

# Religion, Repulsion, and Reaction Formation: Transforming Repellent Attractions and Repulsions

Dov Cohen  
University of Illinois

Emily Kim  
Yale NUS College

Nathan W. Hudson  
Southern Methodist University

Protestants were more likely than non-Protestants to demonstrate phenomena consistent with the use of reaction formation. Lab experiments showed that when manipulations were designed to produce taboo attractions (to unconventional sexual practices), Protestants instead showed greater repulsion. When implicitly conditioned to produce taboo repulsions (to African Americans), Protestants instead showed greater attraction. Supportive evidence from other studies came from clinicians' judgments, defense mechanism inventories, and a survey of respondent attitudes. Other work showed that Protestants who diminished and displaced threatening affect were more likely to sublimate this affect into creative activities; the present work showed that Protestants who do not or cannot diminish or displace such threatening affect instead *reverse* it. Traditional individual difference variables showed little ability to predict reaction formation, suggesting that the observed processes go beyond what we normally study when we talk about self-control.

*Keywords:* culture, defense, reaction formation, religion, repulsion

Humans are attracted and repulsed by various stimuli. Moreover, they can be repulsed by their own attractions and repulsions (Miller, 1997). The question of how people deal with feelings they would rather not have is one that has interested psychologists, beginning with Freud and continuing with modern personality, social, and clinical psychology. Typically, psychologists have talked about how people can control or “defend” against such unwanted feelings. However, “defend” can be a misnomer, because sometimes the best defense is a good offense.

Thus, one of the most intriguing methods of dealing with an unwanted desire is through the phenomenon of reaction formation, by which an undesired feeling is replaced by its opposite. Antipathy or repulsion is replaced by attraction; attraction is replaced by repulsion or antipathy. There is likely individual difference variation in how often such a defense is used. And beyond this, there are likely cultural and religious differences in people's likelihood of using reaction formation to fend off what would be repellent thoughts and desires. In this paper, we explore the possibility that adherents of Protestant denominations are more likely than non-Protestants to defend against forbidden thoughts and emotions by turning them to their opposite.

We begin with a brief review of reaction formation phenomena, explain why one would expect Protestants to be generally more

likely to engage in it, and then outline the present studies. To be clear, we do not think reaction formation is *only* engaged in by Protestants, simply that it is more likely to be engaged in by them (as illustrated by the studies in this paper that concern chronic individual differences) or is more easily induced in them (as illustrated by the studies involving situational manipulations).

## Reaction Formation: Evidence, Distinctions, and Moderation

Freud viewed reaction formation as a defense in which threatening thoughts, feelings, or motives become transformed into their opposite. Reaction formation goes beyond *denial* of such thoughts and feelings and instead represents a psychological “overcompensation.” For example, an attraction that is threatening (e.g., because it is directed toward a target deemed unacceptable) may be transformed into hostility toward that target. Conversely, a hatred that is threatening (e.g., because it is directed toward an inappropriate target or because hatred in general is unacceptable) may be transformed into more positive feelings.

## Evidence

Though the defense of reaction formation is counterintuitive, psychologists have found some evidence for its existence. In their review, Baumeister, Dale, and Sommer (1998, p. 1089) conclude that “when people are publicly or implicitly accused of having socially undesirable sexual feelings, prejudiced attitudes, or failures of competence, some respond by asserting the opposite (and attempting to prove it) to an exceptionally high degree.” That is, studies demonstrate, for example, that some homophobic men

---

This article was published Online First June 12, 2017.

Dov Cohen, Department of Psychology, University of Illinois; Emily Kim, Division of Social Sciences, Yale NUS College; Nathan W. Hudson, Department of Psychology, Southern Methodist University.

Correspondence concerning this article should be addressed to Dov Cohen, Department of Psychology, University of Illinois, 603 East Daniel, Champaign, IL 61820. E-mail: [dovcohen@illinois.edu](mailto:dovcohen@illinois.edu)

have physiological and behavioral responses that seem to imply an attraction to same-sex coupling (Adams, Wright, & Lohr, 1996; Cheval et al., 2016; cf. Mahaffey, Bryan, Ito, & Hutchison, 2011; Weinstein et al., 2012). In experiments, accusations of racism or sexism often lead participants to subsequently show less prejudice (ex. Dutton & Lake, 1973). And initial failure on tasks sometimes leads people—particularly high self-esteem people (McFarlin, Baumeister, & Blascovich, 1984)—to show unwarranted overconfidence (Baumeister, Heatherton, & Tice, 1993; Jordan et al., 2003; McGregor & Marigold, 2003). Baumeister and colleagues concluded that whereas the particular mechanism (or mechanisms) for reaction formation are unknown, it “apparently . . . is one of the most prominent and common responses to esteem threat” (p. 1089).

