The Moral Virtue of Authenticity: How Inauthenticity Produces Feelings of Immorality and Impurity

Francesca Gino1, Maryam Kouchaki2, and Adam D. Galinsky3
1Harvard Business School, Harvard University; 2Kellogg School of Management, Northwestern University; and 3Columbia Business School, Columbia University

Abstract
The five experiments reported here demonstrate that authenticity is directly linked to morality. We found that experiencing inauthenticity, compared with authenticity, consistently led participants to feel more immoral and impure. This link from inauthenticity to feeling immoral produced an increased desire among participants to cleanse themselves and to engage in moral compensation by behaving prosocially. We established the role that impurity played in these effects through mediation and moderation. We found that inauthenticity-induced cleansing and compensatory helping were driven by heightened feelings of impurity rather than by the psychological discomfort of dissonance. Similarly, physically cleansing oneself eliminated the relationship between inauthenticity and prosocial compensation. Finally, we obtained additional evidence for discriminant validity: The observed effects on desire for cleansing were not driven by general negative experiences (i.e., failing a test) but were unique to experiences of inauthenticity. Our results establish that authenticity is a moral state—that being true to thine own self is experienced as a form of virtue.

Keywords
authenticity, morality, compensatory ethics, helping, prosocial behavior, open data, open materials

Received 6/5/13; Revision accepted 2/8/15

In a notable passage of Hamlet, Polonius exhorted his departing son, Laertes, to live to the full extent of his humanity: “This above all: to thine own self be true,… Thou canst not then be false to any man” (Shakespeare, 1603/1885, Act 1, Scene iii). Not just the province of a Shakespearean turn of phrase, the desire to be authentic—to act in accordance with one’s own sense of self, emotions, and values—seems to be a driving force of human nature (Gecas, 1986, 1991). Scholars, writers, and philosophers have argued that authenticity is a fundamental aspect of individuals’ well-being (Harter, 2002). A disconnect between one’s expressions and internal states can be psychologically costly, producing palpable discomfort, dissonance, and exhaustion (Ashforth & Tomiuk, 2000; Festinger, 1957; Grandey, 2000). Indeed, some schools of psychotherapy ascribe to Polonius’s belief that psychological health can be achieved only by expressing one’s true inner thoughts and feelings (Rogers, 1961).

Yet it is also the case that people often profess opinions, modulate their emotional expressions, and act in the service of interpersonal relationships and goal-directed behavior (Ekman & Friesen, 1975; Schlenker, 2002). In fact, the more successful a person is at portraying inauthentic experiences or expressions, the more interpersonally competent he or she is judged to be (Snyder, 1987). Indeed, some scholars have argued that the ability to express thoughts and feelings that contradict one’s mental states is an important developmental adaptation (Harter, Marold, Whitesell, & Cobbs, 1996).

In the current research, we attempted to resolve these contradictory claims by exploring whether there is a link
between feeling inauthentic and feeling immoral and impure. We suggest that inauthenticity poses a challenge to a person’s sense of self. Authenticity involves both owning one’s personal experiences (thoughts, emotions, needs, and wants) and acting in accordance with those experiences. A commitment to one’s identity and values (Erickson, 1995) is important for effective self-regulation. When this commitment is violated, people feel inauthentic.

Though being untrue to oneself is psychologically costly, by definition it does not constitute immoral behavior. Yet, we argue, people do experience inauthenticity as immoral, feeling that it taints their moral self-concept. Our arguments build on the writings of the numerous philosophers—such as Kierkegaard, Nietzsche, Rand, and Sartre—who have discussed authenticity in relation to morality. For instance, Nietzsche and Sartre believed that individuals need to create their own moral code and act in ways consistent with that code (i.e., they should act authentically).

By contrast, morality is commonly defined in social and interpersonal terms (Haidt & Kesebir, 2010). For example, Turiel (1983) defined morality as “prescriptive judgments of justice, rights, and welfare pertaining to how people ought to relate to each other” (p. 3). Philosophers and psychologists alike have treated being untrue to oneself (inauthenticity) differently from being untrue to others (dishonesty), and have suggested that society tolerates or promotes inauthenticity but universally prohibits dishonesty (Harter et al., 1996).

We, however, suggest that inauthenticity and dishonesty share a similar root: They are both a violation of being true, whether to others or oneself. As a result, they elicit similar psychological and behavioral responses. For instance, expressing excitement for an activity or person one does not like or trying to fit in with a group that does not share one’s values is not defined as immoral behavior per se, but we argue that individuals experience those behaviors as immoral. Feeling as if one is an imposter to oneself produces moral distress and feelings of being morally tainted and impure that are similar to those that accompany dishonesty.

Previous studies have shown that moral threats activate the need to cleanse oneself (Lee & Schwarz, 2010a; Zhong & Liljenquist, 2006). Similarly, the sacred-value-protection model (see Tetlock, Kristel, Elson, Green, & Lerner, 2000) suggests that when people violate their own values, they engage in symbolic or literal moral cleansing to purify their contaminated conscience and reaffirm their core values. Building on this research, we suggest that experiencing inauthenticity results in lower moral self-regard and feelings of impurity, which trigger a desire for physical cleansing and acting prosocially to compensate for violating the true self (Fig. 1). We also argue that cleansing breaks the link between inauthenticity and prosocial compensation.

Our hypotheses differ from cognitive dissonance theory and its variants in two ways. First, building on the sacred-value-protection model, we suggest that the mere contemplation of acting inauthentically is sufficient to produce feelings of moral contamination. It is the inauthenticity and impurity experienced in these situations, and not the inconsistency itself, that lead to the desire to cleanse and morally compensate. Second, dissonance processes are often triggered not by mere inconsistency but rather by aversive consequences (Cooper & Fazio, 1984); what provokes dissonance is the knowledge that one’s actions have produced material consequences that violate one’s attitudes.

