nichols, 2014, 2015; Wojciszke, 2005). For example, when people think about close others who have changed, the valence of those changes predicts whether they are seen as movements towards or away from the true self (Bench, Schlegel, Davis, & Vess, 2015). The current research examines whether this belief also influences moral judgments and decisions, by assessing how people respond to morally-charged scenarios when instructed to *be authentic* compared to alternative instructions. If true selves are conceived of as morally good, then instructions to engage with the true self (i.e., to be authentic), should result in greater adherence to moral virtues than alternative instructions.

1. Study 1

1.1. Method

Participants were 192 undergraduate students from a large public university (91 female, 100 male, 1 transgender, $M_{age} = 18.83$, $SD_{age} = 1.08$; 62% White). We did not conduct power analyses. However, for Studies 1, 2, and 4 we set a target minimum of 150 participants and collected data in one-week increments until this minimum was reached. Our aim was to maximize power within the constraints of lab resources.

1.2. Materials and procedure

1.2.1. Decision strategy manipulation (within-subjects)

Participants were told they would make a series of decisions involving their willingness to engage in certain behaviors and that they would make each decision twice using different strategies (i.e., "How much money would it take for you to do this if you were focused on trying to be authentic/thinking rationally?"). The order of these items was randomized for each trial.

1.2.2. Moral foundations

We used the Moral Foundations Sacredness Scale (MFSS; Graham, Haidt, & Nosek, 2009) to assess participant's willingness to violate moral norms within five categories (i.e., Harm, Fairness, In-group Loyalty, Authority, and Purity). Specifically, participants indicated how

Table 1
Descriptive and test statistics (Study 1).

Domain	Authenticity		Rationality		95% CI				Cohen's d
	M	SD	M	SD	t	p	LL	UL	
Total MFSS	6.17	0.99	5.91	1.06	6.05	< .001	0.18	0.35	0.25
Harm	6.94	1.26	6.60	1.42	5.78	< .001	0.22	0.46	0.25
Fair	5.96	1.35	5.70	1.30	3.67	< .001	0.12	0.40	0.20
Ingroup	6.55	1.15	6.23	1.14	5.77	< .001	0.21	0.42	0.28
Authority	4.86	1.56	4.81	1.60	0.66	.508	−0.10	0.20	0.03
Purity	6.57	1.15	6.24	1.29	6.30	< .001	0.23	0.44	0.27

much money it would take for them to commit acts that violate each of the five moral foundations (e.g., "Kick a dog in the head, hard" for harm) on a scale from 1 (*$0 - I'd do it for free*) to 8 (*never for any amount of money*). In addition to creating subscales, responses were averaged to form an overall composite.

All primary measures, manipulations, and exclusions are reported in the manuscript. Exploratory measures, administered after the main tasks, as well as information about four additional studies, are available at https://osf.io/fbtrm/?view_only=c9e5b77457d841dfb2d794fa7e34464a.

1.3. Results and discussion

A paired-samples *t*-test showed that the differences between conditions were significant for the overall composite and for all subscales save authority (see Table 1). In each case, participants required more money to violate moral norms when they were focused on being authentic.

In Study 2, we sought to replicate and extend these findings by including an intuition condition. Intuition is a common counterpart to rationality, and this allowed for a more inclusive test of how authenticity and other common decision-making strategies relate to moral judgments.

2. Study 2

2.1. Method

Participants were 167 undergraduate students from a large public university (119 female, 47 male, 1 other; $M_{age} = 19.51$, $SD_{age} = 1.23$; 54% White). Eleven participants were excluded from analyses for failing two attention checks.

2.2. Materials and procedure

This study used the same basic procedure, except that participants were told to use *three* different strategies when completing the 20 MFSS items (i.e., "How much money would it take for you to do this if you were focused on being authentic/thinking rationally/trusting your gut?"). Responses were averaged to form the subscale and total composite scores.