### Distinctions

Whereas we agree with Baumeister and colleagues that the general phenomenon is well supported, it may be useful to make some distinctions. The first has to do with the reassertion of one’s competence after a threat to one’s ability. Such assertions are likely correctly seen as defensive. However, strictly speaking, it is not clear that this is what is classically meant by reaction formation—or at least what we mean here by reaction formation—which involves the reversal of unacceptable desires or *morally* problematic feelings.

The second distinction applies to the prejudice research, much of which involves the public exposure of one’s undesirable attitudes, the public or semipublic compensatory behavior that results, and occasionally both. Though there are exceptions (ex. Sherman & Gorkin, 1980), as Baumeister and colleagues (1998) note, it is not always clear whether such phenomena involve (a) internal psychological processes or (b) behavior designed to impress others or at least counter their false impressions. It is plausible that the latter goal is in the service of the former (If I can convince you that I’m not prejudiced, then that reassures me that I’m not actually prejudiced; Mead, 1934). It is also plausible that the former is in service of the latter (If I need to convince you that I’m not prejudiced, I should internally suppress my own prejudiced thoughts so they do not leak out during our interaction; Kunda & Spencer, 2003). Thus, both theoretically and practically, it can be hard to separate out who the “audience” is for a given act. Nevertheless, it is generally useful to explore reaction formation as an internal psychological process in settings that reduce concerns with socially desirable responding, do not involve external criticism or public exposure of undesirable attitudes, and lessen any experimental demand characteristics to consciously reflect on and change such attitudes.

The third distinction applies to homophobia research. This research takes prejudice against homosexuals not as the undesirable attitude but as the *response* to an unacceptable attraction to same-sex others. Such research *measures* (rather than experimentally induces) reaction formation by taking people with antigay prejudices and then using physiological or reaction time (RT) tasks as hard-to-repress signs that such people may have same sex attractions (ex. Adams et al., 1996). This research begins with the final attitude (antigay prejudice) and looks to see what might underlie it (ex. same-sex attraction) in some cases. Such research seems to support the existence of reaction formation (though studies were

correlational and there is some uncertainty about what the physiological data imply). However, two of the experimental studies in the present paper usefully add to this work by starting at the other end of the process: they begin with manipulations trying to experimentally induce an unacceptable desire or feeling and then measure participants’ responses. The manipulations try to induce the unacceptable desires subtly, without publicly exposing or calling attention to them, and then (privately) measure participants’ attitudes. The assumption is that, to the extent that participants’ (privately) measured attitudes become more desirable/correct/ socially appropriate when the experimental manipulations should have made them less so, then we are seeing evidence of experimentally created reaction formation (see also Sarnoff, 1960). (We further differentiate reaction formation from simple reactance [Brehm, 1966] later in the paper).

### Moderators

Though self-esteem has been implicated as a moderator in defensive responses to incompetency (ex. Jordan et al., 2005; Sommer & Baumeister, 2002; Vohs & Heatherton, 2004), much less is known about individual differences that might moderate threats of other types, and, to our knowledge, there is no systematic evidence (either way) concerning religious or cultural differences in the use of reaction formation.<sup>1</sup> In the present studies, we examine both individual and religious differences. To anticipate our findings, we hypothesized and found consistent religious differences across 5 studies, with Protestants being more likely to show reaction formation, as compared with Catholics and Jews. We found little in the way of individual differences that produced effects analogous to that of religion—results that also demonstrate some discriminant validity for reaction formation phenomena, suggesting that such phenomena go beyond the “normal” personality processes we discuss when we think of self-control, emotion regulation, agreeableness, dogmatism, general insecurity, and so on. The exception is Study 5, where we identify individual differences in types of defensive responding that differentiate between Protestants who are likely to sublimate inappropriate thoughts and feelings into creative activity (Cohen, Kim, & Hudson, 2014; Hudson & Cohen, 2016; Kim & Cohen, 2017) versus those who are likely to show reaction formation by reversing them.