---

**Fig. 1.** Theoretical model for the link between inauthenticity and moral cleansing. Inauthenticity leads to two main consequences of a threatened moral self-concept—feelings of impurity and lower self-regard—as well as dissonance. However, only a threatened moral self-concept explains the link between experiencing inauthenticity and a heightened desire to cleanse oneself and behave prosocially.
Finally, the research we report here is related to the work by Lee and Schwarz (2010b) showing that the physical act of washing reduces cognitive dissonance by creating a clean slate. However, their research did not examine whether experiencing dissonance increases the desire for physical cleansing, whereas we theorized about and empirically tested the link between inauthenticity and cleansing. Specifically, we directly examined the need for cleansing as a result of feeling morally tainted by experiencing inauthenticity.

**Overview of the Present Research**

We tested our predictions in five studies in which people recalled and wrote about a time when they felt authentic or inauthentic. We measured whether inauthenticity influenced people's moral self-regard and feelings of impurity (Experiments 1 and 3) and their desire to cleanse themselves (Experiments 2, 4, and 5). We also linked inauthenticity to prosocial behavior in the form of helping (Experiment 3) and donating money (Experiment 5). To establish discriminant validity, we compared the effects of inauthenticity with the effects of recalling a morally irrelevant, negative experience (i.e., failing a test) in Experiment 3 and with the effects of cognitive dissonance in Experiment 4.

**Experiment 1: The Impurity of Inauthenticity**

Experiment 1 examined whether inauthenticity produces feelings of immorality and impurity, independently of whether it involves being untrue to others or untrue only to oneself.

**Method**

**Participants and design.** Two hundred sixty-nine individuals (mean age = 30.73 years, SD = 8.07; 143 male) from Amazon Mechanical Turk participated in this study for $1. We calculated our target sample size using an estimated effect size, $f$, of 0.2, which would require a sample size of approximately 270 participants for the study to be powered at 90%. We randomly assigned participants to a 2 (type of behavior: authentic vs. inauthentic) × 2 (type of event: general vs. unrelated to lying) between-subjects design. Two participants did not write an essay and were excluded from the analyses, according to a decision made prior to conducting the study.

**Procedure.** Participants first read initial instructions welcoming them to the study and answered an attention check. Those who failed the attention check were automatically informed that, on the basis of their answers, they did not qualify for the study. Thus, their data were not recorded. Participants were then asked to recall an event and write about it for 5 to 10 min. In the authentic-behavior, general-event condition, the instructions read as follows (word changes in the inauthentic-behavior, general-event condition are shown in brackets):

Please recall a time in your personal or professional life when you behaved in a way that made you feel true [untrue] to yourself, that made you feel authentic [inauthentic]. It should just be a situation in which you felt authentic [inauthentic] with your core self. Please describe the details about this situation that made you feel authentic [inauthentic]. What was it like to be in this situation? What thoughts and feelings did you experience?

In the authentic-behavior, event-unrelated-to-lying condition, the instructions read as follows (word changes in the inauthentic-behavior, event-unrelated-to-lying condition are shown in brackets; boldface is used here for emphasis but was not used in the original instructions):

Please recall a time in your personal or professional life when you behaved in a way that made you feel true [untrue] to yourself, that made you feel authentic [inauthentic]. It is important that you choose a situation that is unrelated to telling the truth to others [unrelated to lying or deceiving others]. It should just be a situation in which you felt authentic [inauthentic] with your core self. Please describe the details about this situation that made you feel authentic [inauthentic]. What was it like to be in this situation? What thoughts and feelings did you experience?

Next, participants completed measures assessing their moral self-regard and feelings of impurity. The order in which these two sets of questions were presented was randomly determined for each participant. Participants then completed manipulation checks and reported their age and gender.

**Moral self-regard.** Participants indicated the extent to which the event they described made them feel moral, generous, cooperative, helpful, loyal to others, dependable, trustworthy, reliable, caring, and respectful ($\alpha = .965$; adapted from Walker & Hennig, 2004). Responses were on a 7-point scale (ranging from 1, *not at all*, to 7, *to a great extent*).

**Feelings of impurity.** Using the same 7-point scale, participants indicated the extent to which the event they described made them feel impure, dirty, and tainted ($\alpha = .94$).
Manipulation check: self-alienation. As a manipulation check, we measured feelings of self-alienation with four items (e.g., “After experiencing the situation I described I felt out of touch with the ‘real me,’” “After experiencing the situation I described I felt as if I did not know myself very well”; $\alpha = .88$) that have been used in prior work to measure inauthenticity (Gino, Norton, & Ariely, 2010). We asked participants to indicate their agreement with each of the four items using a 7-point scale (from 1, strongly disagree, to 7, strongly agree).

**Manipulation check: content of the essay.** As an additional manipulation check, we asked participants to think back to the initial writing task and indicate whether they had written about an event that made them feel authentic, inauthentic, or neutral.

**Manipulation check: self-alienation.** A 2 (type of behavior: authentic vs. inauthentic) × 2 (type of event: general vs. unrelated to lying) between-subjects analysis of variance (ANOVA) using self-alienation as the dependent measure revealed only a main effect of type of behavior. Participants in the inauthentic-behavior condition reported greater self-alienation ($M = 4.04$, $SD = 1.37$, 95% confidence interval, CI = [3.82, 4.26]) compared with participants in the authentic-behavior condition ($M = 1.90$, $SD = 1.19$, 95% CI = [1.70, 2.12]), $F(1, 263) = 186.16$, $p < .001$, $\eta_p^2 = .41$.

