Good, but Not a Goody Two-Shoes: Self-Presentation Concerns of Appearing Overly Moral

A thesis presented to

the faculty of

the College of Arts and Sciences of Ohio University

In partial fulfillment

of the requirements for the degree

Master of Science

Colleen M. Cowgill

August 2017

© 2017 Colleen M. Cowgill. All Rights Reserved.
This thesis titled
Good, but Not a Goody-Two Shoes: Self-Presentation Concerns of Appearing Overall
Moral

by
COLLEEN M. COWGILL

has been approved for
the Department of Psychology
and the College of Arts and Sciences by

Kimberly M. Rios
Associate Professor of Psychology

Robert Frank
Dean, College of Arts and Sciences
Abstract

COWGILL, COLLEEN M., M.S., August 2017, Psychology

Good, but Not a Goody Two-Shoes: Self-Presentation Concerns of Appearing Overly Moral

Director of Thesis: Kimberly M. Rios

People are generally motivated to appear moral to others in order to gain trust and be liked. However, there may be conditions under which people may be motivated to appear less moral to others in order to be liked. Given previous research that people often tend to derogate and dislike "do-gooders" and "moral rebels," we hypothesized that people will be motivated to downplay their level of morality in interpersonal interactions after privately receiving feedback that they are far more moral than their peers. Furthermore, we predicted these effects would occur in the realm of morality, but not the realm of intelligence. Hypotheses were partially supported by the results of four studies. Although studies provided evidence that people prefer to be seen as intellectually superior to their peers rather than morally superior on affective measures, other studies provided no behavioral evidence that fear of being seen as a “goody-two-shoes” leads people to downplay their moral behavior.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>3</td>
</tr>
<tr>
<td>List of Figures</td>
<td>5</td>
</tr>
<tr>
<td>Introduction</td>
<td>6</td>
</tr>
<tr>
<td>Social Comparison Theory and Do-Gooder Derogation</td>
<td>7</td>
</tr>
<tr>
<td>Optimal Distinctiveness</td>
<td>9</td>
</tr>
<tr>
<td>Moral Licensing</td>
<td>11</td>
</tr>
<tr>
<td>Morality Versus Other Positive Traits</td>
<td>12</td>
</tr>
<tr>
<td>The Current Research</td>
<td>13</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>14</td>
</tr>
<tr>
<td>Pilot Study</td>
<td>16</td>
</tr>
<tr>
<td>Study 1</td>
<td>19</td>
</tr>
<tr>
<td>Method</td>
<td>19</td>
</tr>
<tr>
<td>Results</td>
<td>22</td>
</tr>
<tr>
<td>Study 2</td>
<td>25</td>
</tr>
<tr>
<td>Method</td>
<td>25</td>
</tr>
<tr>
<td>Results</td>
<td>29</td>
</tr>
<tr>
<td>Study 3</td>
<td>32</td>
</tr>
<tr>
<td>Method</td>
<td>32</td>
</tr>
<tr>
<td>Results</td>
<td>34</td>
</tr>
<tr>
<td>Study 4</td>
<td>37</td>
</tr>
<tr>
<td>Method</td>
<td>37</td>
</tr>
<tr>
<td>Results</td>
<td>38</td>
</tr>
<tr>
<td>General Discussion</td>
<td>40</td>
</tr>
<tr>
<td>References</td>
<td>44</td>
</tr>
<tr>
<td>Appendix A: Dependent Measures</td>
<td>50</td>
</tr>
<tr>
<td>Appendix B: Morality Test</td>
<td>61</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Hypothetical Morality Score Feedback</td>
<td>20</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Average Morality Score Feedback</td>
<td>27</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Above Average Morality Score Feedback</td>
<td>28</td>
</tr>
</tbody>
</table>
Introduction

People tend to think of being moral as an unambiguously positive trait. Indeed, most psychological research to date has focused on how people self-enhance when it comes to their moral characters (Alicke, 1984), and considerable evidence has been found to support this phenomenon. For instance, an analysis of various impression management scales, including Paulus and Reid's (1991) Self-Deceptive Enhancement Scale, showed that moralistic self-enhancement, defined as a "tendency to see oneself as an exceptionally good member of society," is one of two primary biases when it comes to socially desirable responding (Paulus & John, 1998). Further, studies have shown that people unrealistically overestimate their likelihood of engaging in future prosocial actions (Epley & Dunning, 2000; 2006).

The tendency to enhance one's moral character appears to confer clear psychological and social benefits: unrealistically positive beliefs about one's moral character can contribute to the pleasant feelings that come from maintenance of high self-esteem (Baumeister et al., 2003; Pyszczynski et al., 2004), whereas advertising oneself as trustworthy, responsible, and benevolent to others can aid one in forming valuable relationships (Goodwin, Piazza, & Rozin, 2014; Iredale & Vugt, 2011). Other research shows that being perceived as trustworthy can grant advantages to competitors during economic exchanges (Barney & Hansen, 1994). However, could there be potential drawbacks to exaggerating the nobleness of one's character? In other words, might there be conditions under which people would be motivated to downplay their moral characters?
The current research explores the possibility that there exist situational and social motives that encourage people to deemphasize their level of morality. Like a teenager who shirks parentally instilled moral convictions against smoking and swearing in order to fit with a new social group at school or a vegetarian who wants her meat-eating friends to know that she's not passing moral judgment on their dietary habits, signaling that one is not "holier than thou" or a "goody two-shoes" may have important social advantages.

**Social Comparison Theory and Do-Gooder Derogation**

Social comparison theory posits that we strive to obtain somewhat accurate self-evaluations and are often motivated to self-enhance through comparing ourselves with relevant others (Festinger, 1954; Collins, 1996). Within this framework, previous research has demonstrated that upward social comparisons, or comparisons with others who exceed the self on a given characteristic, can threaten self-esteem, prompt envy, and create negative mood states (Gibbons & Gerrard, 1989; Salovey & Rodin, 1984; Wheeler & Miyake, 1992). As a response to these threats to positive mood and anxiety-buffering self-esteem, people may view very high performing sources of upward social comparison less favorably than those whose performance is less superior (Aronson, Willerman, & Floyd, 1966). Furthermore, sources of upward social comparison may face retaliatory punishments in the form of derogation (Cialdini & Richardson, 1980; Tesser & Campbell, 1982). Apart from the derogation of high performing others in competence domains, the derogation of superior others may also extend to the domain of morality.

In line with previous research on upward social comparison, people tend to dislike others they identify as being exceptionally ethical when it comes to moral choices.
A recent study showed that participants derogated "moral rebels" who refused to participate in a racist task that the participants themselves agreed to take part in (Monin, Sawyer, Marquez, 2008). This effect was driven by the imagined or implicit reproach of the "moral rebels" against the complicit participants. In another study, participants who were asked to freely associate words with a target group (vegetarians) selected more words with negative connotations if they expected that target group to identify themselves as morally superior to the participants, who were meat-eaters (Minson & Monin, 2012).

Other research has shown that this phenomenon of derogating potentially morally superior others even appears in young children. Although children rated their peers as more likable if they were generous, they liked generous children significantly less if those children were more generous than themselves - presumably because those children represented a threatening upward social comparison (Tasimi, Dominquez, & Winn, 2015). These do-gooder derogation effects highlight an important potential drawback to being identified as "holier than thou." People tend not to like those they perceive as morally superior to themselves, or those they imagine as deeming themselves morally superior. Additionally, people in danger of being perceived as “holier than thou” may even face social penalties in the form of antisocial punishment.

