Straying From the Righteous Path and From Ourselves: The Interplay Between Perceptions of Morality and Self-Knowledge

Andrew G. Christy[1], Elizabeth Seto[1], Rebecca J. Schlegel[1], Matthew Vess[2], and Joshua A. Hicks[1]

Abstract

The present research addresses the relationship between morally valenced behavior and perceptions of self-knowledge, an outcome that has received little attention in moral psychology. We propose that morally valenced behavior is related to subjective perceptions of self-knowledge, such that people experience lower levels of self-knowledge when they are reminded of their immoral behaviors. We tested this proposition in four studies (*N* = 1,177). Study 1 used daily-diary methods and indicates that daily perceptions of self-knowledge covary with daily levels of morally valenced behavior. The final three studies made use of experimental methods and demonstrate that thinking about immoral behaviors attenuates current perceptions of self-knowledge. The predicted relationships and effects generally persist when controlling for self-esteem. Based on our findings, we argue that perceived self-knowledge may play a functional role in moral self-concept maintenance and moral regulatory processes.

> "'I'm afraid I can't explain myself, sir,' said Alice, 'Because I am not myself, you see.'"
>
> —Lewis Carroll

Immoral behavior has historically been a prominent concern for psychologists. Researchers have sought to make sense of highly salient and heinous immoral acts, such as the Holocaust (e.g., Milgram, 1963) and the Kitty Genovese murder (e.g., Darley & Latané, 1968), as well as the processes that dispose people to engage in more mundane forms of immoral conduct (e.g., Sachdeva, Iliev, & Medin, 2009). Although behaving immorally clearly has interpersonal consequences, psychologists have also explored the *intra*personal consequences of moral wrongdoing (e.g., Tsang, 2002). This work largely hinges on the idea that the commission of immoral acts challenges people's fundamental tendency to view themselves as morally good, and, as a result, acting immorally has negative consequences for how people view and feel about themselves (e.g., Lickel, Kushlev, Savalei, Matta, & Schmader, 2014). Building on these ideas, the present research examines the consequence of moral transgressions on feelings of self-knowledge.

The perceived true self-knowledge construct refers to people's subjective beliefs about how well they know who they *truly* are (Kernis & Goldman, 2006; Wood, Linley, Maltby, Baliousis, & Joseph, 2008). The subjective feature of the construct is critical. Whereas other work on self-knowledge focuses on the accuracy of people's knowledge about themselves (e.g., Silvia & Gendolla, 2001), research on perceived self-knowledge focuses squarely on the importance of people's beliefs about who they are regardless of whether or not those beliefs are correct. Although previous studies demonstrate a robust link between perceived true-self knowledge and psychological well-being (e.g., Schlegel & Hicks, 2011; Wood et al., 2008), recent research suggests it may also be closely related to self-perceptions of morality (e.g., Strohminger & Nichols, 2014). The current research

[1]Texas A&M University, College Station, USA
[2]Montana State University, Bozeman, USA

Corresponding Author:
Joshua A. Hicks, Department of Psychology, Texas A&M University, 4235 TAMU, College Station, TX 77843, USA.
Email: joshua.hicks@tamu.edu

builds off this emerging evidence to test the novel hypothesis that recollections of one's morally relevant behaviors affect perceptions of self-knowledge.

Moral Behaviors, Person Perception, and Personal Identity

The current studies extend a body of recent work that has documented a close psychological relationship between morality, person perception, and identity. For example, Wojciszke (2005) concluded that perceptions of moral character largely determine global impressions of others and that people automatically construe others' behavior in moral terms. Similarly, Goodwin, Piazza, and Rozin (2014) found that information about moral character traits (e.g., honesty, courage, generosity) was consistently more predictive of identity-related judgments than other kinds of personality information.

Building on these findings, recent work by Strohminger and Nichols (2014, 2015) has provided evidence that moral character is central to how people understand personal identity, an idea they termed the *essential moral self hypothesis*. In an initial series of studies (Strohminger & Nichols, 2014), participants were more likely to judge that a hypothetical target was no longer the same person following changes to the target's moral character traits, compared with changes to other kinds of personal characteristics (e.g., non-moral personality traits, preferences/desires, knowledge/memories, cognitive/perceptual faculties). Furthermore, when considering cases in which the target's identity was transferred from one body to another, moral character traits were judged as being more likely to survive these transitions than other categories of personal characteristics.

More recently, these findings were replicated in samples of friends and family members of individuals with neurodegenerative disorders (Strohminger & Nichols, 2015). Friends and family members of individuals with frontotemporal dementia, Alzheimer's disease, and amyotrophic lateral sclerosis (ALS) were surveyed. Across all three disorders, perceived change in the patient's moral character traits was the strongest predictor of perceived change in the patient's personal identity (see also Heiphetz, Strohminger, & Young, 2016).

Consistent with these findings, recent research has demonstrated that judgments about *true selves* are constrained by moral values (Newman, Bloom, & Knobe, 2014; Newman, De Freitas, & Knobe, 2015). For instance, participants are significantly more likely to regard egalitarian impulses as stemming from a target's true self than racist impulses (Newman et al., 2015). These results indicate that people believe that true selves are morally good, "that deep inside every individual, there is a 'true self' motivating him or her to behave in ways that are virtuous" (Newman et al., 2014, p. 211).

In recent work, Gino, Kouchaki, and Galinsky (2015) examined the relationship between authentic versus inauthentic behaviors (i.e., behaviors that either are or are not true to one's self) and moral self-regard. Across five studies,

participants felt worse about their moral character in situations where they had behaved inauthentically. Self-knowledge is typically conceived of as a component of the higher order construct of authenticity (e.g., Wood et al., 2008). As such, Gino and colleagues' (2015) findings are highly consistent with the hypotheses of the present study. However, where Gino et al. focused on authenticity of *behavior* as a predictor of *perceptions* of morality, the present research focuses on morality of *behavior* as a predictor of *perceptions* of a component of authenticity (i.e., subjective self-knowledge). Given that perceived self-knowledge has been implicated in a number of important outcomes (e.g., Schlegel, Hicks, King, & Arndt, 2011), the present studies have the potential to illuminate a highly consequential result of perceiving one's behaviors as immoral.

Taken together, these findings indicate that moral character traits are perceived as fundamental to personal identity. That is, folk reasoning about personal identity seems to be anchored by moral character; moral information guides global evaluations of others (Goodwin et al., 2014; Wojciszke, 2005), and changes in moral character are sufficient to change the perceived identity of the person in question (Strohminger & Nichols, 2014, 2015). To the extent that people assume that true selves are morally good (Newman et al., 2014, 2015), we hypothesize that when a person's own moral character is called into question, he or she will experience reduced feelings of self-knowledge.

Overview of the Present Studies

We conducted four studies to test the hypothesis that perceptions of one's moral failings and successes will influence perceived self-knowledge. First, we report a daily diary study in which we examined the proposed relationships over a 5-day period. Studies 2 and 3 manipulated the cognitive accessibility of past morally valenced behavior. Finally, Study 4 utilized false feedback to directly manipulate the perceived morality of one's behavior in a decision-making task. Notably, we control for self-esteem in all four studies to determine whether reminders of immoral acts contribute to lowered self-knowledge over and above simply feeling bad about oneself.

Sample sizes for the present studies were determined as follows. For Study 1, we sought to collect a sample of at least 175, similar to previous daily-diary studies of subjective self-knowledge (e.g., Schlegel, Hicks, Davis, Hirsch, & Smith, 2013). For the experimental studies (Studies 2-4), we sought to collect at least 100 participants per condition, working within the constraints of a 1-week time frame for each lab study and available lab resources.