**Impurity and moral self-regard.** Similar 2 × 2 ANOVAs using impurity and moral self-regard as dependent measures also revealed only a significant main effect of type of behavior. Participants in the inauthentic-behavior condition reported greater feelings of impurity ($M = 3.56$, $SD = 1.36$, 95% CI = [3.30, 3.85]) and lower moral self-regard ($M = 2.90$, $SD = 1.50$, 95% CI = [2.61, 3.16]) than did participants in the authentic-behavior condition (impurity: $M = 1.51$, $SD = 1.29$, 95% CI = [1.25, 1.78]; moral self-regard: $M = 4.99$, $SD = 1.68$, 95% CI = [4.72, 5.26]), $F(1, 263) = 115.25$, $p < .001$, $\eta_p^2 = .31$, respectively.

### Results

**Coding of the essays.** Two coders, who were blind to conditions and hypotheses, categorized the situations participants described in their essays. The two coders agreed on the categorization 94% of the time, and disagreements were resolved with a third coder. As Table 1 shows, about 90% of the essays described situations unrelated to ethics. Most were situations in which people expressed emotions, attitudes, or opinions that did not match their internal state or attempted to fit in by conforming to social norms or peer attitudes.

<table>
<thead>
<tr>
<th>Category</th>
<th>Inauthentic-behavior condition</th>
<th>Authentic-behavior condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Expressing emotions, attitudes, or opinions that do not match one’s internal state</td>
<td>39.1% (46.7%)</td>
<td>35.8% (31.0%)</td>
</tr>
<tr>
<td>2. Attempting to fit in by conforming to norms or shared attitudes and behaviors, or in the face of social pressure</td>
<td>53.6% (30.0%)</td>
<td>32.8% (36.6%)</td>
</tr>
<tr>
<td>3. Lying to obtain a material self-interested advantage</td>
<td>0.0% (13.3%)</td>
<td>0.0% (1.4%)</td>
</tr>
<tr>
<td>4. Theft, stealing</td>
<td>0.0% (5.0%)</td>
<td>17.9% (21.1%)</td>
</tr>
<tr>
<td>5. Cheating in a relationship</td>
<td>0.0% (0.0%)</td>
<td>0.0% (1.4%)</td>
</tr>
<tr>
<td>6. Not being able to create something for oneself</td>
<td>0.0% (1.7%)</td>
<td>6.0% (4.2%)</td>
</tr>
<tr>
<td>7. General*</td>
<td>7.2% (3.3%)</td>
<td>7.5% (4.2%)</td>
</tr>
</tbody>
</table>

*Essays in this category were mainly descriptions of general feelings resulting from the experience.

**Table 1. Distribution of Event Descriptions in Experiment 1 by Content Category**
Next, participants completed measures assessing accessibility of cleansing-related words, desire to use cleansing-related products (e.g., Tide detergent), and desire to cleanse through behaviors such as taking a shower. The order in which these three sets of measures were presented was randomly determined. Participants then completed manipulation checks and reported their age and gender.

**Accessibility of cleansing-related words.** Participants completed a word-completion task using the first word that came to mind (Zhong & Liljenquist, 2006). The instructions read,

> You will now be presented with a word completion task. You will be given a list of words with letters missing. Your task is to fill in the blanks to make complete words. Please use the first word that comes to mind.

Three of the word segments (W_ _H, SH_ _ER, and S_ _P) could be completed as cleansing-related words (*wash*, *shower*, and *soap*) or as unrelated, neutral words (*e.g.*, *wish*, *shaker*, and *step*). The remaining three word segments (F_ O _ , B_ _ K, and PA_ _ R) could be completed with neutral words only.

**Cleansing products.** Participants indicated how desirable they found a list of products to be (using a 7-point scale, ranging from 1, *completely undesirable*, to 7, *completely desirable*). The list included five cleansing products (i.e., Dove shower soap, Crest toothpaste, Windex cleaner, Tide detergent, and Lysol disinfectant) and five neutral products (i.e., Post-it Notes, Nantucket Nectars juice, Energizer batteries, Sony CD cases, and Snickers bars). We averaged responses to the five cleansing products to create one aggregate measure ($\alpha = .86$).

**Cleansing behaviors.** Participants indicated the desirability of various behaviors on a 7-point scale (ranging from 1, *completely undesirable*, to 7, *completely desirable*). Some of the behaviors were related to cleansing (taking a shower, washing hands, brushing teeth, and taking a bath), and others were not (taking a walk, having something to eat, watching TV, and listening to music). We averaged responses to the four cleansing behaviors to create one aggregate measure ($\alpha = .75$).

**Manipulation checks.** As a manipulation check, we measured self-alienation using the same four-item measure as in Experiment 1 ($\alpha = .87$). We also asked participants to think back to the initial writing task and indicate the type of essay they wrote, that is, whether they wrote...
about an event that made them feel authentic, inauthentic, or neutral.

**Results**

**Manipulation check: self-alienation.** A 2 (type of behavior: authentic vs. inauthentic) × 2 (type of event: general vs. unrelated to lying) between-subjects ANOVA using self-alienation as the dependent measure revealed only a main effect of type of behavior. Participants in the inauthentic-behavior condition reported greater self-alienation ($M = 4.07, SD = 1.41, 95\% CI = [3.95, 4.19]$) than did participants in the authentic-behavior condition ($M = 1.87, SD = 1.07, 95\% CI = [1.75, 1.99]$), $F(1, 834) = 655.80, p < .001, \eta^2_p = .44$.

**Accessibility of cleansing-related words.** A similar 2 × 2 ANOVA using the sum of cleansing-related words participants generated as the dependent measure revealed only a main effect of type of behavior (authentic vs. inauthentic). Participants who recalled and wrote about an inauthentic behavior ($M = 1.52, SD = 0.99, 95\% CI = [1.23, 1.42]$) generated more cleansing-related words than did those who recalled and wrote about an authentic behavior ($M = 1.11, SD = 0.93, 95\% CI = [1.02, 1.20]$), $F(1, 834) = 10.02, p = .002, \eta^2_p = .012$.