Antisocial punishment occurs when people who behave prosocially are penalized, either economically or through ostracism, in direct response to their prosocial behavior, even though these cooperative group members benefit the group as a whole (Hermann, Thoni, & Gachter, 2008). A recent study originally developed to study the detection and
expulsion of selfish group members from a cooperative game resulted in the unexpected finding that especially altruistic members were also expelled from the group with a much higher frequency (Parks & Stone, 2010). Some researchers theorize that this seemingly irrational behavior occurs when non-cooperators strategically punish generous cooperators, either preemptively or as retaliation for past punishment (Hermann, Thoni, & Gachter, 2008, Baumard et al., 2013). Others suggest that altruistic members of a society might be punished for violating a "norm of self-interest" (Irwin & Horne, 2012; Miller, 1999). In this conception of antisocial punishment, norm violators are punished regardless of the overall positive impact of the violation on the group. These findings provide evidence of further potential motivation for people to avoid appearing overly moral.

**Optimal Distinctiveness**

Aside from avoiding antisocial punishment and incurring dislike from their peers, people may strive to avoid appearing "overly moral" due to social motivations outlined in Optimal Distinctiveness Theory. Brewer's (1991) Optimal Distinctiveness Theory describes the competing needs people have to both be similar to others and to differentiate themselves from others. Within this model, being extreme on either end of the spectrum of "difference" and "sameness" fails to satisfactorily meet both opposing needs. Thus, despite the advantages of possessing a positive trait, such as intelligence, it may be possible to possess an overabundance of the quality if it threatens one's assimilation or belongingness in one's social group.
Indeed, some people who possess an overabundance of a positive trait may find their position relative to others to be uncomfortable. Research has found that being the target of an upward social comparison can produce anxiety related to fearing producing envy and hostility in others (Exline & Lobel, 1999). A field study by Exline and Lobel (1997) determined that those who imagined that they possessed something others did not have reported fears of being envied, disliked, and rejected by others. Other research has shown that those who outperform others on some measure exhibit more anxiety, modesty and downplay their achievements relative to others (Brigham, 1996; Ishiyama & Chabassol, 1984). Some research even suggests that some people may hinder their own performance on tasks in order to avoid the anxiety and stigma associated with outperformance (Pappo, 1983; Peplua, 1976). These findings are consisted with Optimal Distinctiveness Theory in that outperformance, while affirming and conducive to higher self-regard, can threaten belongingness in that it may garner dislike from others. This burden of outperformance has been conceptualized as the "Fear of Positive Evaluation" by some researchers (Weeks, Heimberg, & Rodebaugh, 2008). Other researchers have described this negative affect as a “fear of envy” that stems from fearing ill will from others as a result of other people coveting the outperformed trait (Rodriquez Mosquera, Parrott, & Hurtado de Mendoza, 2010). However, we expect that these optimal distinctiveness concerns will be stronger for morality than for other traits, such as competence, due to the additional threat of do-gooder derogation.
Moral Licensing

An alternative explanation for why people might be motivated to portray themselves as less moral after receiving feedback that they are far more moral than their peers is the theory of "moral licensing." According to this theory, people may feel more permitted or "licensed" to commit immoral acts when they have recently behaved morally or received feedback validating their moral characters (Merritt, Effron, & Monin, 2010; Blanken, Ven, & Zeelenberg, 2015). Research has shown that people who affirmed their non-prejudiced attitudes by expressing their disagreement with sexist statements on a survey were subsequently more likely to demonstrate prejudice by preferring a male candidate for a stereotypically masculine job (Monin & Miller, 2001). In another study, participants who affirmed their moral identities by writing about themselves using positive trait words opted to donate less to charity than those who wrote about themselves using negative trait words or those who wrote about someone they knew (Sachdeva, Iliev, & Medin, 2009). Thus, it could be argued that people who receive affirming feedback about their moral identities may feel "licensed" to report a higher frequency of past deviant behaviors and a higher level of endorsement of such behaviors.

We do not expect moral licensing explanations to account for our findings. Moral licensing makes predictions about future immoral acts, not self-reported past deviant behavior. There is no apparent rationale for why those who feel "licensed" to behave less morally by past altruistic behavior would also be motivated to represent their past, current, or typical behavior as less moral. Indeed, part of the mechanism driving moral
licensing effects requires that people continue to see their past behavior as exceptionally moral in order for them to feel “licensed” to behave immorally in the future.

**Morality Versus Other Positive Traits**

Although people may be averse to learning that they possess higher degrees of other traits than average, such as intelligence, we expect our effects to be especially strong when people receive feedback that they have far higher levels of *morality* than their peers. According to Reeder and Brewer (1979), traits that contain ability components, such as intelligence, carry far different implications when it comes to attribution than traits that do not, such as morality. That is, when someone performs an act that requires ability, others can logically make inferences that the person in question possesses the necessary ability. However, one can perform an act that does not require an ability, such as some type of moral act, without necessarily possessing the trait of morality. This basic difference restricts the self-serving bias when it comes to evaluating one's own traits that require ability, because these traits are experienced as more objective and specific. Indeed, recent research has shown that when asked to report on their morally good versus their intelligent behaviors, people exhibit self-serving biases in both dimensions but much stronger biases in the moral domain (Allison, Messick, & Goethals, 1989). That is, people rate themselves highly in competence relative to their peers, but they rate themselves even more highly in morality relative to their peers when it comes to morality.

Furthermore, previous research has shown that morality plays a more important role in the perceptions and judgment of others than competence. Research conducted by
Ybarra, Chan, and Park (2001) showed that participants of all ages were quicker to respond to personality trait terms that pertained to morality than traits related to competence. Other recent research has shown that information about a social group's level of morality was a stronger predictor of people's overall feelings and impressions about that social group than their level of competence or sociability (Brambilla, et. al, 2011). This research indicates that cues pertaining to a person's moral character are more salient than cues that fall within the domain competence. Thus, people may have a greater interest in effectively managing impressions formed about their level of morality rather than their level of competence.

For these reasons, we did not expect people who receive feedback that they possess far more academic aptitude than their peers to subsequently downplay their intelligence in interpersonal interactions. Indeed, failing to find the similar effects when a competence-relevant domain is emphasized rather than a morality-relevant domain would help rule out alternative explanations that any downplaying of one’s moral character that results is solely attributable to optimal distinctiveness concerns.

The Current Research

The purpose of the current studies is to determine whether people attempt to modify their moral self-presentation after receiving feedback that they are far more moral in comparison with their peers. Firstly, we intend to establish that people will experience more negative affect, particularly related to discomfort, in scenarios wherein their superior morality relative to their peers is emphasized as opposed to their superior intellect relative to their peers. We also aim to determine whether people who receive
feedback that they are on the extremely positive end of the morality spectrum, relative to those who receive feedback that their level of morality falls within an average range, will be motivated to report more "deviant" behaviors, such as heavy drinking, casual sex, use of profanity, and truancy, or express endorsement or approval of these same behaviors in others. Furthermore, we aim to determine whether people who receive the "far more moral" feedback will downplay the extent to which they negatively judge others' immoral behavior relative to people who receive the "average morality" feedback. Lastly, we intend to demonstrate that these effects do not occur for other positive traits, such as competence, about which people tend to receive objective feedback on a regularly.

**Hypotheses**

Hypothesis 1: Participants will express more negative affect in real or imagined scenarios where they behave far more morally than their peers, or are described as far more moral than their peers, than participants in real or imagined scenarios where they behave far more competently/intelligently than their peers, or are described as far more competent/intelligent than their peers.

Hypothesis 2: Participants who receive feedback that they are much more moral than their peers will attempt to downplay their perceived level of morality, relative to participants who receive feedback that they are of average levels of morality in comparison with their peers, by reporting more deviant behavior when asked about their moral behavior face-to-face.

Hypothesis 3: Participants who receive feedback that they are much more moral than their peers will report that they see themselves as less judgmental than their peers
relative to participants who receive feedback that they are of average levels of morality compared to their peers.