Study 1[1]

The goal of Study 1 was to test whether fluctuations in self-knowledge reliably covary with fluctuations in the perceived

morality of behavior. We conducted a daily diary study in which participants reported the moral valence of their behavior and their perceived self-knowledge each day for 5 consecutive days. This methodology allowed us to directly examine whether people felt less self-knowledge on days when they perceived that they had behaved immorally.

Method

Participants. One-hundred eighty-three[2] undergraduates (60% women, M_{age} = 18.92, *SD* = 1.19) recruited from a psychology participant pool participated for partial completion of course requirements. Participants were predominantly White (78%) and non-Hispanic (77%).

Materials and procedure. Participants completed a brief online survey each day for 5 consecutive weekdays (Monday-Friday). Each day, an email was sent to participants in the late afternoon containing a link to that day's survey.[3] The survey remained open to participants until midnight. Response rates were satisfactory, with 93% of participants completing at least three daily surveys, and 84% of participants completing four or five daily surveys.

The five surveys were identical and comprised the following measures, presented in the following order:

Moral behavior. Participants were provided with a list of 15 words and asked, "In the past 24 hours, how much did each of the following words describe your behavior?" The words consisted of five moral words (*fair, courageous, honest, helpful, hard-working*) based on the characteristics that comprise the Moral Identity Questionnaire (Aquino & Reed, 2002) and five face-valid immoral words (*cruel, malicious, uncooperative, untrustworthy, selfish*). Five neutral behaviors (e.g., *conventional*) were included to obscure the purpose of the study. Responses were made on a 7-point scale (1 = *not at all*, 4 = *moderately*, 7 = *very much*). Composite scores were created for responses to the moral (*M* = 4.81, *SD* = 1.02, α = .79) and immoral adjectives (*M* = 2.03, *SD* = 1.01, α = .84).

Self-knowledge. Self-knowledge was assessed using the four-item Self-Alienation subscale of the Authenticity Scale (Wood et al., 2008) and four items from the Awareness subscale of the Authenticity Inventory (Kernis & Goldman, 2006). Responses for the Self-Alienation subscale (e.g., "I feel as if I don't know myself very well") and the Awareness subscale (e.g., "I am able to distinguish those self-aspects that are important to my core or true-self from those that are unimportant") were made on a 7-point scale (1 = *strongly disagree*, 7 = *strongly agree*). The self-alienation items were reverse-coded and averaged with the responses for the Self-Awareness subscale to yield a composite measure in which higher scores indicated greater self-knowledge (*M* = 5.32, *SD* = 1.06, α = .87).

Self-esteem. Self-esteem was assessed using the Rosenberg (1965) Self-Esteem scale. Participants indicated their agreement with 10 statements (e.g., "On the whole, I am satisfied with myself" and "I feel that I have a number of good qualities") using a 7-point scale (1 = *strongly disagree*, 7 = *strongly agree, M* = 5.27, *SD* = 1.07, α = .90).

Results and Discussion

To assess relationships between daily moral and immoral behaviors and subjective self-knowledge, we conducted hierarchical linear modeling (HLM) using HLM7 software (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011). This multilevel modeling technique accounts for the lack of independence among repeated within-person observations. Two levels were included in this analysis. Level 1 represented daily responses nested within individuals, and Level 2 represented mean differences between individuals. All predictors were centered within-person to control for between-person differences in mean levels of these variables (Bryk & Raudenbush, 1992; Fleeson, 2007), thus allowing us to examine the purely within-person relationships. No Level 2 predictors were included in these models. Following guidelines from Rosnow, Rosenthal, and Rubin (2000; Equation 2.5), we used the obtained *t* and *df* to calculate effect size *r* coefficients.

First, we estimated an unconditional model to calculate the intraclass correlation coefficient (e.g., Snijders & Bosker, 1999). According to this model, approximately 72% of the variance in subjective self-knowledge was at Level 2 (between individuals). Approximately 28% of the variance was at Level 1 (within individuals, across days). These results indicate that levels of subjective self-knowledge vary more between different people than within the same person over the course of 5 days. However, the presence of a substantial proportion of variance at Level 1 justified our subsequent efforts to fit a model predicting this variance.

We next estimated a model using restricted maximum likelihood that included daily reports of moral and immoral behavior as predictors of daily self-knowledge (all effects were estimated as random). Results of this model revealed a significant, negative relationship between ratings of immoral behavior and self-knowledge, *b* = −.15, standard error (*SE*) = .04, *t*(178) = −4.33, *p* < .001, *r* = .31, such that immoral behaviors were associated with reduced self-knowledge. A positive relationship was observed between ratings of moral behavior and self-knowledge, *b* = .24, *SE* = .04, *t*(178) = 6.71, *p* < .001, *r* = .45, such that moral behaviors were associated with greater self-knowledge.

We followed up on this initial analysis by estimating another model, this time including self-esteem as a predictor in addition to moral and immoral behavior ratings. The two-way interactions between self-esteem and morality/immorality were also included as predictors in this model, to fully account for the variance in self-knowledge

associated with self-esteem. Effects of moral and immoral behavior were estimated as random, while the effects of self-esteem and the interaction terms were estimated as fixed. Results of this model indicated that self-esteem was positively associated with self-knowledge, $b = .43$, $SE = .05$, $t(255) = 8.07$, $p < .001$, $r = .45$. Even when accounting for self-esteem, however, both moral behavior, $b = .18$, $SE = .03$, $t(178) = 5.83$, $p < .001$, $r = .40$, and immoral behavior, $b = −.09$, $SE = .04$, $t(178) = −2.34$, $p = .021$, $r = .17$, significantly predicted self-knowledge. Neither of the interactions between self-esteem and moral or immoral behavior ratings was significant (both $ps > .60$).

These results are consistent with our hypothesis that immoral behavior is associated with reduced self-knowledge; daily levels of moral and immoral behavior reliably covaried with daily levels of self-knowledge. The fact that both of these relationships persisted when self-esteem was included in the model suggests that they cannot be explained simply in terms of moral behavior promoting favorable self-evaluations.

Study 2

Our findings in Study 1 provide preliminary support for our hypothesis that the perceived morality of behavior affects subjective self-knowledge. In Study 2, we aimed to directly test this causal relationship by manipulating the cognitive accessibility (e.g., Schwarz & Strack, 1999) of past morally valenced behavior. This was accomplished by having participants complete a checklist in which they indicated whether they had ever committed various behaviors. Participants were randomly assigned to either a moral, immoral, or neutral version of the checklist. We predicted that thinking about one's past immoral behaviors would attenuate feelings of self-knowledge compared with thinking about morally good acts or neutral behaviors.

Method

Participants. Four-hundred eighty-eight participants (221 women, three unreported; $M_{age} = 31.95$, $SD = 17.15$) were recruited from Amazon's Mechanical Turk platform (Buhrmester, Kwang, & Gosling, 2011). They were paid US$ 0.50 for their participation. Participants were predominantly White (79%) and non-Hispanic (91%).

Materials and procedure

Moral salience manipulation. Participants were randomly assigned into one of three conditions. In the immoral behavior condition, participants read a list of 20 immoral behaviors including modified items taken from the Conventional Morality Scale (Tooke & Ickes, 1988), the Moralization of Everyday Life Scale (Lovett, Jordan, & Wiltermuth, 2012), and the Aggression Questionnaire (Buss & Perry, 1992). They indicated whether they had ever committed

each behavior in the past. Examples of immoral behaviors included, "I have lied to my parents about something," "I have made an offhanded racist or sexist comment," and "I have physically hurt another person." On average, participants indicated committing 10.69 ($SD = 3.98$) behaviors.