**Desirability of cleansing products.** Similarly, a 2 × 2 ANOVA using participants’ desirability ratings of cleansing products as the dependent measure revealed only a main effect of type of behavior (authentic vs. inauthentic). Recalling an inauthentic rather than an authentic behavior led to greater desirability of cleansing products ($M = 3.47, SD = 1.48, 95\% CI = [3.33, 3.61]$), vs. $M = 3.11, SD = 1.39, 95\% CI = [2.97, 3.24]$), $F(1, 834) = 13.03, p < .001, \eta^2_p = .015$, but the desirability of noncleansing products did not differ between the inauthentic-behavior condition ($M = 3.08, SD = 1.21, 95\% CI = [2.96, 3.20]$) and the authentic-behavior condition ($M = 3.09, SD = 1.18, 95\% CI = [2.98, 3.21]$), $F < 1$. The effect of inauthenticity on the desirability of cleansing products but not noncleansing ones was confirmed by a significant interaction between type of behavior and type of product (i.e., cleansing related or neutral), $F(1, 834) = 23.94, p < .001, \eta^2_p = .028$.

**Desirability of cleansing behaviors.** Similarly, recalling an inauthentic experience increased the desirability of cleansing behaviors ($M = 4.36, SD = 1.37, 95\% CI = [4.22, 4.50]$), vs. $M = 4.04, SD = 1.46, 95\% CI = [3.91, 4.18]$), $F(1, 834) = 10.19, p = .001, \eta^2_p = .012$, but the desirability of noncleansing behaviors did not differ between the inauthentic-behavior condition ($M = 4.77, SD = 1.26, 95\% CI = [4.65, 4.89]$) and the authentic-behavior condition ($M = 4.70, SD = 1.19, 95\% CI = [4.58, 4.82]$), $F < 1$. The effect of inauthenticity on the desirability of cleansing behaviors but not noncleansing ones was confirmed by a significant interaction between type of behavior in the writing task (authentic vs. inauthentic) and type of behavior in the rating task (i.e., cleansing related vs. neutral), $F(1, 834) = 7.92, p = .005, \eta^2_p = .009$.

**Discussion**

Recalling and writing about an inauthentic experience enhanced a desire for physical cleanliness as measured both implicitly and explicitly. Thus, experiencing inauthenticity heightens the desire to cleanse oneself.

**Experiment 3: Prosocial Compensation and Discriminant Validity**

One concern with the previous experiments is the possibility that the results were driven by recalling a negative, or uncomfortable, event. In Experiment 3, we compared effects of inauthenticity and effects of a morally irrelevant negative experience—failing a test—to test whether the observed link between inauthentic behavior and moral cleansing generalizes to any negative experience. By so doing, we tested for discriminant validity and furthered our understanding of the triggers of moral cleansing. We also tested whether inauthenticity produces moral compensation, leading people to act prosocially, and whether feelings of impurity but not dissonance mediate this effect.

**Method**

**Participants and design.** Two hundred ninety-one individuals (mean age = 30.06 years, $SD = 7.87$; 47% male) from local universities in the northeastern United States participated in this study for pay. We calculated our target sample size using an estimated effect size, $f$, of 0.2, which would require a sample size of approximately 280 participants for the study to be powered at 85%. At some of the experimental sessions, however, participants showed up at a higher rate than expected. Experiment 3 was the first in an hour-long series of experiments for which participants received $20 as compensation. Participants were randomly assigned to one of three conditions: inauthenticity, failure, or control. Three participants failed the manipulation check asking them to indicate the type of essay they wrote and were thus excluded from the analyses, according to a decision made prior to conducting the study. We conducted analyses on the remaining 288 participants.

**Procedure.** Participants first read some general instructions welcoming them to the study, answered one
attention-check question, and then, if they successfully responded to it, moved on to the writing task. In the inauthenticity condition, the instructions read (as in the inauthentic-behavior, general-event condition of Experiments 1 and 2):

Please recall a time in your personal or professional life when you behaved in a way that made you feel untrue to yourself, that made you feel inauthentic. It should just be a situation in which you felt inauthentic with your core self.

Please describe the details about this situation that made you feel inauthentic. What was it like to be in this situation? What thoughts and feelings did you experience?

In the failure condition, we asked participants to describe a time when they failed in an activity, test, or project. The instructions read:

Please recall a time in your personal or professional life when you failed in an activity, test, or project in a way that made you feel disappointed.

Please describe the details about this situation in which you did not succeed on a task. What was it like to be in this situation? What thoughts and feelings did you experience?

Finally, in the control condition, we asked participants to describe their activities from the previous day. The instructions read:

Please recall what happened yesterday, throughout the day.

Please describe the details about this situation. What was it like to be in this situation? What thoughts and feelings did you experience?

After the writing task, participants completed a questionnaire with a few measures of interest (i.e., feelings of impurity, psychological discomfort, negative and positive affect, and embarrassment), two manipulation-check questions, and demographic questions (age and gender). They then indicated their willingness to help the experimenter with another survey that would take 15 min of their time.

Feelings of impurity. As in Experiment 1, participants used a 7-point scale to indicate the extent to which the event they described made them feel impure, dirty, and tainted (α = .94).

Cognitive dissonance. To assess cognitive dissonance, we used a measure developed by Elliot and Devine (1994) that includes psychological discomfort, negative and positive affect, and also embarrassment. In their work, Elliot and Devine found that psychological discomfort was the distinct affective consequence of engaging in counterattitudinal behavior. For completeness, however, we included all the original items. All items were rated on 7-point scales. Psychological discomfort was assessed through three items: Participants rated how uncomfortable, uneasy, and bothered they felt (α = .94). Negative affect was assessed with three items: “angry toward myself,” “disgusted with myself,” and “annoyed with myself” (α = .93). Three items measured positive affect (“happy,” “good,” and “energetic”; α = .95), and two items measured embarrassment (“embarrassed” and “ashamed”; α = .90).