Hypothesis 4: Similar effects will not be present among those given feedback that they are far more intelligent than their peers. Participants who are given feedback that they are far more intelligent, rather than far more moral, than their peers will not show motivation to downplay their intelligence in subsequent interactions.
Pilot Study

We have argued that, in the context of fear of envy, a combination of factors makes receiving feedback that one is more competent than one’s peers potentially less aversive than receiving feedback that one is more moral than one’s peers. Therefore, it was necessary to establish if there was an initial evidentiary basis for our predictions reflected in the general self-reported feelings of the population. In order to ascertain whether people would imagine experiencing less positive affect and more negative affect if they received feedback that they were far more moral than their peers rather than far more intelligent than their peers, we recruited 55 people to participate in a brief pilot study on Amazon Mechanical Turk in exchange for a small compensation of $0.40. Four participants were deleted from the analysis for failing to complete the study, leaving a final sample of 51 participants (20 females, 31 males; 34 White participants, 4 Black participants, 8 Asian participants, 3 participants who identified as “Other,” and 2 participants who declined to respond.) The mean age for the sample was 29 years old, with a standard deviation of 9.39 years.

The survey consisted of 26 items asking participants to imagine different scenarios wherein they received feedback that they were either far more moral than their peers, far more intelligent than their peers, morally average compared to their peers, or of average intelligence compared to their peers. These items were evaluated on both bipolar and unipolar 7-point scales. Sample items included, “How anxious would you be to receive feedback that you were far more intelligent than your peers?” and “How good
versus bad would you feel to receive feedback that you were far more moral than your peers?” Full measures for all scales can be found in Appendix A.

To evaluate our results, we condensed morality-relevant items and intelligence-relevant items into composite measures and conducted paired samples t-tests pitting intelligence-relevant measures against morality-relevant measures. To get a composite measure of negative affect towards receiving a type of feedback, the measures of discomfort, anxiety, good versus bad feelings (reverse coded), and happiness (reverse coded) were averaged together for each type of feedback (intelligent $\alpha = .69$, moral $\alpha = .64$). Consistent with our predictions, a paired-samples t-tests of these composites revealed that people reported imagining that they would experience significantly more negative affect after receiving feedback that they were more moral than their peers ($M = 3.80$, $SD = 1.06$) rather than after receiving feedback that they were more intelligent than their peers ($M = 2.93$, $SD = 1.06$), $t(50) = 2.04$, $p < .05$.

The same composites were created for measures of receiving feedback that one is of average intelligence compared to their peers and receiving feedback that one is of average morality compared to their peers. Participants reported that they would feel significantly more overall negative affect receiving that they were of average intelligence compared to their peers ($M = 3.94$, $SD = .99$) rather than receiving feedback that they were of average morality compared to their peers ($M = 3.46$, $SD = .78$), $t(50) = 3.38$, $p < .001$. This again suggests that people have an overall preference in receiving feedback that they are better-than-average in the competence domain rather than the morality domain. Indeed, one-sample t-tests of two bipolar measures directly measuring the type
of feedback one would prefer to receive found that (compared to a neutral mean value of 4) participants significantly preferred to receive feedback that they were far more intelligent than their peers rather than far more moral than their peers, $t(50) = 2.59, p < .05$, and also they significantly preferred to receive feedback that they were far less moral than their peers rather than far less intelligent, $t(50) = 2.45, p < .05$.

In light of these findings providing support for our initial predictions, we ran another study on Amazon Mechanical Turk manipulating the type of feedback people received and measuring the effect on our proposed dependent variables of self-reported participation in deviant behaviors.
Study 1

Study 1 sought to establish whether imagining receiving feedback that they were far more moral than their peers would lead participants to report participating in more deviant or morally questionable behaviors, relative to participants who imagined receiving feedback that they were far more intelligent compared to their peers or of average morality compared to their peers. Participants viewed an image of a standard normal curve that either placed them a few standard deviations above the median or right at the median of a distribution of morality scores or intelligence scores prior to responding to our dependent measures.

Method

Participants and design. One-hundred and seventy-eight college students were recruited through Amazon Mechanical Turk. Twenty-three participants were removed from the final analysis for failing to complete all measures, leaving 155 people in the final sample (75 male, 80 female, 23 Black, 15 Asian, 13 Hispanic/Latino, 52 White, one American Indian, two “Other, and 52 declining to respond). The mean age of the sample was 26 years old with a standard deviation of 6.68.

Participants were randomly assigned to imagine themselves receiving three different types of feedback: that they were far more moral than their peers, that they were far more intelligent than their peers, or that they were morally average in comparison with their peers. They were each presented with an image visually representing where they would lie on a distribution of scores for the assigned trait. An example of the feedback image is included below:
After being presented with the image, participants were asked to take a few moments to reflect on what they might think and feel after receiving such feedback and to write a brief hypothetical reaction. Then, they were asked to evaluated on a 7-point scale how surprising versus unsurprising they would find the feedback and how expected versus unexpected they would find the feedback. Averaging these two items together
created a composite measure of unexpectedness with a Cronbach’s alpha value of .77 ($M = 4.1$, $SD = 1.59$).

We then informed participants that they would be asked about some of their habits and behaviors relative to their peers as part of a brief separate study. In this survey, we gave participants bogus “facts” about how frequently fellow college students reported engaging in certain behaviors, such as drinking alcohol, binge drinking, how frequently they studied, and how much effort they put into classes, and then asked how frequently they themselves engaged in the behavior. We also asked participants to rate themselves on 7-point scales on how frequently they engaged in other so-called “deviant” behaviors, such as the use of profanity of social media. A full list of the dependent measures can be found in Appendix A. Furthermore, we asked participants to rate on 7-point scales how negatively they would judge their fellow college students for engaging in a variety of moral violations and deviant behaviors. Finally, we asked participants, “On a scale from 1-7, with 1=not at all judgmental and 7=very judgmental, how judgmental of a person would you say you are relative to your peers?” and “On a scale from 1-7, with 1=not at all accepting and 7=very accepting, how accepting of imperfections in others would you say you are relative to your peers?” At the conclusion of the study, participants were thanked and fully debriefed.

**Results**

We predicted that participants in the “far more moral” condition would report more drinking, greater use of profanity, and greater engagement in behaviors that would meet disapproval from their parents or supervisors relative to their other two conditions.
We also predicted that participants in this condition would report being less likely to volunteer their time or donate to charity. Finally, we predicted that participants in the “far more moral” condition would report judging their peers less negatively for drinking alcohol, using drugs, using profanity, cheating on a test, or disrespecting their instructor, and would also report being less judgmental and more accepting overall.

To test these predictions, we first created more composite measures for our dependent variables, standardizing and averaging similar variables. We averaged the two standardized self-report measures pertaining to drinking ($\alpha = .79$, $M = .01$, $SD = .91$), the four standardized self-report measures related to being a responsible or “model” student ($\alpha = .73$, $M = .01$, $SD = .75$), the two measures concerning altruistic behaviors ($\alpha = .62$, $M = 3.45$, $SD = 1.55$), and the eight measures assessing how negatively participants would judge their peers ($\alpha = .75$, $M = 3.81$, $SD = .99$). Two measures, “How often do you post what might be considered “profanity” on social media?’ and “How often do you engage in behaviors your parents or supervisors might disapprove of?” were analyzed as stand-alone items, because they were uncorrelated with any of the other items.

We then submitted each of these composites to a one-way ANOVA. The strongest significant effect to emerge in these analyses were participants’ evaluations of how unexpected the feedback would be, $F(2, 152) = 8.87$, $p < .001$, $\eta^2 = .11$. Post-tests revealed that there were significant differences between the “far more moral” and “average morality” conditions, such that participants in the “far more moral” condition reported finding the feedback more unexpected ($M = 4.54$, $SD = 1.54$) than the “average morality” feedback ($M = 3.26$, $SD = 1.61$), $t(149) = 4.10$, $p < .001$. There were also
significant differences between the “far more intelligent” and “average morality” conditions, such that those in the “far more intelligent” condition reported finding the feedback more unexpected ($M = 4.26, SD = 1.43$) than the “average morality” feedback ($M = 3.26, SD = 1.61$), $t(149) = 3.19, p < .05$. No significant difference emerged between the “far more intelligent” and “far more moral” condition, $t(149) = .99, p = .33$.