In the moral behavior condition, participants read a list of 20 moral behaviors (e.g., "I have returned a valuable item that I found, rather than keeping it for myself.") and indicated whether they had ever committed each behavior in the past ($M = 17.67$ items selected, $SD = 2.68$).

In the neutral control condition, participants read a list of 20 common grocery-shopping behaviors (e.g., "I have used the self-checkout lane at a grocery store.") and indicated whether they had ever committed each behavior in the past ($M = 13.97$ items selected, $SD = 3.16$).

*Self-knowledge.* After completing the manipulation, participants completed the same Self-Alienation subscale (Wood et al., 2008) as Study 1 and the full 12-item Self-Awareness subscale of the Authenticity Inventory (Kernis & Goldman, 2006). Self-alienation was assessed using a 7-point scale, while self-awareness was assessed using a 5-point scale. We reverse-coded the self-alienation items and appropriate items from the Self-Awareness scale, standardized both measures at the item level, and averaged the standardized items to yield a composite measure in which higher values reflect greater self-knowledge ($M = −0.001$, $SD = 0.65$, $α = .91$).

Self-esteem. Participants completed the Rosenberg (1965) Self-Esteem scale ($M = 4.87$, $SD = 1.30$, $α = .94$) prior to completing demographic measures and being debriefed.

Results and Discussion

As a test of our main hypothesis, we ran a one-way ANOVA to determine whether self-knowledge varied by condition. The results revealed significant differences[4] in self-knowledge across the conditions, $F(2, 485) = 4.59$, $p = .011$, partial $η^2 = .019$. Post hoc analyses revealed that participants in the immoral condition reported significantly less self-knowledge ($M = −0.09$, $SD = 0.63$) compared with the moral condition ($M = 0.12$, $SD = 0.67$, $p = .004$; 95% confidence interval [CI]$_{difference} = [.07, .35]$). Unexpectedly, the immoral and neutral conditions ($M = −0.04$, $SD = 0.64$) were not significantly different from each other ($p = .476$; 95% CI$_{difference} = [−.19, .09]$), although the moral and neutral conditions did significantly differ from each other ($p = .028$; 95% CI$_{difference} = [.02, .30]$).

We next computed an ANCOVA including self-esteem as a covariate. Results indicated that self-esteem significantly predicted self-alienation, $F(1, 481) = 338.77$, $p < .001$, partial $η^2 = .41$. The omnibus effect of condition remained significant[5] after accounting for self-esteem, $F(2, 481) = 3.86$, $p = .022$, partial $η^2 = .016$. Controlling for self-esteem, the difference between the immoral ($M = −0.08$, $SD = 0.49$) and
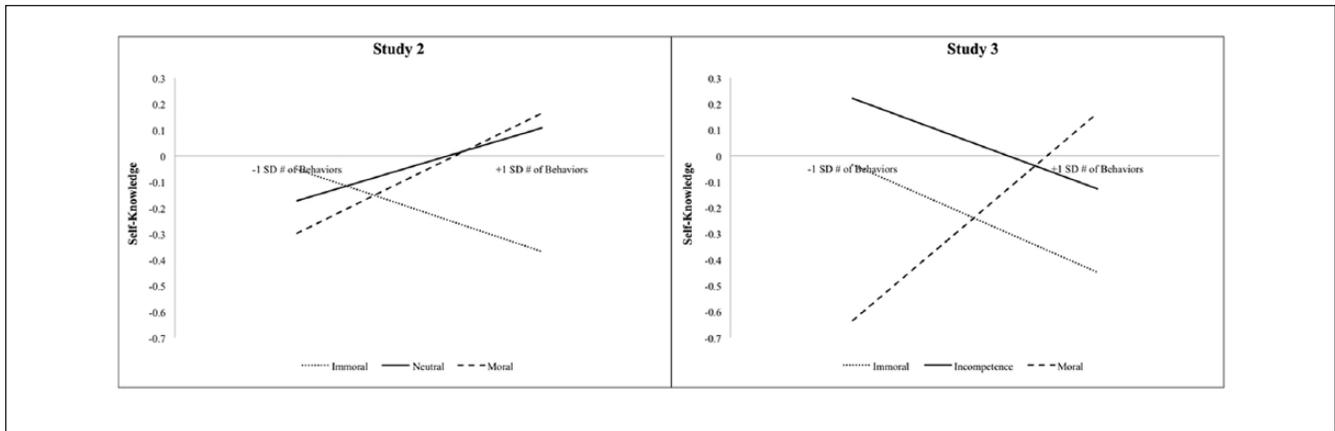
Figure 1. Interactions between condition and number of behaviors endorsed (no covariates), Studies 2 and 3.

moral conditions ($M = 0.07$, $SD = 0.49$) remained significant ($p = .006$, 95% CI$_{\text{difference}}$ = [.04, .26]). The difference between the immoral and neutral conditions ($M = 0.01$, $SD = 0.70$) remained non-significant ($p = .101$, 95% CI$_{\text{difference}}$ = [−.20, .02]), although the mean difference was in the predicted direction. No differences were observed between the moral and neutral conditions after controlling for self-esteem ($p = .258$, 95% CI$_{\text{difference}}$ = [−.05, .17]).[6] The neutral condition ($M = −0.01$, $SD = 0.50$) did not differ from either the immoral condition ($p = .258$, 95% CI$_{\text{difference}}$ = [−.05, .17]) or the moral condition ($p = .101$, 95% CI$_{\text{difference}}$ = [−.02, .20]).

We explored whether the effect of the manipulation differed depending on the number of behaviors participants endorsed by conducting a regression analysis assessing the interaction between number of behaviors endorsed and condition. Indeed, in this study, our analysis predicts that making these behaviors salient should most strongly affect people who have actually committed more of the behaviors in question. We therefore entered two dummy-coded condition variables (Dummy 1: Moral/Neutral = 0, Immoral = 1; Dummy 2: Immoral/Neutral = 0, Moral = 1) to capture the main effect of condition and the number of behaviors (mean-centered) in Step 1. The two-way interaction terms were entered in Step 2. The results revealed that the main effects accounted for a significant amount of variance in Step 1 ($\Delta R^2 = .019$, $p = .026$). However, this was qualified by a significant Condition × Number of Behaviors interaction in Step 2 ($\Delta R^2 = .04$, $p < .001$).[7]

To probe this effect, we conducted predicted means tests that compared the experimental condition effects at high and low (±1 *SD*) numbers of behaviors endorsed. At low (−1 *SD*) numbers of behaviors endorsed, there were no significant differences in self-knowledge between the experimental conditions ($|\beta s| < .13$, $ps \geq .220$). At high (+1 *SD*) numbers of behaviors endorsed, however, significant differences consistent with our hypotheses emerged. Self-knowledge was reduced in the immoral condition relative to both the moral ($\beta = −.39$, $p < .001$) and neutral conditions ($\beta = −.35$, $p = .001$). In contrast, there was no difference in self-knowledge

between the moral and neutral conditions ($\beta = .04$, $p = .580$). These results are depicted in Figure 1.