Manipulation Check 1: self-alienation. As a manipulation check, we measured feelings of self-alienation as in Experiments 1 and 2 (α = .90).

Manipulation Check 2: content of the essay. As an additional manipulation check, we asked participants to think back to the initial writing task and indicate whether they wrote about an event that made them feel inauthentic, what they did the day before, or a time when they did not succeed.

Helping. At the conclusion of the experiment, participants were told that the “research team is interested in understanding how people make choices across various domains (health care, work, food purchases). We have prepared a 15-minute survey. We would love your help. If you can help us out, please click yes below and you will be redirected to the survey. Otherwise, please press No. Note that you will receive no extra payment for completing it.” If participants decided to help, they received a message thanking them for choosing to help the research team and then were asked to answer a short questionnaire with general bogus questions.

Results

Table 2 reports the means and confidence intervals for the variables in this study, separately for each condition.

Manipulation check: self-alienation. A one-way ANOVA using self-alienation as the dependent measure revealed a main effect of condition, F(2, 285) = 43.23, p < .001, η_p^2 = .23. Pairwise comparisons (with Bonferroni adjustment) revealed that participants reported greater self-alienation when they recalled and wrote about an inauthentic experience (M = 3.83, SD = 1.51) than when
they recalled and wrote about either a failure ($M = 3.21$, $SD = 1.62; p = .012$) or what they had done the previous day ($M = 1.92$, $SD = 1.19; p < .001$). Participants also reported greater self-alienation in the failure than in the control condition ($p < .001$).

**Feelings of impurity.** Feelings of impurity also differed by condition, $F(2, 285) = 72.29, p < .001$, $\eta^2 = .34$. Pairwise comparisons (with Bonferroni adjustment) revealed that participants reported feeling more impure in the inauthenticity condition ($M = 3.66, SD = 1.82$) than in either the failure condition ($M = 2.09, SD = 1.57; p < .001$) or the control condition ($M = 1.21, SD = 0.61; p < .001$). Participants also reported greater feelings of impurity in the failure than in the control condition ($p < .001$).

**Psychological discomfort.** Psychological discomfort, which has been tied to cognitive dissonance, varied across conditions, $F(2, 285) = 82.67, p < .001$, $\eta^2 = .37$. Pairwise comparisons (with Bonferroni adjustment) revealed that participants reported less psychological discomfort in the control condition ($M = 5.11, SD = 1.53; p < .001$) than in either the inauthenticity condition ($M = 4.90, SD = 1.64; p < .001$). Participants felt the same amount of psychological discomfort in the failure and inauthenticity conditions ($p = 1.00$).

**Negative and positive affect, and embarrassment.** Our manipulation also led to differences across conditions in negative affect, $F(2, 285) = 116.76, p < .001$, $\eta^2 = .41$; positive affect, $F(2, 285) = 80.77, p < .001$, $\eta^2 = .45$; and embarrassment, $F(2, 285) = 82.37, p < .001$, $\eta^2 = .36$. As shown in Table 2, participants in the control condition reported lower negative affect, higher positive affect, and lower embarrassment compared with participants in both the failure and the inauthenticity condition (all $ps < .001$), whereas participants in the latter two conditions did not differ on these measures (all $ps > .71$).

### Table 2. Means and 95% Confidence Intervals (in Brackets) for the Variables Assessed in Experiment 3

<table>
<thead>
<tr>
<th>Condition</th>
<th>Inauthenticity</th>
<th>Failure</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feelings of impurity</td>
<td>3.66 [3.37, 3.95]</td>
<td>2.09 [1.81, 2.37]</td>
<td>1.21 [0.93, 1.49]</td>
</tr>
<tr>
<td>Discomfort</td>
<td>5.11 [4.78, 5.45]</td>
<td>4.90 [4.57, 5.23]</td>
<td>2.41 [2.09, 2.73]</td>
</tr>
<tr>
<td>Positive affect</td>
<td>1.99 [1.72, 2.27]</td>
<td>1.84 [1.57, 2.11]</td>
<td>4.46 [4.29, 4.73]</td>
</tr>
<tr>
<td>Embarrassment</td>
<td>4.40 [4.07, 4.74]</td>
<td>4.69 [4.36, 5.01]</td>
<td>1.97 [1.64, 2.29]</td>
</tr>
<tr>
<td>Helping</td>
<td>33.7% [25.3, 42.1]</td>
<td>17.5% [9.4, 25.7]</td>
<td>16.2% [8.1, 24.3]</td>
</tr>
</tbody>
</table>

Note: Within a row, means with different subscripts are significantly different, $p < .05$.

**Moral compensation through helping.** The percentage of participants who decided to help the experimenter varied by condition, $\chi^2(2, N = 288) = 10.35, p = .006$. Participants who recalled and wrote about an inauthentic experience were more likely to help the experimenter (33.7%, 31 of 92 participants) than those in the failure condition (17.5%, 17 of 97 participants), $\chi^2(1, N = 189) = 6.48, p = .011$, and those in the control condition (16.2%, 16 of 99 participants), $\chi^2(1, N = 191) = 6.88, p = .009$.

**Mediation analysis.** Next, we examined whether feelings of impurity or psychological discomfort due to cognitive dissonance explained the link between inauthenticity and greater helping. In the logistic regressions, we included a dummy variable for both the inauthenticity condition and the failure condition, using the control condition as the condition of reference. When feelings of impurity and psychological discomfort were included in the equation (in addition to the dummies for the failure condition and the inauthenticity condition), the effect of inauthenticity on helping was reduced (from $b = −0.97$, $SE = 0.38$, Wald = 7.63, $p = .006$, to $b = 0.37$, $SE = 0.49$, Wald = 0.57, $p = .45$). Feelings of impurity predicted helping ($b = 0.38$, $SE = 0.11$, Wald = 12.25, $p < .001$), but psychological discomfort did not ($b = 0.14$, $SE = 0.11$, Wald = 1.67, $p = .20$). We conducted bootstrap analyses with 10,000 iterations using a macro provided by Preacher and Hayes (2008) for situations involving multiple mediators. The bootstrapped 95% bias-corrected CI around the indirect effect for impurity, [0.38, 1.56], did not contain zero, but the 95% bias-corrected CI around the indirect effect for psychological discomfort did, [−0.20, 1.01].