The one-way ANOVA also showed a marginally significant difference in the standardized composite of self-reported drinking behavior based on condition, $F(2, 152) = 2.44, p = .09, \eta^2 = .03$. Consistent with predictions, post-tests showed the largest difference between the “far more moral” and “average morality” condition, such that those in the “far more moral” condition reported higher levels of drinking ($M = .20, SD = .91$) than those in the “average morality” condition ($M = -.20, SD = .91$), $t(149) = 2.13, p = .03$. Post-tests showed a non-significant difference between the “far more intelligent” ($M = -.04, SD = .89$) and the “far more moral” condition ($M = .20, SD = .91$), $t(149) = 1.43, p = .16$, as well as a non-significant difference between the “far more intelligent” and “average morality” condition, $t(149) = .87, p = .39$. There were also trending effects for self-reported use of profanity on social media, $F(2, 152) = 1.66, p = .19, \eta^2 = .02$, and standardized composite measure of being a good student, $F(2, 152) = 1.70, p = .19, \eta^2 = .02$. These trends were also in the predicted direction, such that participants who were asked to imagine receiving feedback that they were “far more moral” reported posting more profanity online ($M = 3.16, SD = 1.88$) than those who imagined receiving feedback that they were “morally average,” ($M = 2.53, SD = 1.80, t(149) = 1.33, p = .07$, and that those in the “far more moral” condition reported being less responsible students ($M = -
.07, $SD = .77$) than those in the “average morality” condition ($M = .20, SD = .73$), $t(149) = 1.33, p = .08$. Analyses of how negatively one would judge one’s peers, $F(2, 152) = .22, p = .80, \eta^2 = .00$, self-reported altruistic behaviors, $F(2, 152) = .30, p = .74, \eta^2 = .00$, and how frequently participants engaged in behaviors their parents or supervisors might disapprove of, $F(2, 152) = .70, p = .50, \eta^2 = .01$, showed no significant effects.

Although these results provided partial support for our hypotheses, we conjectured that the tenuous findings may have occurred because of a lack of believability of the hypothetical scenario. Participants finding the “far more moral” feedback highly unexpected and surprising may have contributed to them discounting this feedback as being unrealistic. Furthermore, our manipulation may not have been salient enough to elicit the predicted results due to its hypothetical nature and the lack of social pressure in the form of potential evaluation from one’s peers. To overcome these limitations, we conducted a far more immersive lab study of undergraduate college students wherein participants took what they thought was an actual test of their morality, received what they believed was real feedback, and then responded to our survey of dependent variables face-to-face with another undergraduate research assistant.
Study 2

We designed Study 2 to be a test of our hypotheses with a higher degree of experimental realism than our previous attempts. To achieve this, we devised an ostensible “morality test” for participants to take prior to receiving bogus feedback about their “morality score” relative to their peers who took the same test. The test consisted of questions about what choices the participant would make in various moral dilemmas, what moral violations participants would commit in exchange for money, and measures of empathy. Participants took the test, received one of two kinds of feedback, and then answered a paper survey face-to-face with another undergraduate student about their habits and behaviors.

Method

Participants and design. Seventy-four undergraduate students from Ohio University were recruited to participate in the study in exchange for course credit. Two participants were eliminated from the final analyses for suspecting that the feedback they received was fake, leaving a total sample 72, including 17 males, 55 females, 51 White participants, and 11 participants of various other racial backgrounds. Due to a computer error with the randomization of conditions, the data for “morally average” condition was collected almost entirely in the fall semester, while the data for the “far more moral” condition was collected entirely in the following semester. This also led to higher number of participants in the “morally average” conditions ($N = 46$) than the “far more moral” condition ($N = 26$). We attempted to address this problem in our analyses by randomizing the condition of the last few participants collected in the latter semester to assure that the
time delay did not introduce undue additional variance, and also by randomly selecting a
group of 26 participants from the “morally average” condition to compare against the “far
more moral” condition to make sure that the results derived from the full sample were not
significantly different from the reduced sample.

**Procedure.** Participants read that they would be asked to complete surveys for
two different studies. First, they completed a brief survey that included their demographic
information, including religious and political affiliation. Participants then took a
computerized test described to them as a “morality test” to learn about how ethical they
are in decision-making.

The morality test consisted of three separate components: a moral choices test, an
empathy test, and an evaluation of which (arguably) immoral acts participants would
perform for money. These test components were all taken from the Moral Sense Test,
developed by Cushman, Young, and Hauser (2006). The length of the test and the
diversity of questions asked on the test was meant to A) obfuscate the test's bogus nature
by making it appear more realistic and believable, and B) to encourage participants to pay
closer attention to the instructions of the experiment. The full “Morality Test”
administered to participants can be found in Appendix B.

After completion of the morality test, participants received one or two forms of
bogus feedback, depending on random assignment. In the "average morality score"
feedback condition, participants received the following feedback:

“Comparing your responses with those of your fellow students shows that your
‘morality score’ falls within the **average range** of Ohio University students. The red line
on the graph below indicates where your score falls among the scores of your fellow Ohio University students.”

Figure 2: Average Morality Score Feedback.

By contrast, in the "above average morality score" feedback condition, participants received the following feedback:

“Comparing your responses with those of your fellow students shows that your ‘morality score’ falls very far above the average of Ohio University students. The red
Upon receiving their bogus morality test feedback and "morality score," participants responded to two measures designed to ascertain how much participants expected or were surprised by their feedback. These measures served as a proxy for how believable participants found the feedback, as in Study 1. After completing the computer-
based portion of the experiment, participants were told they would be asked to complete a brief “student life” survey. An undergraduate research assistant sat face-to-face with participants and asked questions related to how frequently participants engage in behaviors that could be considered “deviant” or “immoral” by some, as well as how negatively they judged their peers for similar behaviors or moral violations. These questions were identical to the dependent measures used in Study 1. Full dependent measures for studies 1 and 2 are located in Appendix A. At the conclusion of this face-to-face paper survey, participants were thanked and fully debriefed.

Results

In order to counteract the threat of being seen as a “goody-two shoes” by the undergraduate experimenter, we again predicted that participants in the “far more moral” condition would report engaging in more supposedly “immoral” or “deviant” behavior than participants in the “morally average” condition. We also predicted that participants in the “far more moral” condition would report judging their peers less negatively for the same behaviors, and would judge themselves as more accepting and less judgment than their peers. To test these predictions, we first created composites of the dependent measures identical to those created in Study 1, such that composites for altruistic behavior ($\alpha = .62$, $M = 4.84$, $SD = 1.55$), standardized drinking behavior ($\alpha = .69$, $M = 0$, $SD = .87$), standardized measures of being a “responsible” or “model” student ($\alpha = .56$, $M = 0$, $SD = .66$), negative judgment of one’s peers ($\alpha = .80$, $M = 3.54$, $SD = .99$), and unexpectedness of the feedback ($\alpha = .80$, $M = 3.55$, $SD = 1.64$) were created. We again analyzed use of profanity on social media and frequency of behaviors that one’s parents
or supervisors might disapprove of as stand-alone items. Independent samples t-tests were then run on each of these dependent measures.