We also conducted a similar regression analysis controlling for differences in self-esteem. Self-esteem accounted for significant variance in self-knowledge at Step 1 ($\Delta R^2 = .42$, $p < .001$), and adding the main effects of condition and number of behaviors reported in Step 2 accounted for a significant increase in the amount of variance explained ($\Delta R^2 = .013$, $p = .014$). In addition, a significant Condition × Number of Behaviors interaction in Step 3 emerged ($\Delta R^2 = .008$, $p = .035$; see Figure 2).[8] Mirroring the effects above, there were no significant differences in self-knowledge between the experimental conditions at low (−1 *SD*) numbers of behaviors endorsed ($|\beta s| < .04$, $ps > .519$). At high (+1 *SD*) numbers of behaviors endorsed, however, self-knowledge was lower in the immoral condition relative to both the moral ($\beta = .16$, $p = .019$) and neutral conditions ($\beta = .21$, $p = .010$). Self-knowledge did not differ between the moral and neutral conditions ($\beta = .04$, $p = .461$). Simple slopes for the relationships between number of behaviors and self-knowledge in each condition in both regression models can be found in Table S2 in the online supplement.

As expected, moral and immoral behavior was significantly linked to self-knowledge in the predicted directions. Moral behaviors predicted greater self-knowledge, and immoral behaviors predicted less self-knowledge. Critically, and perhaps unsurprisingly, the number of immoral and moral behaviors people reported appears to moderate this effect. Participants whose immoral behaviors were made salient reported reduced feelings of true self-knowledge compared with the moral and neutral conditions; this was particularly true for individuals who endorsed a greater number of immoral acts.

Study 3

The findings from Study 2 demonstrate that reflecting on morally relevant behaviors influenced perceived self-knowledge.
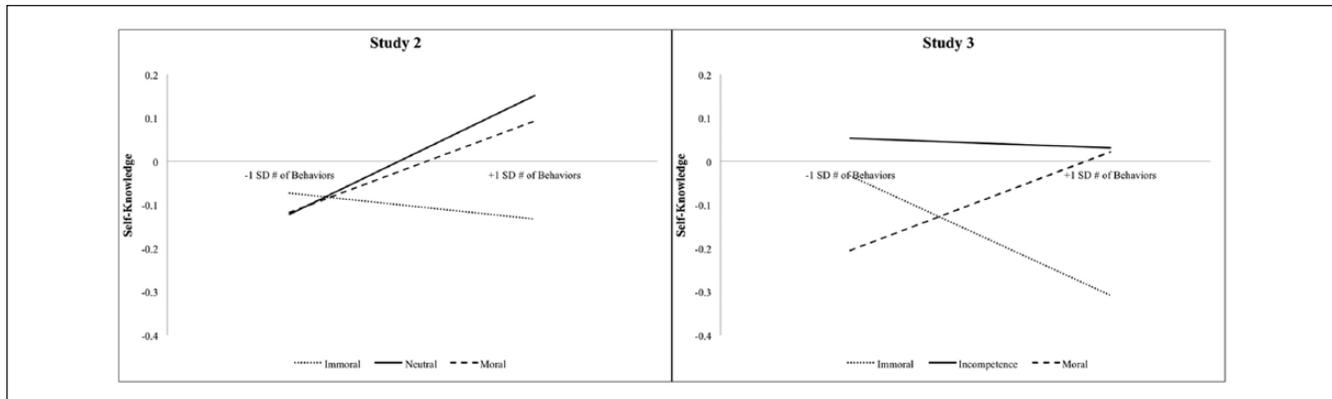
Figure 2. Interactions between condition and number of behaviors (controlling self-esteem), Studies 2 and 3.

This effect was particularly pronounced for those who indicated committing many of the respective behaviors. Although this interaction effect is in line with our hypotheses, we did not predict this pattern of results a priori. As such, the main purpose of Study 3 was to replicate this finding. A secondary purpose of Study 3, however, was to test whether other self-threats or self-inconsistencies would also reduce subjective self-knowledge. In our view, moral behaviors should directly influence perceived true-self knowledge due to the belief that the true self is morally good and research indicating that moral information is essential to true self-conceptions. Nevertheless, it is possible that any behavior that is threatening to or inconsistent with one's self-views will similarly affect subjective self-knowledge. We tested this possibility in Study 3 by including a condition that had participants reflect on behaviors related to incompetence (e.g., failing an exam). Competence, along with morality, is central to evaluations of others (Fiske, Cuddy, & Glick, 2007). In a review of the literature on person perception and self-perception, Wojciszke (2005) concluded that although people tend to construe others' behavior in moral terms, they tend to construe their own behavior primarily in terms of competence. Therefore, this comparison provides a more robust test of our hypotheses that immoral behaviors are strong predictors of perceived self-knowledge. Overall, we predicted that thinking about immoral behaviors one has committed would lead to reduced self-knowledge compared with thinking about either morally good behaviors or instances of incompetence.

Method

Participants. Three-hundred seventy-eight undergraduates (279 women, one unreported; M_{age} = 19.05, SD = 9.45) recruited from a psychology participant pool participated for partial completion of course requirements. Participants were predominantly White (71%) and non-Hispanic (74%).

Materials and procedure. Participants were informed that they would be participating in a study exploring their personality and attitudes. The procedures were identical to Study 2 except we replaced the neutral control condition with a competence condition to assess how this type of self-threat would influence perceived true self-knowledge compared with threats to one's morally good true self. Participants were randomly assigned into one of three conditions. As in Study 2, participants in the immoral behavior condition completed a checklist of 20 immoral behaviors (M = 10.21 items selected, SD = 3.64), and participants in the moral behaviors condition completed a checklist of 20 moral behaviors (M = 18.08 items selected, SD = 1.86). In the incompetence condition, participants indicated whether they had committed each of 20 incompetence-related behaviors (e.g., "I have forgotten about an appointment or assignment") in the past (M = 16.63 items selected, SD = 2.98).

Participants then completed the same measures of self-knowledge described in Study 2. Again, a composite score for self-knowledge was created by reverse-coding the self-alienation items, standardizing both scales at the item level, and averaging across the standardized items (M = 0.00, SD = 0.58, α = .87). After completing the self-knowledge measures, participants completed the Rosenberg Self-Esteem Scale (M = 5.04, SD = 1.14, α = .91).

Finally, participants in the immoral and incompetence conditions rated the perceived severity of any behaviors they had endorsed on a 7-point scale (*1 = not severe at all*, *7 = extremely severe*). This step was taken to account for the possibility that effects on self-knowledge are driven by the perception that immoral behaviors are more severe than failures of competence, rather than by the centrality of the information to the self per se.

Results and Discussion

Preliminary analyses. We first ran an independent-samples *t* test to examine whether severity ratings differed between the immoral and incompetence conditions. A significant difference was found, $t(250)$ = 3.26, p = .001, d = .41, but unexpectedly, the mean severity rating was higher in the

incompetence condition ($M = 3.29$, $SD = 1.02$) than in the immoral condition ($M = 2.86$, $SD = 1.07$).[9] This rules out the possibility that immoral behaviors might have a greater impact on self-knowledge simply because they are perceived as more severe.

Primary analyses. We ran a one-way ANOVA to determine whether self-knowledge varied by condition. In contrast to Study 2, a main effect of condition on self-knowledge was not observed, $F(2, 375) = 0.51$, $p = .603$, partial $\eta^2 = .003$.