### Why Protestantism? The Moral Culture of a “New” Religion

The religious differences we explored are in line with theological bases of Protestantism versus Catholicism and Judaism. Within the Judeo-Christian tradition, Protestantism is relatively new, being “only” about 500 years old. Catholicism is 2000 years old, if one dates it to the beginning of Christianity. Judaism is thought to be about 3500 years old. Protestantism, as the latest developing branch, has historically held with what Darwin (2008, p. 11) called: “The highest stage in moral culture at which we can arrive . . . when we recognize that we ought to control our

<sup>1</sup> A. Cohen (2009; see also A. Cohen, 2013; A. Cohen & Varnum, 2016) argued for conceiving of religions in terms of cultures. We think this conception is apt, and because we do not define cultures in geographic terms here, we use the words religion and culture interchangeably.

thoughts, and ‘not even in inmost thought to think again the sins that made the past so pleasant to us.’” The older religions—Judaism, Catholicism, as well as Hinduism, the oldest surviving world religion—put a strong emphasis on behavior; the “new” religion of Protestantism gave prominent place to thoughts (A. Cohen & Rozin, 2001; A. Cohen, 2015; Hughes, Grossmann, & Cohen, 2015; also Abramowitz et al., 2004; Berman et al., 2010; Laurin & Plaks, 2014).

To be clear, no religion is likely to condone antisocial thoughts (Ritter & Preston, 2011; Ritter et al., 2016; Vishkin et al., 2016). However, Protestantism takes the focus on internal mental processes much further than the other Judeo-Christian religions do. Generally, the emphasis in Protestantism is on an individual’s faith as a route to salvation. This contrasts with Catholicism, where works play a larger role, and Judaism, where there is a strong emphasis on behavior (A. Cohen, 2003; A. Cohen & Rankin, 2004; A. Cohen & Rozin, 2001; A. Cohen, Siegel, & Rozin, 2003; Siev & A. Cohen, 2007).

Moreover, contemporary Protestants make judgments in line with what A. Cohen and Rozin call a “morality of mentality,” regarding illicit thoughts and desires as sinful in and of themselves, even if they have virtually no possibility of being realized. Protestant respondents are also far more likely to endorse the belief that consciously entertaining thoughts about doing something immoral is as bad as actually doing it, whereas Jews, and to a lesser extent Catholics, made stronger distinctions between thinking and doing (A. Cohen, 2003; A. Cohen & Rozin, 2001; Kim & Cohen, 2017; Kim, Zeppenfeld, & Cohen, 2013; also Abramowitz et al., 2004; Laurin & Plaks, 2014; Siev, Chambless, & Huppert, 2010; more generally, see Sanchez-Burks, 2002, 2004; Uhlmann, Poehlman, Tannenbaum, & Bargh, 2011; Uhlmann & Sanchez-Burks, 2014 on the continued relevance of Protestant theology for stricter emotion regulation in the workplace and a strain of “implicit Puritanism” in American life).

Because of this “morality of mentality,” we expect Protestants would be more likely than non-Protestants to be threatened by inappropriate feelings and desires and would be more likely to engage in reaction formation. Reaction formation is not the only way one can deal with such feelings; there are a number of other defenses. We suspect, however, that reaction formation is a likely outcome because much religious doctrine—particularly Christian religious doctrine—reads like an incitement to reaction formation. The section in the Sermon on the Mount known as the “Antitheses” (from Matthew 5) is particularly revealing. After telling his listeners to not commit “adultery in your heart” (because merely entertaining lust is wrong), Jesus goes on to reverse what might seem to be “natural” sentiments:

Do not resist an evil person. If anyone slaps you on the right cheek, turn to them the other cheek also. And if anyone wants to sue you and take your shirt, hand over your coat as well. If anyone forces you to go one mile, go with them two miles. . . . You have heard that it was said, ‘Love your neighbor and hate your enemy’. But I tell you, love your enemies and pray for those who persecute you. . . . If you love those who love you, what reward will you get? Are not even the tax collectors doing that? (Matthew 5, NIV).

Jesus seems to acknowledge that turning the other cheek and loving one’s enemies might not be a natural act, but adds, “If you love those who love you, what credit is that to you? Even sinners

love those who love them. And if you do good to those who are good to you, what credit is that to you? Even sinners do that” (Luke 6, New International Version). These passages are, of course, well-known, and the phrase “turning the other cheek” has even been used as a shorthand for reaction formation responses (Vaillant, 2012, p. 268; see also Sarnoff, 1960).

Jesus encourages listeners to “Be perfect, therefore, as your heavenly Father is perfect” (Matthew 5). The contrast here is most clear with Judaism, which does not urge perfection, does not ask one to love one’s enemies, and is actually quite skeptical about attempts to control one’s thoughts. The Baal Shem Tov, the 18th-century founder of Chasidic Judaism, was once asked how to distinguish a true spiritual leader from a false one. “Ask him if he knows a way to banish impure thoughts,” The Baal Shem Tov said. “If he does, he’s a fraud.” (paraphrase from Telushkin, 1994, p. 132; Wolpe, 2004).