Contrary to our predictions, our manipulation did not have a significant impact on self-reported altruistic behavior, $t(70) = .10, p = .92$, use of profanity, $t(70) = .15, p = .88$, or negative judgments of one’s peers, $t(70) = .13, p = .20$. Trending effects emerged, but in the opposite direction of the predictions. Participants who received feedback that they were “morally average” compared to their peers reported higher frequencies of drinking behaviors ($M = .12, SD = .90$) than those who received feedback that they were “far more moral” than their peers ($M = -.21, SD = .80$), $t(70) = 1.52, p = .13$, and they also scored lower on measures of being a “responsible” or “model” student ($M = -.09, SD = .65$) than those who received feedback of being morally superior to their peers ($M = .16, SD = .66$), $t(70) = 1.53, p = .13$. Furthermore, participants in the “morally average” condition reported higher frequency of engaging in behaviors their parents or supervisors might disapprove of ($M = 3.37, SD = 1.64$) than participants in the “far more moral” condition ($M = 2.81, SD = 1.50$), $t(70) = 1.44, p = .15$. Once again, those in the “far more moral” condition reported finding the feedback much more surprising/unexpected ($M = 4.50, SD = 1.32$) than those in the “average morality” condition ($M = 3.00, SD = 1.56$), $t(70) = 4.12, p < .001$.

The results from Study 2 suggested two major limitations with our experimental design. Firstly, because participants still reported finding their “morality score” feedback much more surprising and unexpected in the “far more moral” condition than the “average morality” condition, we thought perhaps that once again the manipulation may
not have been realistic enough to be believable. Furthermore, it is possible that participants saw the undergraduate research assistant not as a peer in this context, but rather an authority figure. To bypass these issues and replicate our initial pilot study findings, we again turned to evaluating people’s anticipated reactions to hypothetical scenarios.
Study 3

Study 3 was designed to replicate our pilot study findings more extensively. We asked participants to report how they would prefer to be evaluated relative to their peers on various trait dimensions. Then, we asked participants to respond to items on the Moral Identity Scale (Aquino & Reed, 2002), which measures the centrality of being a moral person one’s overall identity. We included this variable with the intention of exploring it as a potential moderator of our hypothesized effects, such that participants rating higher on the Moral Identity Scale would report preferring to be rated as more moral relevant to their peers and would report imagining themselves feeling less negative emotions in a scenario wherein the acted more morally than their peers. Finally, we asked participants to imagine various scenarios that either featured them intellectually outperforming their peers or acting with moral superiority relative to their peers, and then to imagine to what degree they would experience various positive and negative emotions.

Method

Participants and design. One hundred and three participants were recruited to complete the survey from Amazon Mechanical Turk in exchange for $0.40. Three participants were eliminated from the final analyses for failing to complete all items, leaving a final sample of 50 males, 50 females, eight Black participants, ten Asian participants, 7 Hispanic/Latino participants, 72 White participants, two American Indian participants, one Arab American participant, and one participant self-described as “Other.” The mean age for the sample was 34 years old with a standard deviation of 11.42 years.
For the first portion of the study, participants were given the following task instructions: “On the following questions, pull the slider to indicate how you would prefer to be seen in comparison with your closest peer or friend group on various traits. The scale point of -10 represents that you'd like to be seen as possessing this trait to a MUCH LOWER DEGREE than the rest of your friends/peer group. The scale point of 0 represents that you'd like to be seen as possessing this trait EQUALLY AS MUCH AS the rest of your friends/peer group. The scale point of 10 represents that you'd like to be seen as possessing this trait to a MUCH GREATER DEGREE than the rest of your friends/peer group.” Then, they were presented with a list of eleven traits, including “honest,” “intelligent,” “caring,” “responsible,” “intellectual,” “attractive,” “creative,” “shrewd,” “selfish,” “knowledgeable,” and “judgmental.”

We then asked participants to complete the Moral Identity Scale, which consisted of 13 items measured on 7-point scale where 1 = “strongly disagree” and 7 = “strongly agree.” Prior to completing the survey items, participants were given a list of nine morality-related traits, including “caring,” “helpful,” and “honest,” and were instructed to “visualize in your mind the kind of person who has these characteristics. Imagine how that person would think, feel, and act.” The sample items from the proceeding survey included, “It would make me feel good to be a person who has these characteristics,” and “Being someone who has these characteristics is an important part of who I am.”

Finally, participants were told to consider four different scenarios, which were counterbalanced, and asked to what degree they would experience a list of emotions in response to being in the described scenario. Two of the scenarios pertained to
intellectually outperforming one’s peers, including either receiving mostly correct answers on an exam while one’s friends received mostly incorrect answers or taking a training session with one’s peers wherein it became obvious that one was learning the material at a much faster pace. The other two scenarios related to acting in a morally superior manner to one’s peers, including either making the “right” choice in an ethical dilemma while one’s friends made the “wrong” choice or having to excuse oneself from an activity one’s peers were participating in due to religious or moral convictions. In response to the given scenarios, participants were asked to rate on a 5-point scale the degree to which they would feel uncomfortable, happy, anxious, good, sad, angry, and confident. The happy, good, and confident items were reverse-scored.

Results

We predicted that participants would prefer to be seen as more intelligent relative to their peers, and that participants would imagine themselves experiencing more discomfort in response to the “morally superior” scenarios rather than the “intellectually superior” scenarios. Furthermore, we thought these results might be moderated by moral identity.

To test the first prediction, we created a composite of “moral traits” and “intellectual traits” from the first portion of the study by combining the “caring,” “honest,” “selfish” (reverse coded) and “judgmental” (reverse coded) traits into one composite measure for morality (α = .82), and also combining “intelligent,” “intellectual,” and “knowledgeable” into one composite measure for intelligence (α = .93). Paired-samples t-tests revealed that participants marginally preferred to be seen as
being intellectually superior to their peers \((M = 4.26, SD = 3.43)\) rather than morally superior \((M = 3.73, SD = 3.45)\), \(t(98) = -1.62, p = .10\).

We next tested our second prediction about the hypothetical scenarios by conducting paired-samples t-tests by “morally superior” scenarios versus “intellectually superior” scenarios on a composite measure of discomfort created by combining, “How anxious would you feel?” and “How uncomfortable would you feel?” Consistent with our predictions, participants reported that they would feel more discomfort making the “right” moral choice in an ethical dilemma when their friends were making the wrong choice \((M = 2.06, SD = .93)\) than they would feel if they received mostly correct answers on an exam while their friends received mostly incorrect answers \((M = 1.72, SD = .86)\), \(t(98) = 3.65, p < .001\). Participants also reported that they would feel more discomfort declining to participate an activity with friends due to moral convictions \((M = 2.30, SD = 1.01)\) than they would feel if they were learning material in a training session at a much faster rate than their peers \((M = 1.74, SD = .91)\), \(t(98) = 5.02, p < .001\). Furthermore, combining the scenarios by either moral relevance or intellectual relevance revealed that participants overall felt more discomfort in scenarios wherein they behaved morally superior to their peers \((M = 2.17, SD = .80)\) rather than intellectually superior to their peers \((M = 1.73, SD = .81)\), \(t(98) = 5.74, p < .001\).

There were also significant effects for other, non-discomfort-related emotions. For morality-related scenarios and intelligence-related scenarios, we created two composites, such that feeling “happy,” “good,” “confident,” “sad” (reversed coded), and “angry” (reversed coded) were combined into one scale for each type of scenario.
(intelligence $\alpha = .79$, morality $\alpha = .77$). A paired samples t-test on these composites showed that participants reported feeling more overall positive affect when imaging the intelligence-related scenarios ($M = 4.76$, $SD = .76$) rather than the morality-related scenarios ($M = 4.26$, $SD = .85$), $t(98) = 4.51$, $p < .001$.

Correlational analyses revealed that moral identity was significantly negatively related to reported discomfort with the “morally superior” scenarios, such that those who reported a stronger moral identity reported less imagined discomfort with the scenarios, $r = -.24$, $p < .05$. Participants who reported a stronger moral identity were also more likely to report preferring to be seen as superior to their peers on morality-related traits, $r = .34$, $p < .001$. Moral identity was uncorrelated with imagined discomfort in intelligence-related scenarios ($r = -.17$, $p = .09$) and preference for being seen as possessing intelligence-related traits, $r = .18$, $p = .08$.