We next conducted regression analyses examining whether the experimental manipulation interacted with the number of behaviors participants endorsed. We again entered two dummy-coded condition variables to capture the main effect of condition, and the number of behaviors (mean-centered) in Step 1. The two-way interaction terms were entered in Step 2. The results revealed that the main effects accounted for a marginally significant amount of variance in Step 1 ($\Delta R^2 = .019$, $p = .067$). However, this was qualified by a significant Condition × Number of Behaviors interaction in Step 2 ($\Delta R^2 = .051$, $p < .001$).[10]

We conducted predicted means tests that compared the simple effects of experimental condition at high and low (± 1 *SD*) numbers of behaviors endorsed (see second panel of Figure 1). At low (−1 *SD*) numbers of behaviors endorsed, participants in the immoral condition reported less self-knowledge than those in the incompetence condition ($\beta = -.21$, $p = .044$) and more self-knowledge than those in the moral condition ($\beta = .49$, $p = .006$). Participants in the moral condition also reported significantly less self-knowledge than those in the incompetence condition ($\beta = -.70$, $p < .001$) at low (−1 *SD*) numbers of behaviors endorsed. In contrast, at high (+1 *SD*) numbers of behaviors endorsed, self-knowledge was lower in the immoral condition relative to both the moral ($\beta = -.50$, $p < .001$) and incompetence conditions ($\beta = -.26$, $p = .037$). Self-knowledge was also significantly lower in the incompetence condition relative to the moral condition ($\beta = -.24$, $p = .002$).

As in Study 2, we also ran a parallel regression analysis that controlled for differences in self-esteem. Self-esteem accounted for significant variance in self-knowledge a Step 1 ($\Delta R^2 = .451$, $p < .001$), and adding the main effects of condition and number of behaviors reported in Step 2 did not result in a significant increase in the amount of variance explained ($\Delta R^2 = .006$, $p = .233$). Critically, however, a significant Condition × Number of Behaviors interaction in Step 3 emerged[11] ($\Delta R^2 = .010$, $p = .033$; see Figure 2). At low (−1 *SD*) numbers of behaviors endorsed, participants in the immoral condition did not differ in self-knowledge compared with the incompetence condition ($\beta = -.07$, $p = .371$) and the moral condition ($\beta = .17$, $p = .306$). In addition, participants in the moral condition did not differ in self-knowledge relative to those in the incompetence condition ($\beta = -.26$, $p = .169$) at low (−1 *SD*) numbers of behaviors endorsed. In contrast, at high (+1 *SD*) numbers of behaviors

endorsed, self-knowledge was lower in the immoral condition relative to both the moral ($\beta = -.28$, $p = .003$) and incompetence conditions ($\beta = -.27$, $p = .005$). Self-knowledge did not differ between the incompetence condition and moral conditions ($\beta = .01$, $p = .893$). Simple slopes from both regression models are presented in Table S2 in the online supplement.

These results are consistent with our hypothesis. Although we did not find a main effect of condition on self-knowledge in Study 3, this is not necessarily surprising. Depending on the number of behaviors participants endorse, the effects of condition could vary considerably. The two negative conditions (immoral and incompetence) pose much less of a threat to participants who endorse very few of the items. Endorsing very few items in these conditions may function as a self-affirmation, reinforcing participants' confidence that they are moral or competent individuals. In contrast, the opposite pattern should hold in the moral condition—participants who endorse very few items may be threatened, because their responses imply that they have acted less morally than possible.

This reasoning is consistent with the observed patterns of results, particularly for the regression model that did not control for self-esteem. This model indicated that self-knowledge was attenuated in the moral condition at low levels of behavior relative to both the competence and immoral conditions. At high levels of behavior, however, the immoral condition exhibited less self-knowledge than both the moral and incompetence conditions. Although the incompetence condition also reduced self-knowledge relative to the moral condition at high levels of behavior, this effect disappeared when self-esteem was statistically controlled. Controlling for self-esteem did not eliminate the effects of the immoral condition on self-knowledge. In addition, when controlling for self-esteem, the relationship between number of behaviors and self-alienation was only significant in the immoral condition (see Table S2 of the supplement for simple slopes), suggesting that immorality may have a qualitatively distinct relationship with self-knowledge that is not reducible to self-esteem threat. In contrast, these results imply that morally good and competence-related behaviors may bear more directly on self-esteem than self-knowledge.

Study 4

The results of Studies 2 and 3 demonstrate that reminders of past immoral behavior lead to a decline in subjective self-knowledge. In both of these studies, participants were asked to consider whether they had ever committed a relatively large number of behaviors at any prior point in their lives. Therefore, the results of these studies do not directly capture how people might respond in the moment to a single instance of immoral behavior. We attempted to provide this evidence in Study 4 using a generosity-based false feedback paradigm. Generosity is a key ingredient of moral character

and virtue (Peterson & Seligman, 2004). Participants made a series of hypothetical decisions between donating money and receiving money to keep for themselves and were then provided with false feedback on how their responses compared with those of other students at their university. Participants were randomly assigned to two feedback conditions. In the immoral condition, participants were told that the majority of their fellow students had been more generous on the same task, while in the moral condition, participants were told that they had been more generous than most other students. We predicted that participants in the immoral condition would exhibit reduced self-knowledge compared with those in the moral condition.

Method

Participants. One-hundred twenty-eight undergraduates[12] (76 women; M_{age} = 19.31, SD = 1.63) recruited from a psychology participant pool participated for partial completion of course requirements. Participants were predominantly White (57.8%) and non-Hispanic (78.1%).

Materials and procedure. Participants either completed the study online or in the laboratory. Participants were informed that they would be participating in a study exploring their decision-making behaviors and personality.

Monetary decision task. Participants were first informed that they would complete a decision-making task involving money. They read that this decision-making task has been administered to other students at their university as part of a large-scale study of college students' monetary decision making. They also read that, at the end of the task, their data would be processed and they would receive personalized feedback about how their responses compare with the average student at their university.

Participants completed a series of 44 trials in which they were asked to imagine that a third party was giving them the opportunity to either receive some money to keep for themselves or to have a charitable donation made in their name, and that they had to choose one of these alternatives. The instructions further specified that decisions to keep money for oneself were final, and that kept money could not be later donated to a charitable organization. The trials varied with respect to how much money was allocated to each response option (e.g., "receive $50 to keep for yourself versus a $50 donation made in your name to a worthy charity," "receive $100 to keep for yourself versus a $200 donation made in your name to a worthy charity") and were presented in a random order. Descriptive statistics indicated a roughly even split between decisions to keep for oneself and decisions to donate to charity (M = 26.35 decisions to donate, SD = 11.39).

Morality manipulation. Participants were randomly assigned to the immoral or moral condition. The personalized feedback

for the decision-making task served as our manipulation. Participants were provided with a mock distribution of all the responses from students at their university and their unique position in the distribution. In the immoral condition, participants were told they were "in the *10th percentile* in terms of decisions to donate. That is, 90% of A&M students made *more* decisions to donate to charity than you did." In the moral condition, participants were told they were "in the *91st percentile* in terms of decisions to donate. That is, 90% of A&M students made *fewer* decisions to donate to charity than you did."

Manipulation check. To assess the manipulation's efficacy, participants were provided with a list of 12 words and asked to indicate how well each word describes the choices they made in the decision-making task. Five of the words (e.g., *fair, selfish, generous, moral*, and *immoral*) served as our manipulation check items, and the rest were filler items (e.g., *logical*). Responses followed a 7-point scale (1 = *not at all*, 4 = *moderately*, 7 = *perfectly*). Responses for "selfish" and "immoral" were reverse-coded, and all items were averaged together to create a composite variable in which higher scores indicated greater perceived morality (M = 4.48, SD = 0.75, α = .80).