2.3. Results and discussion

Repeated-measures ANOVAs revealed a significant effect of strategy on the composite as well as all subscales except authority (see Table 2). A planned Helmert contrast on the composite revealed a significant difference between authenticity and the other strategies, $F(1, 155) = 13.57$, $p < .001$, partial $\eta^2 = 0.08$. Nonetheless, follow-up analyses revealed this difference was mostly driven by the difference between the true self and rational thinking conditions as the difference

than differences between authenticity and rationality. This may be because participants interpreted the authenticity and intuition instructions similarly; instructions to be authentic might be psychologically equivalent to instructions to be intuitive.

To address this possibility, we explored whether people *think* being authentic is intuitive ($n = 77$). Participants completed the MFSS, but only answered the items once under instructions to be authentic. Afterwards, they indicated the extent to which they had used rational and intuitive processing. Results indicated no significant difference in how rational versus intuitive participants thought they had been, $t(76) = 0.11$, $p = .913$, $d = 0.02$, 95% CI = [−0.30, 0.33] (for a more complete description, see supplementary materials). These findings speak against the idea that people equate authentic and intuitive processing. When instructed to be authentic, people reported using both rational *and* intuitive processes.

Another concern not addressed in the previous studies is that the rational thinking condition drives the previously observed differences by *decreasing* the amount of money needed to violate moral norms. In order to address this concern, Study 3 compared instructions to be authentic versus *realistic*. This provided another strong test of our hypothesis considering that people generally perceive themselves as moral (Aquino & Reed II, 2002), even without explicit reference to the true self.

3. Study 3

3.1. Methods

Participants were 220 American adults recruited from Amazon's Mechanical Turk (89 female, 130 male, 1 not reporting; $M_{age} = 33.43$, $SD_{age} = 9.67$; 75% White). We pre-registered this study on AsPredicted.org (http://aspredicted.org/blind.php?x=sk8bf2). Seven participants were excluded from analyses for failing two attention checks (see pre-registration).

3.2. Materials and procedure

Participants completed the MFSS. For each item, they responded under instructions to be authentic *and* under instructions to realistically estimate how much money they would accept if they were actually confronted with each situation. Responses were averaged to form subscales and a total composite.

3.3. Results and discussion

A paired samples *t*-test indicated a significant difference between conditions on the overall composite and across all the subscales but authority (see Table 3).

While we believe that the patterns observed across Studies 1–3 are due to a tendency to ascribe morally good content to true selves, a plausible alternative explanation is that these effects are driven by self-

between the true self and intuition strategies did not reach significance.

These results are generally consistent with hypotheses. However, the differences between authenticity and intuition were notably smaller

serving motivations. People may report enhanced moral intentions under instructions to be authentic not because they actually conceive of their authentic selves as morally virtuous, but because they want to

Table 2
Descriptive and test statistics (Study 2).

Domain	Authenticity		Rationality		Intuition				Partial η^2
	M	SD	M	SD	M	SD	F	p	
Total MFSS	6.54_a	0.97	6.33_b	0.96	6.46_a	0.94	10.68	< .001	0.06
Harm	7.34_a	0.93	7.10_b	1.06	7.29_a	0.93	11.42	< .001	0.07
Fair	6.48_a	1.32	6.22_b	1.27	6.26_b	1.34	9.44	< .001	0.06
Ingroup	6.68_a	1.15	6.50_b	1.06	6.69_a	1.05	6.20	.002	0.04
Authority	5.38_a	1.70	5.33_a	1.47	5.29_a	1.62	0.71	.49	0.01
Purity	6.79_a	1.15	6.52_b	1.15	6.78_a	1.08	14.90	< .001	0.09

Note. Within each row, means not sharing a subscript differ at $p < .05$.

Table 3
Descriptive and test statistics (Study 3).