Study 3 conceptually replicated and extended the findings of our pilot study. As an additional replication, in Study 4 we attempted to find evidence that participants primed with the threat of do-gooder derogation would be motivated to downplay their own moral characters. We did this through either inducing participants to have thoughts related to do-gooder derogation prior to answering morality-related questions about themselves or not.
Study 4

The foundation of the experimental manipulation in Study 4 was derived from relevant research on do-goober derogation (Minson & Monin, 2012). In this research, participants were asked to free-associate words with vegetarians. The free association of negatively-valenced words like “preachy” or “holier-than-thou” with vegetarians was found to be related to the degree that participants expected vegetarians to cast negative moral judgments on meat-eaters like themselves. We used this research as a basis for the central independent variable in our study, which was whether participants were randomly assigned to imagine a hypothetical scenario of a dinner party at which a guest declines to partake in the meat entrée and announces that his choice is due to moral reasons, or to imagine a hypothetical scenario where the guest announces that his choice is due to a doctor’s orders. Then, we measured participants’ scores on the Moral Identity Scale and the degree they wished that others saw them as possessing various morality-related characteristics.

Method

Participants and design. One hundred and five participants were recruited through Amazon Mechanical Turk in exchange for $0.40. All participants were included in the final analyses, leaving a sample of 55 males, 50 females, 12 Black participants, 15 Asian participants, seven Hispanic/Latino participants, 74 White participants, one American Indian participant, and one Arab American participants. When asked about their own personal dietary restrictions, ten participants responded that they were either
vegetarian or vegan, but removing these participants did not change the results of the analysis.

Participants were randomly assigned to either imagine a scenario where a dinner party guest declines a meat entrée as a result of personal moral convictions or as a result of health recommendations from his doctor. Then, participants rated this hypothetical person from 0 to 100 on various trait dimensions, including moral, annoying, self-righteous, honest, ethical, thoughtful, weird, preachy, judgmental, arrogant, compassionate, and popular. After the manipulation, participants filled out a shortened, 9-item version of the Moral Identity Scale survey, which can be found in Appendix A, and also rated to what degree, on a 0 to 100 scale, they would like to be seen by others as possessing the aforementioned characteristics. At the conclusion of the study, participants were thanked and debriefed.

Results

We predicted that participants presented with the scenario wherein the dinner party guest declines the meat for moral reasons would have a lower mean moral identity score than participants presented with the scenario wherein the dinner party guest declines the meat for medical reasons. We also predicted that these participants would report less of a desire to be seen as moral, ethical, preachy, and judgmental.

To test these predictions, we first created two separate composite measures for the traits that participants reported that they would like to be seen by others as possessing. The traits “preachy,” “judgmental,” and “self-righteous” made up one composite ($\alpha = .88$, $M = 20.23$, $SD = 23.40$), while “moral,” “honest,” “ethical,” and “compassionate” made
up the second ($\alpha = .92, M = 76.59, SD = 22.15$). We conducted a series of independent samples t-tests on these composites, along with participants’ mean score on the Moral Identity Scale. There was no significant difference in the composite measure for the desire to be seen as moral between the conditions, $p = .29$. There were effects somewhat trending in the opposite of the predicted direction, such that those in the “moral vegetarian” condition reported a stronger moral identity ($M = 5.2, SD = 1.07$) than those in the “health vegetarian” condition ($M = 4.99, SD = .84$), $t(102) = 1.21, p = .23$, and that those in the “moral vegetarian” condition reported a stronger desire to be seen as possessing traits placed into our composite measure for “preachy” ($M = 22.59, SD = 24.76$) than those in the health vegetarian condition, ($M = 17.26, SD = 21.45$), $t(102) = 1.16, p = .25$. However, this could be due to participants demonstrating an anchoring or assimilation effect towards the target in the “moral vegetarian” condition.
General Discussion

Although the present studies supported some of our predictions, the package of studies presented here provides mixed support for the presented claims overall. In line with our hypotheses, participants across two studies imagined experiencing less positive affect and more negative affect, such as discomfort, upon receiving feedback that they were far more moral than their peers rather than far more intelligent. Further, participants reported having an overall preference for receiving feedback that they were far more intelligent than their peers over receiving feedback that they were far more moral than their peers. Contrary to our hypotheses, however, participants who were given ostensibly real feedback that they were morally superior in comparison with their peers were no more likely to emphasize their participation in morally deviant behaviors to a peer, nor were they more likely to portray themselves as less judgmental and more accepting, than participants given feedback that they were morally average in comparison with their peers. Although it is possible that the hypothesized effects do not exist in reality, there were some serious limitations in this package of studies that may have contributed to our failure to detect effects.

Firstly, a necessary component of achieving our hypothesized effect was fearing some type of negative evaluation or ill will from one’s peers. The only study to contain an experimental realistic element of potential evaluation from a peer was Study 2, and it is possible that participants in this study viewed the undergraduate experimenter not as a peer, but as an authority figure due to the context of the experimental setting. Improving on this study design would require the use of confederates posing as undergraduate
students taking the same survey to present while the participants were responding to the dependent variable measures. Another way to elicit the fear of envy that would cause participants to downplay their level of morality would be to lead them to believe that their very high “morality score” would be visible to members of their peer group in some way before publicly answering a set of questions about their own behaviors.

Another serious limitation of this package of studies was the lack of experimental realism in the majority of the studies, which simply asked participants to imagine hypothetical scenarios rather than directly experience a situation that would make moral self-presentation concerns salient. These designs primarily relied on collecting online samples of participants using a remote platform wherein it is impossible to ensure proper debriefing, rendering the use of deception ethically impermissible. Although our lab study, Study 2, improved on this limitation by having participants take what they believed to be a real “morality test,” the degree to which participants in the “far more moral” condition found their feedback to be surprising and unexpected in comparison with participants in the “average morality” condition may be an indication of the lack of believability of such feedback. Indeed, although many people may desire to believe that they are somewhat above average on various positive trait dimensions, it may be hard to believe that one is an outlier or truly exceptional on these same dimensions. A possible way to improve upon this limitation might be to recruit participants who are seen by others as “morally superior” in some way, such as vegetarians, and then experimentally manipulate the threat of being viewed as “holier-than-thou” by pairing them with other vegetarians versus non-vegetarians.
Finally, the present research was also limited by the complications inherent in operationalizing moral behavior and moral character. Definitions of morality vary considerably from person to person. Recent research on Moral Foundations Theory argues strongly for moral pluralism, or the idea that conceptions of morality stem from a variety of basic foundations that can differ widely in relative emphasis among cultures, as well as the idea that people derive their ideas of morality in part through cultural learning (Graham et al., 2012). Thus, it seems highly plausible, if not certain, that participants in studies that including taking our “morality test” had very different ideas about what in our given scenarios constituted correct moral behavior, as well as whether our dependent measures of so-called “deviant behaviors” contained examples of what they would consider “immoral behavior,” and to what degree. Future studies on this topic should attempt to reduce this variability by extensively pretesting and validating measures and manipulations across a wide variety of samples prior to conducting full studies.

Despite the limitations of the presented studies, it is important to acknowledge that people are more uncomfortable with being seen as morally superior to their peers rather than intellectually superior to their peers, and that they indeed report that they prefer being seen as intellectually superior to their peers rather than morally superior to their peers. This suggests that, as predicted, morality is a trait dimension distinct from other traits subject to the better-than-average effect. Future research should extend these findings to see whether this is true of morality in comparison with other positive traits, such as attractiveness, popularity, or tenacity. The initial findings of these studies provide evidence that being morally superior to one’s peers is to be in a distinctly
ambiguous affective state, wherein one experiences negative affect, mainly related to discomfort and anxiety, along with the positive affect accompanies feeling one is better-than-average. This finding is worthy of further exploration in future studies designed to ascertain what mechanism drives these experiences of negative affect in relation to being morally superior to one’s peers, whether it be thoughts related to a fear of envy or thoughts related to being taken advantage of or antisocial punishment.