**Discussion**

Inauthenticity produced greater feelings of impurity and greater moral compensation compared with failing a test. This study demonstrates that the effect of inauthenticity
on moral compensation cannot be attributed to general negative experiences. It also shows that feeling impure, not cognitive dissonance, explains the relationship between inauthenticity and moral compensation through helping.

**Experiment 4: Inauthenticity Is Not Dissonance**

Experiment 3 provided preliminary evidence that inauthenticity is distinct from cognitive dissonance. In Experiment 4, we explored this issue further using a cognitive dissonance paradigm. In a typical dissonance study, participants are asked to write a counterattitudinal essay on a personally relevant topic, and perceived choice is manipulated. In the high-choice condition, participants are persuaded to write a counterattitudinal essay, but the request provides a feeling of choice. In the low-choice condition, participants are instructed to write the counterattitudinal essay, which gives them little choice. Dissonance studies show a positive correlation between perceived choice and attitudes toward the counterattitudinal topic (Cooper & Fazio, 1984).

Whereas choice is critical in producing cognitive dissonance, we suggest that choice does not play a role in increasing the desire for cleanliness that is associated with feeling inauthentic. We tested our hypothesis in Experiment 4 by including three conditions: high-choice, proattitudinal; low-choice, counterattitudinal; and high-choice, proattitudinal. We predicted that participants would experience a greater sense of choice in the high-choice condition than in the low-choice condition. But we also predicted that participants would express a greater desire for cleanliness whenever they wrote essays that were not consistent with their internal beliefs, regardless of their perceived level of choice. We expected to observe a greater desire for cleanliness in both the high-choice, counterattitudinal condition and the low-choice, counterattitudinal condition compared with the high-choice, proattitudinal condition.

**Method**

**Participants and design.** Four hundred ninety-one college students (mean age = 20.42 years, SD = 1.90; 43% male) from Harvard University participated in the study in return for a $10 Amazon gift card. Fifty-four additional students started the study, but dropped out after reading the initial instructions and before the manipulation took place; their data were thus not recorded. We calculated our target sample size using an estimated effect size, $f$, of 0.15, which would require a sample size of approximately 490 participants for the study to be powered at 85%. We recruited 550 participants, knowing—from prior experience running online studies with this population—that about 10% to 15% of them likely would not complete the study after reading the initial instructions. We randomly assigned participants to one of three conditions: high-choice, counterattitudinal; low-choice, counterattitudinal; or high-choice, proattitudinal.

**Procedure.** Participants first read initial instructions welcoming them to the study. They were then asked to confirm that they were college students at Harvard. Next, as part of the cognitive dissonance manipulation, we asked participants for their opinion whether or not difficulty ratings should be a part of the Q guide (in which all Harvard courses are rated and reviewed by students who have taken them in the past). This issue was topical and familiar because it was a common topic of debate at the college at the time of the study; most students supported the inclusion of difficulty ratings, and most faculty were against it. Participants indicated whether they were for or against the inclusion of difficulty ratings in the Q guide and reported how strongly they held their opinion (from 1, *not at all*, to 7, *very much so*).

Next, participants were asked for their age, gender, and year in school. They were then told that their first task was to write an essay on a current topic, a task that would take about 5 to 10 min to complete. We manipulated dissonance by giving some participants a choice and other participants no choice regarding whether to write a counterattitudinal essay. All participants were told, “We are interested in the effectiveness of writing on current topics of interest to students.” The rest of the instructions varied by condition.

Instructions in the low-choice, counterattitudinal condition indicated,

> We are randomly assigning people to write either a short essay that indicates they are in favor of including difficulty ratings in the Q guide or a short essay that indicates they are against it. You have been assigned to write a list of arguments in favor of/against [depending on their initial opinion] including difficulty ratings in the Q guide. Therefore, you must argue in support of/against [depending on their initial opinion] including difficulty ratings in the Q guide.

In contrast, the instructions in the high-choice, counterattitudinal condition indicated,

> We are asking people to write a short essay about including difficulty ratings in the Q guide. While we would like to stress the voluntary nature of your decision regarding which side of the issue to write on, we would like you to list arguments in favor of/
Participants in this condition had to check a box to confirm their willingness to write the counterattitudinal essay.

Finally, the instructions in the high-choice, proattitudinal condition were the same as the instructions in the high-choice, counterattitudinal condition except that participants were asked to write about the perspective they supported.

In all three conditions, the last part of the instructions read,

We will be using the essay you write to describe this issue to current undergraduates at Harvard. So it is important that you be as persuasive and convincing as possible to convey the message that difficulty ratings should be included in the Q guide.

Participants in all conditions were instructed to start their essay with the same statement, which appeared at the top of the open box where they wrote their essay: “I believe that Harvard College should [should not] include difficulty ratings in the Q guide because. . . .”

After the writing task, participants received a list of five cleansing products to create one aggregate measure ($\alpha = .84$). Next, participants indicated the extent to which the writing task they had completed earlier made them feel inauthentic. We measured inauthenticity using the measure of self-alienation we employed in Experiments 1, 2, and 3 ($\alpha = .91$).

Finally, we asked participants, “How much choice did you have in writing the essay you wrote?” (1 = none at all, 7 = a lot).

**Results**

Table 3 reports the means and confidence intervals for the variables measured in this study, separately for each condition.