Domain	Authenticity		Actual behavior		95% CI				Cohen's d
	M	SD	M	SD	t	p	LL	UL	
Total MFSS	5.93	1.14	5.76	1.18	5.92	< .001	0.12	0.23	0.15
Harm	6.79	1.20	6.57	1.28	5.37	< .001	0.14	0.30	0.17
Fair	5.91	1.40	5.66	1.41	5.19	< .001	0.15	0.34	0.17
Ingroup	5.82	1.40	5.64	1.43	4.75	< .001	0.11	0.26	0.13
Authority	4.98	1.67	4.92	1.67	1.55	.124	−0.02	0.15	0.04
Purity	6.17	1.42	6.00	1.41	4.66	< .001	0.10	0.24	0.12

think of and publicly present their authentic selves in positive terms. In Study 4, we sought to reduce self-enhancement motives by having participants reflect on what *other people* would do.

4. Study 4

4.1. Methods

Participants were 244 undergraduate students (166 female, 78 male; $M_{age} = 18.88$, $SD_{age} = 1.66$; 64% White) recruited from a large public university.

4.2. Materials and procedure

In a within-subjects design, participants reported how *a typical person* would respond to each MFSS item under three different conditions: under ordinary circumstances, being authentic, and being rational. As before, the 20 items were presented in random order, accompanied by sub-items asking how a typical person would respond under each of the three conditions, and item scores were averaged to yield the subscale scores and global composite.

4.3. Results and discussion

A repeated-measure ANOVA indicated a significant effect of condition on the MFSS composite and on all five subscales (see Table 4). A planned Helmert contrast on the composite indicated that the authentic condition differed significantly from the other two conditions, $F(1, 243) = 114.31$, $p < .001$, partial $\eta^2 = 0.32$. Post-hoc tests indicated a generally consistent pattern across the dependent variables, such that scores in the authenticity condition were typically higher than scores in both the ordinary and rational conditions (see Table 4). Once again, the authority subscale was an exception to this general pattern, with scores in the authentic and rational conditions being higher than the ordinary

or the previously observed results. Indeed, the differences between conditions were even more pronounced in this study than the previous studies. This suggests that these effects emerge not because of, but in spite of, self-serving motives that may have pushed responses towards the ceiling in Studies 1–3.

5. Internal meta-analysis

Finally, we meta-analyzed all 8 studies (4 reported here in the text and 4 available on OSF). Details of this meta-analysis are on OSF, but both a fixed ($p < .001$) and random effects ($p = .018$) model yielded significant overall effects.

6. General discussion

Across the present studies, people consistently reported that they themselves (Studies 1–3) and other people (Study 4) would be less willing to engage in immoral activities when responding *authentically*, compared to other modes of responding. These findings are consistent with the idea that people conceive of their own and others' *true selves* as morally good entities. The fact that the patterns were highly consistent across the various subscales of the MFSS suggests that the moral goodness ascribed to the true self is quite broad, encompassing multiple dimensions of moral normativity.

A notable exception to this pattern was the authority subscale, which tended not to differ between conditions. It may be that participants do not see the authority items as representing serious moral violations (indeed, smaller payments were required to defy authorities compared to other moral norms). This possibility suggests that differences may emerge among individuals who hold authority in greater moral regard. [1] Another possibility is that the folk concepts of authenticity and respect for authority are to some extent contradictory; behaving authentically may often be perceived as antithetical to obeying authority figures. If this is the case, then even if

condition, but not differing from each other.

These results suggest that people believe even the typical person would be less willing to engage in immoral behavior when being authentic relative to being rational or under ordinary circumstances. Since self-enhancement motives should be greatly reduced in this version of the paradigm, it seems unlikely that self-enhancement can explain this,

authenticity is broadly associated with morality, this association may not extend to the domain of authority. Further research is needed to more accurately examine potential moderators of this relationship.

---

[1] However, exploratory follow-up analyses provided limited evidence of this possibility (see supplementary material).

Table 4
Descriptive and test statistics for each subscale (Study 4).