A potential individual difference variable that may impact the degree to which people are sensitive to feedback about how moral they are in comparison with the rest of the group is the need to belong (NTB; Baumeister & Leary, 1995). This construct has been defined by Baumeister and Leary (1995) as an innate “need to form and maintain at least a minimum quantity of interpersonal relationship,” partially captures the drive to not be too distinct from relevant others. Researchers have shown that high NTB people are more highly perceptive about social cues compared to their low NTB counterparts (Pickett, Gardner, & Knowles, 2004). Further, some research has revealed that people high in NTB are more likely to conform to group norms during intergroup negotiations (Steinel et al., 2010). Thus, it seems reasonable to expect that people high in NTB will be more likely to avoid violating social norms. Future studies should seek to discover whether participants high in NTB are more sensitive to feedback about their level of morality relative to their peers, leading them to show more effort to avoid the possible backlash of being seen as "too moral" by a) more heavily emphasizing their deviant or immoral behaviors and/or b) downplaying the degree to which they socially judge others.


comparison on mood states. *Journal of Social and Clinical Psychology, 8*(1), 14.


Appendix A: Dependent Measures

This appendix includes complete scales for the outcome variables measured in this thesis. The sets of measures are separated by study.

Pilot Study

1) On a scale of 1-7, with 1="extremely unhappy" and 7="extremely happy" how happy would you be to receive feedback that you were far more intelligent than your peers?

2) On a scale of 1-7, with 1="not at all anxious" and 7="extremely anxious" how anxious would you be to receive feedback that you were far more intelligent than your peers?

3) On a scale of 1-7, with 1="not at all uncomfortable" and 7="extremely uncomfortable" how uncomfortable would you be to receive feedback that you were far more intelligent than your peers?

4) On a scale of 1-7, with 1="extremely bad" and 7="extremely good" how good versus bad would you feel to receive feedback that you were far more intelligent than your peers?

5) On a scale of 1-7, with 1="extremely bad" and 7="extremely good" how good versus bad would you feel if your peers discovered that you tested as far more intelligent than them?

6) On a scale of 1-7, with 1="not at all uncomfortable" and 7="extremely uncomfortable" how uncomfortable would you be if your peers discovered that you tested as far more intelligent than them?
7) On a scale of 1-7, with 1="extremely unhappy" and 7="extremely happy" how **happy** would you be to receive feedback that you were **of average intelligence**?

8) On a scale of 1-7, with 1="not at all anxious" and 7="extremely anxious" how **anxious** would you be to receive feedback that you were **of average intelligence**?

9) On a scale of 1-7, with 1="not at all uncomfortable" and 7="extremely uncomfortable" how **uncomfortable** would you be to receive feedback that you were **of average intelligence**?

10) On a scale of 1-7, with 1="extremely bad" and 7="extremely good" how **good versus bad** would you feel to receive feedback that you were **of average intelligence**?

11) On a scale of 1-7, with 1="extremely bad" and 7="extremely good" how **good versus bad** would you feel **if your peers discovered** that you tested **of average intelligence**?

12) On a scale of 1-7, with 1="not at all uncomfortable" and 7="extremely uncomfortable" how **uncomfortable** would you be **if your peers discovered** that you tested **of average intelligence**?

13) On a scale of 1-7, with 1="extremely unhappy" and 7="extremely happy" how **happy** would you be to receive feedback that you were **far more moral** than your peers?
14) On a scale of 1-7, with 1="not at all anxious" and 7="extremely anxious" how anxious would you be to receive feedback that you were far more moral than your peers?

15) On a scale of 1-7, with 1="not at all uncomfortable" and 7="extremely uncomfortable" how uncomfortable would you be to receive feedback that you were far more moral than your peers?

16) On a scale of 1-7, with 1="extremely bad" and 7="extremely good" how good versus bad would you feel to receive feedback that you were far more moral than your peers?

17) On a scale of 1-7, with 1="extremely bad" and 7="extremely good" how good versus bad would you feel if your peers discovered that you tested as far more moral than them?

18) On a scale of 1-7, with 1="not at all uncomfortable" and 7="extremely uncomfortable" how uncomfortable would you be if your peers discovered that you tested as far more moral than them?

19) On a scale of 1-7, with 1="extremely unhappy" and 7="extremely happy" how happy would you be to receive feedback that you were morally average compared to your peers?

20) On a scale of 1-7, with 1="not at all anxious" and 7="extremely anxious" how anxious would you be to receive feedback that you were morally average compared to your peers?
21) On a scale of 1-7, with 1="not at all uncomfortable" and 7="extremely uncomfortable" how uncomfortable would you be to receive feedback that you were morally average compared to your peers?

22) On a scale of 1-7, with 1="extremely bad" and 7="extremely good" how good versus bad would you feel to receive feedback that you were morally average compared to your peers?

23) On a scale of 1-7, with 1="extremely bad" and 7="extremely good" how good versus bad would you feel if your peers discovered that you tested as morally average compared to them?

24) On a scale of 1-7, with 1="not at all uncomfortable" and 7="extremely uncomfortable" how uncomfortable would you be if your peers discovered that you tested as morally average compared to them?

25) Which type of feedback would you prefer to get: that you were far more moral than the rest of your peers or that you were far more intelligent than the rest of your peers?

- Definitely prefer "far more moral"
- Moderately prefer "far more moral"
- Slightly prefer "far more moral"
- No preference between "far more moral" and "far more intelligent"
- Slightly prefer "far more intelligent"
- Moderately prefer "far more intelligent"
26) Which type of feedback would you prefer to get: that you were far less moral compared to your peers than the rest of your peers or that you were far less intelligent compared to your peers?

☐ Definitely prefer "far less moral"

☐ Moderately prefer "far less moral"

☐ Slightly prefer "far less moral"

☐ No preference between "far less moral" and "far less intelligent"

☐ Slightly prefer "far less intelligent"

☐ Moderately prefer "far less intelligent"

☐ Definitely prefer "far less intelligent"

**Study 1 and Study 2**

After receiving their “Morality Test” score, participants were asked the following questions:

1) On a scale of 1-7, where 1=totally expected and 7=totally unexpected how much would you expect your score to fall within the range that it did?

2) On a scale of 1-7, where 1=not at all surprised and 7=totally surprised, how surprised would you be that your score fell within the range that it did?
In the next part of our study, you will answer some questions about yourself, your habits, and your personality. Your answers will then be shown to an anonymous partner, whose answers you will also see, as part of an impression formation task.

1) The average college student says they study about 5-7 hours per week. To what extent would you say that you study less versus more than this amount per week?

- Much less than
- Less than
- A little less than
- About the same as
- A little more than
- More than
- Much more than

2) The average college student says they drink alcohol about two nights per week. How many nights per week do you drink alcohol?

3) Binge drinking is defined as five or more drinks per sitting. On a scale of 1-7, with 1="never," 4="sometimes," and 7="often," how often would you say you engage in binge drinking?
4) Do you ever post what might be considered "profanity" on social media? On a scale of 1-7, with 1="never," 4="sometimes," and 7="often," how often would you say this occurs?

5) Do you and your friends engage in behaviors that your parents or supervisors would disapprove of? On a scale of 1-7, with 1="never," 4="sometimes," and 7="often," how often would you say this occurs?

6) On a scale of 1-7, with 1 being definitely not and 7 being definitely, how much do you think you reflect what would be considered a "model student," that is, a student who is an exemplary citizen of your institution?