Self-knowledge. As in the previous studies, participants completed the Self-Alienation and Self-Awareness subscales to assess perceived self-knowledge. The self-alienation items and appropriate items from the Self-Awareness scale were reverse-coded, and all items were standardized and averaged together to form a composite self-knowledge variable (M = 0.00, SD = 0.61, α = .88).

Self-esteem. As in the previous studies, participants completed the Rosenberg Self-Esteem scale (M = 5.08, SD = 1.10, α = .90) prior to completing additional exploratory measures and demographic items and being debriefed.

Results and Discussion

Preliminary analyses. We ran an independent-samples t test on the manipulation check items to verify that the manipulation was effective. There was a significant difference in perceived morality, $t(126)$ = 4.078, p < .001, d = .72; 95% $CI_{difference}$ = [.26, .75], such that participants receiving moral feedback reported that they had behaved more morally during the task (M = 4.76, SD = 0.70) than participants who received the immoral feedback (M = 4.26, SD = 0.73).

Primary analyses. For our main analysis, we ran an independent-samples t test on our dependent variable. As predicted, there was a significant difference in self-knowledge between conditions, $t(126)$ = 3.09, p = .002, d = .55; 95% CI = [.12, .53], such that participants in the immoral condition reported feeling less self-knowledge (M = −0.16, SD = 0.55) than
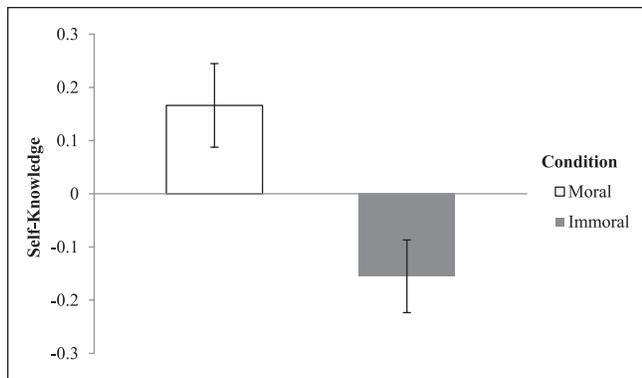
Figure 3. Differences in self-knowledge by condition in Study 4.
Note. Error bars represent standard errors.

participants in the moral condition ($M = 0.17$, $SD = 0.62$; see Figure 3).

Finally, an ANCOVA revealed that the manipulation marginally predicted true self-knowledge over and above the contribution of self-esteem, $F(1, 124) = 3.35$, $p = .070$, partial $\eta^2 = 0.03$, 95% $CI_{difference} = [-.31, .01]$. This result was statistically significant when multivariate outliers were excluded from the analysis.[13]

The results of Study 4 provide further support for the hypothesized relationship between the perceived moral valence of behavior and subjective feelings of self-alienation. Whereas the first three studies dealt with aggregated perceptions of behavioral morality over the previous day (Study 1) and over participants' entire lives (Studies 2 and 3), Study 4 demonstrated that the perceived moral valence of a single recent behavior causally affects perceptions of self-knowledge. Participants who were led to believe they had behaved selfishly on the monetary decision task reported lower levels of self-knowledge.

General Discussion

Across four studies, we found robust evidence that the perceived moral valence of behavior is associated with, and influences, subjective feelings of self-knowledge. In Study 1, daily reports of both moral and immoral behaviors significantly covaried with daily reports of self-knowledge, suggesting that people feel more or less in touch with their true self depending on how morally/immorally they believed they behaved. Similarly, two experimental studies demonstrated that reflecting on past morally valenced behaviors elicited corresponding changes in perceived self-knowledge. Recalling past immoral behaviors reduced self-knowledge relative to recalling moral behaviors or incompetent behaviors. These latter effects were moderated by the number of behaviors that participants had endorsed, which suggests that the manipulation was either (a) only effective to the extent that it resonated with participants or (b) that the interaction capitalized on existing individual differences. Study 4

resolved this issue by manipulating perceptions of behavior on the same task and finding that experimentally induced feelings of immorality (vs. morality) led to lower self-knowledge.[14] Taken together, these findings suggest that when people feel they have committed a moral transgression, they feel uncertain of who they truly are, and, conversely, when they perceive their behavior as moral, they experience feelings of self-understanding.

Implications for Research on the True Self, Moral Self-Regulation, and Moral Emotions

The true self. Our findings add to the growing body of evidence indicating that moral information is central to reasoning about the self, the *true self* in particular. Most people seem to operate on the assumption that their own true self (Bench, Schlegel, Davis, & Vess, 2015) and the true selves of others are morally good (Newman et al., 2014, 2015). This widespread assumption fits with the present findings as instances of immoral conduct directly contradict the assumed goodness of the true self. Moral information carries a great deal of weight in reasoning about identity (e.g., Goodwin et al., 2014; Strohminger & Nichols, 2014, 2015), making doubts about one's morality especially likely to spill over into a more generalized self-doubt. Our findings suggest that this is the case and represent some of the first evidence of the influence of moral information on perceptions of one's *own* self (see also Heiphetz et al., 2016).

Moral self-regulation. The current findings are consistent with a broader theoretical contention that the true self plays a functional role in moral self-regulation. Specifically, self-alienation following immoral deeds may signal to people that they have done something wrong. Indeed, Gino et al. (2015) found that feelings of inauthenticity reduce moral self-regard. Thus, rather than having to be explicitly aware of the ways in which their action has violated norms, people may simply have a sense that what they have done is "not them." In this way, feelings of self-alienation may serve as a "moral barometer" much like feelings of shame and guilt (Tangney, Stuewig, & Mashek, 2007). It is plausible that self-alienation in these situations functions to motivate prosocial compensatory responses, ultimately bringing behavior back into alignment with moral norms. However, another possibility is that self-alienation is a mechanism of moral rationalization (e.g., Tsang, 2002), enabling people to distance themselves from the threatening implications of their own immoral conduct. If this is the case, denying self-knowledge may allow people to resolve the self-threat posed by immoral behavior; consequently, self-alienation may actually lead to reduced prosocial intent following immoral behavior.

Interestingly, Gino and colleagues (2015) found that inauthentic behavior activates compensatory motivations to engage in prosocial behavior. This suggests that the true self has a clear role to play in moral self-regulation, as feelings of

alienation may signal that one has behaved immorally and motivate reparative action. For example, in a typical hypocrisy induction (e.g., Stone, Aronson, Crain, Winslow, & Fried, 1994), participants affirm their commitment to a particular value (e.g., safe sex practices) and are asked to reflect on past instances in which they have failed to live up to this value. Based on the present findings, we would predict that reflecting on these past failures would reduce perceived self-knowledge. We suspect that reduced self-knowledge contributes to the subsequent motivation to behave better, similar to typical findings in hypocrisy research. Similarly, people should become motivated to self-verify (i.e., to confirm pre-existing beliefs about themselves; Swann & Brooks, 2012) when their sense of self-knowledge is called into doubt. Behaving morally and expressing moral intentions are fairly direct means by which people can produce evidence of their morality, confirming their self-views and restoring a sense of self-knowledge. Thus, our findings suggest an avenue by which broader self-related motivations may be recruited in moral self-regulation.