Domain	Authentic		Rational		Ordinary				Partial η^2
	M	SD	M	SD	M	SD	F	p	
Total MFSS	6.13_a	0.86	5.70_b	0.98	5.45_c	0.86	79.28	< .001	0.25
Harm	6.91_a	1.18	6.41_b	1.31	6.25_b	1.21	43.57	< .001	0.15
Fair	5.84_a	1.38	5.44_b	1.28	5.02_c	1.18	55.04	< .001	0.19
Ingroup	6.52_a	1.17	6.09_b	1.10	5.93_c	1.08	59.81	< .001	0.20
Authority	4.75_a	1.61	4.53_a	1.43	3.90_b	1.29	55.60	< .001	0.19
Purity	6.63_a	1.12	6.00_b	1.19	6.13_b	1.13	57.97	< .001	0.19

Note. Within each row, means not sharing a subscript differ at $p < .05$.

One major limitation of the current work is our reliance on within-subjects methodology. Indeed, two between-subjects studies yielded null effects (see supplementary material). Although these null findings raise the possibility of demand characteristics influencing the results reported in the main text, we believe those concerns are more related to the size of the effect as opposed to the direction.[2] That is, even if the within-subjects design exaggerates the differences between conditions, it does so in a consistent way that still demonstrates most people possess a lay theory that true selves are morally good. Nevertheless, it is possible that these effects only emerge when people simultaneously consider other decision-making strategies. Dwyer, Dunn, and Hershfield (2017) recently showed that differences in two related constructs (meaning and happiness) are stronger only when people consider both constructs at the same time (e.g., meaning *without* happiness). It may be that the current decision-making strategies are so similar that the saliency of other possible options is needed to produce reliable differences.

While there are limitations, our findings suggest that many people *believe* that following their true self will engender moral behaviors. Future studies will ultimately reveal whether subtle reminders to act authentic actually enhance the importance of universal cultural norms and engender corresponding behavioral responses.

Open practices

See https://osf.io/fbtrm/?view_only=c9e5b77457d841dfb2d794fa7e34464a for materials and data for all experiments. The Preregistration report for Study 3 is available at http://aspredicted.org/blind.php?x=sk8bf2.

References

Aquino, K., & Reed, A., II (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*, 1423–1440.

Bench, S., Schlegel, R. J., Davis, W. E., & Vess, M. (2015). Thinking about change in the self and others: The role of self-discovery metaphors and the true self. *Social Cognition, 33*, 169–185.

Christy, A. G., Seto, E., Schlegel, R. J., Vess, M., & Hicks, J. A. (2016). Straying from the righteous path and from ourselves: Moral behavior and perceived self- knowledge. *Personality and Social Psychology Bulletin, 42*, 1538–1550.

Christy, A. G., Kim, J., Schlegel, R. J., Vess, M., & Hicks, J. A. (2017). The reciprocal relationship between perceptions of moral goodness and knowledge of others' true selves. *Social Psychological and Personality Science, 8*, 910–917.

De Freitas, J., & Cikara, M. (2018). Deep down my enemy is good: Thinking about the true self reduces intergroup bias. *Journal of Experimental Social Psychology, 74*, 307–316.

De Freitas, J., Cikara, M., Grossmann, I., & Schlegel, R. J. (2017). Origins of the belief in good true selves. *Trends in Cognitive Sciences, 21*, 634–636.

Dwyer, R., Dunn, E., & Hershfield, H. E. (2017). Cousins or conjoined twins: How different are meaning and happiness in everyday life? *Comprehensive Results in Social Psychology* (pp. 1–17). .

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*, 1029–1046.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224–253.

Schlegel, R. J., & Hicks, J. A. (2011). The true self and psychological health: Emerging evidence and future directions. *Social and Personality Psychology Compass, 5*, 989–1003.

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition, 131*, 159–171.

Strohminger, N., & Nichols, S. (2015). Neurodegeneration and identity. *Psychological Science, 26*, 1469–1479.

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science, 12*, 551–560.

Wojciszke, B. (2005). Morality and competence in person- and self-perception. *European Review of Social Psychology, 16*, 155–188.

---

[2] Notably, the effect size was considerably larger in the "typical person" (Study 4) version of the study compared to the "self" versions (Studies 1–3). In Study 4, we also tried to limit the role of demand by explicitly stating to participants that their answers may not necessarily differ across decision strategy.