7) On a scale of 1-7, with 1 being definitely not and 7 being definitely, to what degree would you say that you're a responsible student?

8) The average college student says they generally put about 60% effort into their classes overall. From 0-100%, what percentage of effort do you put into your classes?

9) On a scale of 1-7, with 1 being very unlikely and 7 being very likely, how likely are you to volunteer your time to bettering your community in the next month?

10) On a scale of 1-7, with 1 being very unlikely and 7 being very likely, how likely are you to donate to a charitable cause in the next month?

11) On a scale of 1-7, how negatively would you judge a fellow student who binge drinks or "parties" often?
12) On a scale of 1-7, how negatively would you judge a fellow student who uses recreational drugs?

13) On a scale of 1-7, how negatively would you judge a fellow student who frequently uses profanity?

14) On a scale of 1-7, how negatively would you judge someone for being an irresponsible student?

15) On a scale of 1-7, how negatively would you judge a fellow student for being disrespectful towards their instructor?

16) On a scale of 1-7, how negatively would you judge a fellow student for cheating on an exam?

17) On a scale from 1-7, with 1=not at all judgmental and 7=very judgmental, how judgmental of a person would you say you are relative to your peers?

18) On a scale from 1-7, with 1=not at all accepting and 7=very accepting, how accepting of imperfections in others would you say you are relative to your peers?

**Study 3**

Below are the full instruction set followed by the complete 9-item version of the Moral Identity Scale presented to participants in Study 3. Each of the items were evaluated on a 7-point scale, where 1=“strongly disagree” and 7=“strongly agree.”

Listed below are some characteristics that may describe a person [list of nine traits]. The person with these characteristics could be you or it could be someone else. For a moment, visualize in your mind the kind of person who has these characteristics.
Imagine how that person would think, feel, and act. When you have a clear image of what this person would be like, answer the following questions.

Caring
Compassionate
Fair
Friendly
Generous
Hardworking
Helpful
Honest
Kind

1) It would make me feel good to be a person who has these characteristics.
2) Being someone who has these characteristics is an important part of who I am.
3) A big part of my emotional well-being is tied up in having these characteristics.
4) I would be ashamed to be a person who has these characteristics.
5) Having these characteristics is not really important to me.
6) Having these characteristics is an important part of my sense of self.
7) I strongly desire to have these characteristics.
8) I often buy products that communicate the fact that I have these characteristics.
9) I often wear clothes that identify me as having these characteristics.

10) The types of things I do in my spare time (e.g., hobbies) clearly identify me as having these characteristics.

11) The kinds of books and magazines that I read identify me as having these characteristics.

12) The fact that I have these characteristics is communicated to others by my membership in certain organizations.

13) I am actively involved in activities that communicate to others that I have these characteristics.

**Study 4**

In Study 4, we used a shortened, 9-item version of the Moral Identity Scale, which is included below.

1) It would make me feel good to be a person who has these characteristics.

2) Being someone who has these characteristics is an important part of who I am.

3) A big part of my emotional well-being is tied up in having these characteristics.

4) I would be ashamed to be a person who has these characteristics.

5) Having these characteristics is an important part of my sense of self.

6) I strongly desire to have these characteristics.

7) The types of things I do in my spare time (e.g., hobbies) clearly identify me as having these characteristics.
8) The fact that I have these characteristics is communicated to others by my membership in certain organizations. The types of things I do in my spare time (e.g., hobbies) clearly identify me as having these characteristics.

9) I am actively involved in activities that communicate to others that I have these characteristics.
Appendix B: Morality Test

This appendix includes the complete “Morality Test” given to participants, which was described as assessing participants’ level of morality relative to their peers. The morality test consisted of three separate components: a moral choices test, an empathy test, and an evaluation of which moral violations participants would perform for a large money. Participants read the following directions:

The following test is designed to evaluate how you make moral choices and how frequently you choose to behave morally. Please answer the following questions as honestly as possible.

Moral Choices Test

Evaluate the following scenarios based on what you think is the right or wrong thing to do. Keep in mind that "morally obliged" means that you must perform the action in order to do the right thing, and that "morally permissible" means that performing the action is not the wrong thing to do.

1) You pass someone in the street who is in severe need and you are able to help them at little cost to yourself. Are you morally obligated to do so?
   A) Strongly Morally Obligated
   B) Weakly Morally Obligated
   C) Not Morally Obligated

2) You have a brother. You know that someone has been seriously injured as a result of criminal activity undertaken by him. You live in a country where the
police and legal system are generally trustworthy. Are you morally obliged to inform them about your brother's crime?

A) Strongly Morally Obligated
B) Weakly Morally Obligated
C) Not Morally Obligated

3) You see a charity advertisement in a newspaper about a person in severe need in India. There is no state welfare available to this person, but you can help them at little cost to yourself. You have good reason to believe that any help you offer will make a difference. Are you morally obliged to help the person?

A) Strongly Morally Obligated
B) Weakly Morally Obligated
C) Not Morally Obligated

4) A situation arises where you can either save your own child from death or contact the emergency services in order to save the lives of ten other children. You cannot do both, and there is no way to save everybody. Which course of action are you morally obliged to follow?

A) Save your own child
B) Save 10 children

5) You can save the lives of a million innocent people by killing a hundred thousand others. Are you morally obliged to do so?

A) Yes
B) No
6) A runaway trolley is heading down the tracks toward five workmen who will be killed if the trolley proceeds on its present course. You are on a footbridge over the tracks, in between the approaching trolley and the five workmen. Next to you on this footbridge is a stranger who happens to be very large. The only way to save the lives of the five workmen is to push this stranger off the bridge and onto the tracks below where his large body will stop the trolley. The stranger will die if you do this, but the five workmen will be saved. Pushing this man is...

A) Morally permissible  
B) Morally obligatory  
C) Morally forbidden

**Empathy Test**

Please indicate whether the following statements describe you well.

1) I sometimes find it difficult to see things from the 'other guy's' point of view.
   A) Does not describe me well  
   B) Neutral  
   C) Describes me very well 

2) I sometimes feel helpless when I am in the middle of a very emotional situation.
   A) Does not describe me well  
   B) Neutral  
   C) Describes me very well
3) Sometimes I don't feel very sorry for other people when they are having problems.
   A) Does not describe me well
   B) Neutral
   C) Describes me very well

4) I believe that there are two sides to every question and try to look at them both.
   A) Does not describe me well
   B) Neutral
   C) Describes me very well

5) Other people's misfortunes do not usually disturb me a great deal.
   A) Does not describe me well
   B) Neutral
   C) Describes me very well

6) I often have tender, concerned feelings for people less fortunate than me.
   A) Does not describe me well
   B) Neutral
   C) Describes me very well

7) When I see someone being taken advantage of, I feel kind of protective towards them.
   A) Does not describe me well
   B) Neutral
   C) Describes me very well
8) Before criticizing somebody, I try to imagine how I would feel if I were in their place.
   
   A) Does not describe me well
   
   B) Neutral
   
   C) Describes me very well

**What Would You Do for Money?**

For each of the actions below, indicate whether you would be willing to perform the action for $10,000 by selecting either “yes” or “no.”

1) After your pet dog dies of natural causes, cook and eat it.

2) When a family member tells you they have been diagnosed with cancer, don’t say anything and start watching television instead.

3) Tell someone who has been attacked by a dog that the scars on their face are ugly.

4) Slap a stranger in the face.

5) Break into a public building, where trespassing is not allowed.

6) Walk around a small pond.

7) Tell a friend who is about to get married that their partner is overweight.

8) Sneak into a stranger's house, look through their family photographs, dressers, and closet, then leave without a trace.

9) Take candy without paying because the check-out line at the store is too long.

10) Punch a friend in the stomach as hard as you can (with the friend’s permission).