**Manipulation check: self-alienation.** A one-way ANOVA using self-alienation as the dependent measure revealed a main effect of condition, $F(2, 487) = 21.14, p < .001, \eta_p^2 = .08$. Pairwise comparisons (with Bonferroni adjustment) revealed that participants reported lower self-alienation in the proattitudinal condition ($M = 1.88, SD = 0.87$) than in both the high-choice, counterattitudinal condition ($M = 2.50, SD = 1.31; p < .001$) and the low-choice, counterattitudinal condition ($M = 2.70, SD = 1.40; p < .001$). Participants reported the same perceived self-alienation in the two counterattitudinal conditions ($p = .94$).

**Perceived choice.** A one-way ANOVA using perceived amount of choice as the dependent measure revealed a main effect of condition, $F(2, 487) = 62.35, p < .001, \eta_p^2 = .20$. Pairwise comparisons (with Bonferroni adjustment) revealed that participants reported lower perceived choice in the low-choice, counterattitudinal condition ($M = 2.85, SD = 1.98$) than in the high-choice, counterattitudinal condition ($M = 3.63, SD = 2.16; p = .001$) and in the proattitudinal condition ($M = 5.24, SD = 1.78; p < .001$). Perceived choice was higher in the proattitudinal condition than it was in the high-choice, counterattitudinal condition ($p < .001$).

**Desirability of cleansing products.** A one-way ANOVA using participants’ desirability ratings of cleansing products as the dependent measure revealed a main effect of condition, $F(2, 487) = 8.24, p < .001, \eta_p^2 = .033$. Pairwise comparisons (with Bonferroni adjustment) revealed that participants reported less desire for cleansing products in the proattitudinal condition ($M = 3.72, SD = 1.35$) than in both the high-choice, counterattitudinal condition ($M = 4.18, SD = 1.51; p = .012$) and the low-choice, counterattitudinal condition ($M = 4.34, SD = 1.44; p < .001$).

### Table 3. Means and 95% Confidence Intervals (in Brackets) for the Variables Assessed in Experiment 4

<table>
<thead>
<tr>
<th>Variable</th>
<th>Low-choice, counterattitudinal</th>
<th>High-choice, counterattitudinal</th>
<th>High-choice, proattitudinal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-alienation</td>
<td>2.70 [2.49, 2.91]</td>
<td>2.56 [2.36, 2.77]</td>
<td>1.88 [1.75, 2.02]</td>
</tr>
<tr>
<td>Desirability of neutral products</td>
<td>3.84 [3.65, 4.03]</td>
<td>3.81 [3.61, 4.01]</td>
<td>3.64 [3.46, 3.88]</td>
</tr>
</tbody>
</table>

Note: Within a row, means with different subscripts are significantly different, $p < .05$. 

Finally, the instructions in the high-choice, proattitudinal condition made them feel more inauthentic ($M = 2.70, SD = 1.35; p = .012$) and the low-choice, counterattitudinal condition ($M = 4.18, SD = 1.51; p = .012$) and the low-choice, counterattitudinal condition ($M = 4.34, SD = 1.44; p < .001$).
Desirability ratings of cleansing products did not differ between the latter two conditions ($p = .94$). There were no differences across conditions in desirability ratings of the noncleansing products, $F(2, 487) = 1.21, p = .30, \eta^2_p = .005$.

**Discussion**

Whereas choice is a critical ingredient in producing cognitive dissonance, it played no role in increasing the desire for cleanliness. When participants wrote essays that were not consistent with their internal beliefs, regardless of choice, they showed a greater desire for cleanliness.

**Experiment 5: Reducing Prosocial Compensation Through Cleansing**

We have demonstrated that inauthenticity makes people feel morally tainted and leads to a greater desire for cleanliness. In Experiment 5, we used moderation to test whether the relationship between inauthenticity and prosocial compensation is explained through a greater desire for cleansing. We manipulated the opportunity to cleanse to examine whether having this opportunity eliminated the link between inauthenticity and helping.

**Method**

**Participants and design.** Two hundred ninety-one individuals (mean age = 22.38 years, $SD = 2.99$; 45% male) from local universities in the northeastern United States participated in this study for pay ($20). We calculated our target sample size using an estimated effect size, $f$, of 0.2, which would require a sample size of approximately 310 participants for the study to be powered at 85%, but the rate at which participants showed up for some of our experimental sessions was lower than expected. We randomly assigned participants to a 2 (behavior recalled: authentic vs. inauthentic) × 2 (opportunity for cleansing: cleansing vs. control) between-subjects design.

**Procedure.** We manipulated authenticity using the same instructions as in the authentic-behavior general-event conditions of Experiments 1 and 2. After completing the writing task, participants were told that the second part of the study consisted of evaluating a product that had been randomly chosen for them. In the cleansing condition, participants were asked to clean their hands carefully with a hand sanitizer placed next to their computer. In the control condition, they were instead asked to place a pen in their hands for a few seconds and examine it carefully. In both conditions, participants were told that they would answer questions about the product later on—which they did, as a filler task.

Following this task, we informed participants that they could donate money to a charity of their choosing. We used willingness to donate money and the amount participants actually donated (from their pay for participating in the experiment) as our main dependent measures.

Next, we asked participants to indicate the extent to which the writing task they had completed earlier made them feel inauthentic. We measured inauthenticity using the measure of self-alienation we employed in our other studies ($\alpha = .88$). Finally, participants reported their age and gender.

**Results**

**Manipulation check: self-alienation.** As expected, participants reported feeling more self-alienated in the inauthentic-behavior condition ($M = 3.12, SD = 1.42, 95\% CI = [2.89, 3.35]$) than in the authentic-behavior condition ($M = 2.36, SD = 1.25, 95\% CI = [2.15, 2.57]$), $F(1, 287) = 22.82, p < .001, \eta^2_p = .074$.