The broadest such motivation may well be cognitive dissonance, which is the mechanism thought to underlie hypocrisy effects (e.g., Stone et al., 1994; Aronson, 1999). The present findings, and the essential moral self-view more broadly, imply that moral transgressions are particularly dissonance-arousing compared with other sorts of self-inconsistency, owing to the centrality of morality to the self-concept. Thus, perceptions of immorality in one's behavior should lead to more pronounced dissonance-reduction efforts than perceptions of other inconsistencies. Given the flexibility of dissonance-reduction processes (e.g., Festinger, 1957), attempts to resolve immorality-induced dissonance might manifest in diverse ways, including the expression of prosocial intent documented by Gino et al. (2015). The present findings may also reflect dissonance processes at work. Reductions in self-knowledge might be symptomatic of dissonance itself (i.e., feelings of self-confusion that naturally accompany exposure to self-inconsistent information), or an attempt to resolve dissonance (i.e., making the self-concept less definitive and determinate, thereby reducing the impact of self-inconsistency).

Shame. A large literature on guilt and shame documents the emotional consequences of behaving immorally (e.g., Tangney et al., 2007). We speculate that perceptions of self-knowledge may mediate the experience of shame. Indeed, some work suggests that conditions particularly likely to erode feelings of true-self knowledge (i.e., being driven by others' expectations) trigger guilt-free shame responses in response to personal shortcomings, whereas activating self-knowledge produces shame-free guilt in response to personal shortcomings (Vess, Schlegel, Hicks, & Arndt, 2014). We thus suggest that people experience shame following an immoral action to the extent that the action causes them to feel subjectively distanced from their true self. This is consistent with social–cognitive models of self-conscious emotions, which suggest that feelings such as shame and guilt stem from perceived discrepancies between identity and behavior (e.g., Tracy & Robins, 2004).

In recent work, Lickel et al. (2014) found that, compared with guilt and regret, feelings of shame were uniquely associated with desire for self-change. When considered in light of the present findings, this desire for change could reflect an underlying state of self-alienation. We have suggested that a perceived lack of self-knowledge activates motivations to re-establish certainty about oneself. These motivations may be expressed verbally as a desire for self-change. Our reasoning here again relies on the idea that most people believe that they are fundamentally good—that they possess a moral *true self*.

Limitations and Future Research

Although our findings support the contention that morally valenced behaviors bear on perceived self-knowledge, the present studies leave several questions open for future investigations to address. Perhaps the greatest limitation of the present studies is their reliance on face-valid self-report measures of perceived self-knowledge across all of the studies. This methodological choice was grounded in our desire to operationalize our dependent measure as straightforwardly and unequivocally as possible, because the present studies were our preliminary attempts to investigate this research question. As such, our primary concern was ensuring that our approach was consistent across studies to best gauge the reliability of the observed effects.

The primary limitation associated with this dependent measure, which applies to any self-report measure, is that the accuracy of participants' reports is an open question. Although we favor the view that experiencing self-alienation is a real and direct consequence of perceiving immorality in one's own behavior, the present studies do not provide direct evidence as to the genuineness of participants' reported declines in self-knowledge. Thus, it is possible that the effects observed are the result of self-presentational processes. Rather than genuinely feeling self-alienated, participants in the immoral conditions may simply have been claiming to lack self-knowledge as a means to externalize blame for their transgressions and maintain a morally good façade. This is consistent with Tsang's (2002) argument that there is a "universal value of morality" resulting in a motivation to maintain "the semblance of being moral" (p. 26).

Although we cannot conclusively rule out a self-presentation account of the present findings, there are some considerations we believe speak against such an account. First, the studies were conducted under conditions of anonymity and relative privacy; this should have dampened participants' self-presentational motives, which are primarily activated by the (perceived or real) presence of others (e.g., Baumeister & Hutton, 1987). Furthermore, the results of Study 3 suggest

that the observed effects on perceived self-knowledge are particularly pronounced for immorality, and that non-immoral negative behaviors (i.e., competence failures) do not affect self-knowledge to the same degree (particularly after accounting for self-esteem). A pure self-presentation view would seem to predict a similar response to any instance of negative behavior, whether immoral or incompetent in nature. Beyond Study 3, the fact that the effects on self-knowledge across the present studies generally persist controlling for self-esteem speaks against a strong self-presentation account; the effects do not seem reducible to a desire to maintain favorable self-views.

To conceptually replicate the present findings and more conclusively determine the nature of the underlying processes involved, future research might use more diverse and subtle outcome measures such as a reaction time task (e.g., "me-not me task"; see Schlegel, Hicks, Arndt, & King, 2009), or assess people's desire to increase their self-knowledge following morally relevant behaviors. For example, perhaps, after recalling immoral conduct, people might report a greater desire to learn about who they really are. Moreover, it is possible that reminders of immoral behavior enhance self-verification motives (e.g., Swann & Read, 1981). These diverse dependent measures could provide more converging evidence that recollections of immoral behavior make people feel less in touch with their true selves and initiate a desire to restore self-understanding.

In spite of our use of the same dependent measure across studies, inconsistencies were found in the results. Specifically, while Study 2 found a main effect of condition, this was not observed in Study 3. Although these results raise concerns regarding the reproducibility of our findings, it is important to point out the consistent pattern of interactions between the moral versus immoral conditions and the number of behaviors participants endorsed, as well as the exploratory analyses that revealed main effects of condition when controlling for the number of behaviors endorsed. These findings, along with the results of Study 4, suggest that our effects are likely to replicate only to the extent that participants perceived that they have actually behaved morally (vs. immorally).

Finally, future studies should also aim to identify moderators and boundary conditions for the observed effects. Such efforts could clarify the nature of these effects and the underlying processes involved. A possible moderator that should be explored in future research is the attributions participants make for their immoral behavior. It is plausible that the effects observed in the present study will be limited to cases where individuals attribute their moral transgressions to internal factors, as these are the cases in which the behavior in question becomes potentially self-diagnostic (e.g., Johnson, Robinson, & Mitchell, 2004). In contrast, if immoral behaviors are attributed to external factors, these behaviors become much less informative about the self and should therefore pose less of a threat to individuals' moral self-conceptions.

Conclusion

Benjamin Franklin famously wrote, "There are three things extremely hard: steel, a diamond, and to know one's self" (Franklin, 2004, pp. 195). Although Franklin and many others have emphasized the difficulty of attaining self-knowledge, our research indicates that this endeavor becomes even more challenging in the wake of moral transgressions. By demonstrating that discrepancies between one's moral compass and behavior engender feelings of self-doubt, the present findings draw clear links between research on perceived self-knowledge and moral psychology. Future studies will help clarify whether self-alienation following immoral acts plays an adaptive role in moral regulation, or whether these feelings of uncertainty simply allow us to attribute our misdeeds to something other than who we are "at the core."

Notes

1. See osf.io/npa35 for complete list of measures and items assessed in all studies, appendix, data sets for each study, and supplementary materials, including exploratory analyses and studies not reported in this article due to space limitations.
2. Due to a study programming error, demographic data were not collected until Day 5 of the study. As such, information about gender and race was only obtained for 148 participants out of the original 183. Descriptive statistics in this section reflect the reduced sample of 148.
3. A small number of participants completed the survey twice on the same day (five each on Days 1 and 2, four each on Days 3 and 4, and one on Day 5), likely by accessing the link from multiple IP addresses. Because HLM can accommodate varying numbers of observations across individuals, we opted to include all data in our analyses.
4. This effect remained significant when multivariate outliers with studentized residuals $\geq |3|$ ($n = 1$) were excluded from the analysis ($p = .015$).
5. This result also remained significant after multivariate outliers with studentized residuals $\geq |3|$ ($n = 4$) were excluded from the analysis ($p = .031$).
6. We also ran a regression analysis testing whether self-esteem interacted with condition to predict self-knowledge. Self-esteem was entered in Step 1 of the regression, along with two dummy-coded condition representing condition. The two-way interaction terms were entered on Step 2. The addition of these interaction terms did not result in a significant increase in variance in Study 2 or 3 ($p \geq .234$).
7. These results were unchanged when multivariate outliers with studentized deleted residuals $\geq |3|$ ($n = 2$) were excluded from

the analysis (*p*s = .019 and <.001 for the main effect and interaction effect, respectively).