**Likelihood of donating.** We examined whether having the opportunity to cleanse would moderate the effect of inauthenticity on donations. There was a marginally significant interaction between the type of behavior recalled and opportunity for cleansing in predicting the likelihood of donating, $b = 1.65, SE = 0.93, Wald(1) = 3.16, p = .076$. As depicted in Figure 2, participants in the inauthentic-behavior condition were more likely to donate when they did not clean their hands (25.3%, 95% CI = [16, 35]) than when they did (4.5%, 95% CI = [0, 10]), $\chi^2(1, N = 149) = 11.72, p = .001, Cramér’s V = .28$.

Participants who recalled and wrote about an authentic behavior decided to donate about as often whether they cleaned their hands (6.0%, 95% CI = [0, 12]) or did not (6.0%, 95% CI = [0, 12]) or did not (8.0%, 95% CI = [2, 14]; see Fig. 2), $\chi^2(1, N = 142) = .02, p = .84, Cramér’s V = .04$. Thus, increased helping was observed in the inauthentic-behavior condition only among those participants who were not given an opportunity to cleanse themselves. Our results suggest that the act of cleaning their hands assuaged participants’ feelings of impurity from acting inauthentically and reduced their motivation to compensate for these feelings by acting prosocially.

**Amount donated.** The results for the amount of money participants actually donated mirrored the results for the likelihood of donating. There was a significant interaction between the type of behavior recalled and opportunity for cleansing in predicting the amount donated, $F(1, 287) = 6.17, p = .014, \eta^2_p = .021$. Participants in the inauthentic-behavior condition donated a larger amount of money when they did not clean their hands than when...
they did ($M = $1.33, $SD = $2.76, 95% CI = [$0.72, $1.93]), vs. $M = $0.24, $SD = $1.37, 95% CI = [−$0.09, $0.58]), $F(1, 287) = 12.09, p = .001$. But when participants recalled and wrote about an authentic behavior, they tended to donate the same amount of money whether they cleaned their hands with the hand sanitizer ($M = $0.42, $SD = $1.84, 95% CI = [−$0.03, $0.87]) or they did not ($M = $0.35, $SD = $1.42, 95% CI = [$0.02, $0.67]), $F(1, 287) < 1, p = .77$.

**Discussion**

Experiment 5 further established that the relationship between inauthenticity and moral compensation is explained through cleansing behavior. When participants had the opportunity to cleanse themselves, the relationship between inauthenticity and prosocial behavior was eliminated.

**General Discussion**

People often act inauthentically, in various ways, from arguing for a cause they do not believe in to expressing affection toward someone they truly dislike. Our five experiments establish that authenticity is linked to a moral state. When participants recalled a time that they behaved inauthentically, rather than authentically, they felt more impure and less moral, and experienced a greater desire for physical cleanliness. This heightened desire, in turn, made them more likely to behave prosocially to compensate for their feelings of impurity. We established the role of cleanliness as the link between inauthenticity and moral compensation through both mediation and moderation. Our results for feelings of impurity, the desire to cleanse, and prosocial behavior cannot be attributed to negative experiences more generally (e.g., failing a test), but rather must be attributed to inauthenticity. Our findings provide the first empirical evidence of discriminant validity in the literature on moral cleansing and moral compensation. We also found that the effects of inauthenticity were not reducible to cognitive dissonance or driven by psychological distress.

Our research contributes to the literature on moral psychology and behavioral ethics. Past research has found that morality is malleable and dynamic, that situational and social pressure can lead moral people to act dishonestly (Monin & Jordan, 2009). It is commonly assumed that unethical behavior involves people violating a norm shared by others and that this violation produces negative feelings. We have shown that violating internal norms can lead to very similar consequences. When people behave in ways that are inconsistent with their own sense of self, they feel morally tainted and engage in behaviors to compensate for these feelings.

Our results also contribute to the literature examining compensatory behaviors that follow threats, and aversive states that accompany threats. Proulx and Inzlicht's (2012; see also Proulx, Inzlicht, & Harmon-Jones, 2012) meaning-maintenance model integrates various social-psychological theories about compensatory behaviors following threats and expectancy violations. Our results are consistent with this model: Inauthenticity serves as a threat and leads people to experience a greater desire for cleanliness, to compensate for the aversive experience that made them feel immoral and impure.

Although we have demonstrated that inauthenticity is not reducible to dissonance, we have not established that inauthenticity is distinct from other inconsistency-related threats (e.g., ambivalence, self-uncertainty). It is possible that the dissonance participants experienced in the low-choice condition of Experiment 4 resulted from a more general sense of ambivalence, inconsistency, or self-uncertainty (e.g., van Harreveld, Schneider, Nohlen, & van der Pligt, 2012). Future research should establish the unique characteristics that differentiate inauthenticity from these other inconsistency-related threats. We expect that ambivalence or self-uncertainty would not increase feelings of impurity or desire for cleanliness but would lead to compensation through other pathways.

From Shakespeare to Sartre to Rand, writers and philosophers alike have suggested that authenticity is a moral state. Our research provides the first empirical demonstration that there is indeed a link between authenticity and morality. Our results suggest why laughing at the jokes of detested colleagues or dancing when one feels blue makes one run for the showers and behave more prosocially.
Author Contributions

All authors developed the study concept and contributed to the study design. Testing and data collection were performed by F. Gino and M. Kouchaki. F. Gino and M. Kouchaki drafted the manuscript, and A. D. Galinsky provided critical revisions. All authors approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Open Practices

All data and materials have been made publicly available via the Harvard Dataverse Network and can be accessed at https://osf.io/sd76g/. The complete Open Practices Disclosure for this article can be found at http://pss.sagepub.com/content/buy/supplemental-data. This article has received badges for Open Data and Open Materials. More information about the Open Practices badges can be found at https://osf.io/tvyxz/wiki/view/ and http://pss.sagepub.com/content/25/1/3.full.

Note

1. We used a high level of power for the first study we conducted and then adjusted power levels as we conducted more studies.

References