8. These results were unchanged when multivariate outliers with studentized deleted residuals ≥ |3| (*n* = 1) were excluded from the analysis (*p*s = .037 and .025, for the main effect and interaction effect, respectively).

9. We also obtained severity ratings for all immoral and incompetence items from a separate sample of undergraduate coders (*N* = 14). When coders' ratings were used to compute severity ratings for the items endorsed by each participant, the average was found to be higher in the immoral condition (*M* = 4.52) than in the incompetence condition (*M* = 3.07), *p* < .001.

10. No multivariate outliers with studentized deleted residuals ≥ |3| were found in this analysis.

11. These results were unchanged when multivariate outliers with studentized deleted residuals ≥ |3| (*n* = 2) were excluded from the analysis (*p* = .012, for the interaction term).

12. As noted earlier, it was our aim to collect at least 100 participants per cell, or as many participants as possible in a 1-week time frame. Unfortunately, although we extended data collection well beyond 1 week, the sample we were able to collect fell short of our target.

13. When multivariate outliers with studentized deleted residuals > |3| (*n* = 2) were excluded from the analysis, a significant effect of condition was observed, *p* = .011.

14. We also conducted two other experimental studies that yielded similar results (see supplemental materials).

References

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

Aronson, E. (1999). Dissonance, hypocrisy, and the self-concept. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive dissonance: Progress on a pivotal theory in social psychology* (pp. 103-126). Washington, DC: American Psychological Association.

Baumeister, R. F., & Hutton, D. G. (1987). Self-presentation theory: Self-construction and audience pleasing. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior*. New York, NY: Springer-Verlag.

Bench, S. W., 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.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models in social and behavioral research: Applications and data analysis methods* (1st ed.). Newbury Park, CA: SAGE.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science*, *6*, 3-5.

Buss, A. H., & Perry, M. P. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, *63*, 452-459.

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*, 377-383.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*, 77-83.

Fleeson, W. (2007). Situation-based contingencies underlying trait-content manifestation in behavior. *Journal of Personality*, *75*, 825-862.

Franklin, B. (2004). *Poor Richard's almanack*. New York, NY: Barnes & Noble.

Gino, F., Kouchaki, M., & Galinsky, A. D. (2015). The moral virtue of authenticity: How inauthenticity produces feelings of immorality and impurity. *Psychological Science*, *26*, 983-996.

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*, 148-168.

Heiphetz, L., Strohminger, N., & Young, L. L. (2016). The role of moral beliefs, memories, and preferences in representations of identity. *Cognitive Science*. Advance online publication.

Johnson, J. T., Robinson, M. D., & Mitchell, E. B. (2004). Inferences about the authentic self: When do actions say more than mental states? *Journal of Personality and Social Psychology*, *87*, 615-630.

Kernis, M. H., & Goldman, B. M. (2006). A multicomponent conceptualization of authenticity: Theory and research. In M. P. Zanna, (Ed.), *Advances in experimental social psychology* (Vol. 38, pp. 283-357). San Diego, CA: Elsevier Academic Press.

Lickel, B., Kushlev, K., Savalei, V., Matta, S., & Schmader, T. (2014). Shame and the motivation to change the self. *Emotion*, *14*, 1049-1061.

Lovett, B. J., Jordan, A. H., & Wiltermuth, S. S. (2012). Individual differences in the moralization of everyday life. *Ethics & Behavior*, *22*, 248-257.

Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, *67*, 371-378.

Newman, G. E., Bloom, P., & Knobe, J. (2014). Value judgments and the true self. *Personality and Social Psychology Bulletin*, *40*, 203-216.

Newman, G. E., De Freitas, J., & Knobe, J. (2015). Beliefs about the true self explain asymmetries based on moral judgment. *Cognitive Science*, *39*, 96-125.

Peterson, C., & Seligman, M. (Eds.). (2004). *Character strengths and virtues: A handbook and classification*. New York, NY: Oxford University Press.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM7: Hierarchical Linear and Nonlinear Modeling*. Chicago, IL: Scientific Software International.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rosnow, R. L., Rosenthal, R., & Rubin, D. B. (2000). Contrasts and correlations in effect-size estimation. *Psychological Science*, *11*, 446-453.

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science*, *20*, 523-528.

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

Schlegel, R. J., Hicks, J. A., Arndt, J., & King, L. A. (2009). Thine own self: True self-concept accessibility and meaning in life. *Journal of Personality and Social Psychology*, *96*, 473-490.

Schlegel, R. J., Hicks, J. A., Davis, W. E., Hirsch, K. A., & Smith, C. M. (2013). The dynamic interplay between perceived true self-knowledge and decision satisfaction. *Journal of Personality and Social Psychology*, *104*, 542-558.

Schlegel, R. J., Hicks, J. A., King, L. A., & Arndt, J. (2011). Feeling like you know who you are: Perceived true self-knowledge and meaning in life. *Personality and Social Psychology Bulletin*, *37*, 745-756.

Schwarz, N., & Strack, F. (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 61-84). New York, NY: Russell-Sage.

Silvia, P. J., & Gendolla, G. H. E. (2001). On introspection and self-perception: Does self-focused attention enable accurate self-knowledge? *Review of General Psychology*, *5*, 241-269.

Snijders, T. A., & Bosker, R. J. (1999). Multilevel analysis: An introduction to basic and advanced multilevel modeling. London, England: SAGE.

Stone, J., Aronson, E., Crain, A. L., Winslow, M. P., & Fried, C. B. (1994). Inducing hypocrisy as a means of encouraging young adults to use condoms. *Personality and Social Psychology Bulletin*, *20*, 116-128.

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

Strohminger, N., & Nichols, S. (2015). Neurodegeneration and identity. *Psychological Science*, *26*, 1468-1479.

Swann, W. B., & Brooks, M. (2012). Why threats trigger compensatory reactions: The need for coherence and quest for self-verification. *Social Cognition*, *30*, 758-777.

Swann, W. B., & Read, S. J. (1981). Self-verification processes: How we sustain our self-conceptions. *Journal of Experimental Social Psychology*, *17*, 351-372.

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, *58*, 345-372.

Tooke, W. S., & Ickes, W. (1988). A measure of adherence to conventional morality. *Journal of Social & Clinical Psychology*, *6*, 310-334.

Tracy, J. L., & Robins, R. W. (2004). Putting the self in self-conscious emotions: A theoretical model. *Psychological Inquiry*, *15*, 103-125.

Tsang, J. (2002). Moral rationalization and the integration of situational factors and psychological processes in immoral behavior. *Review of General Psychology*, *6*, 25-50.

Vess, M., Schlegel, R. J., Hicks, J. A., & Arndt, J. (2014). Guilty, but not ashamed: "True" self-conceptions influence affective responses to personal shortcomings. *Journal of Personality*, *82*, 213-224.

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

Wood, A. M., Linley, P. A., Maltby, J., Baliousis, M., & Joseph, S. (2008). The authentic personality: A theoretical and empirical conceptualization and the development of the authenticity scale. *Journal of Counseling Psychology*, *55*, 385-399.