Again, contrasts are most clear between Protestantism and Judaism, but similar differences hold between Protestants and Catholics (A. Cohen, personal communication; reanalyzed data from Kim et al., 2013). Catholics are less likely than Protestants to believe that bad thoughts are equivalent to bad acts or that bad deeds derive from bad souls (Li et al., 2012).<sup>2</sup>

Because of Protestantism’s relatively greater emphasis on mental states (compared with its closest cousins, Catholicism and Judaism), we expected Protestants would be more likely to endorse “love your enemy/ turn the other cheek” sentiments. In Study 1, we explored this issue and examined the basis for such sentiments, reanalyzing data from a national survey of the U.S. In Study 2, we examined clinicians’ judgments about whether study participants engaged in reaction formation, reanalyzing data from the only publicly available study we could find that employed clinicians’ assessments of reaction formation as a defense. In Studies 3 and 4, we moved to the lab, creating conditions that would induce taboo feelings in participants and then examining the extent to which participants seemed to reverse such feelings—either feeling repulsion at what they were induced to be attracted to (Study 3) or feeling attracted to what they were induced to be repulsed by (Study 4). These experimental studies attempted to minimize impression management concerns, and in Study 4, used evaluative conditioning (which participants were not aware of) to induce the forbidden feelings as well as dependent measures that were not transparently about these feelings. Study 5 was a paper-and-pencil study of participants’ defense mechanisms. It let us to look at levels of reaction formation as well as examine when participants were likely to use reaction formation as opposed to other defenses.

<sup>2</sup> Whereas Protestantism and Catholicism both have the same basic scripture (the New Testament), it is not unusual for different branches of a religion to come to different understandings of a text (witness the differences between, say, Ultraorthodox Judaism and Reform Judaism or between Sunni and Shia branches of Islam). Perhaps Protestantism’s greater emphasis on mental states derives from Protestants’ greater belief in the literal interpretation of the Bible (including the “Antitheses”), greater emphasis on salvation through faith, or its more individualistic, psychological orientation, as compared with Catholicism (A. Cohen, et al., 2005). Likely the religions were shaped by the milieu at their founding—Late Antiquity in the Ancient Near East for Catholicism versus Early Modern Europe for Protestantism. Deriving “essential” source(s) of Protestant versus Catholic differences in the “morality of mentality” (A. Cohen & Rozin, 2001) is beyond the scope here, however.

In Studies 3 through 5, we attempted to identify some standard individual difference variables that would predict reaction formation. We did not succeed in identifying such variables in Studies 3 and 4. This helps establish discriminant validity for reaction formation phenomena as distinct processes that go beyond the usual individual differences we discuss when we think of “normal” self-control and emotion regulation (e.g., neuroticism and conscientiousness), desire to be good or please others (agreeableness, social desirability), closed-mindedness or conventionality (dogmatism, need for closure, low openness to experience), general insecurity (neuroticism or attachment insecurity), and sensitivity to either reward or punishment (behavioral inhibition, behavioral activation, extroversion, and again, neuroticism). The effects in Studies 3 and 4 cannot be described as “merely” a product of these factors.

However, in Study 5, we did identify 2 individual difference factors that—among Protestants—predicted the tendency to engage in reaction formation. These individual difference factors predicted which Protestants seemed to engage in reaction formation and which tended to engage in sublimation (another defense on which there are religious group differences).

### A Final Distinction and a Final Note

In addition to having conventional individual difference measures that provided some discriminant validity for reaction formation processes (as opposed to other personality processes that are more frequently examined), we also used (mostly experimental) methods that minimized alternative interpretations. As described in the General Discussion, the results we find do not seem to derive from impression management concerns, are unlikely to be attributable either to reactance or demand characteristics (especially in Study 4), do not result from either implicit or explicit accusations from others, are not attributable to experimental manipulations that simply fail to influence Protestants (especially in Study 3), and cannot be entirely the product of non-Protestants simply being better able to hide their reaction formation in nonclinical settings (Study 2). The last point is particularly relevant for skeptics who believe that defensive processes are best ferreted out by trained clinicians, who can detect phenomena better than many more standard measures can (Shedler, Mayman, & Manis, 1993).

A second point: There is a distinction between the classic view of reaction formation in which the person is basically unaware that they even possess the socially undesirable feeling and a more modern view that seems to also include conscious attempts to cope with conscious, undesirable thoughts. We take a middle-range view. First, we doubt that there is much in the way of people’s feelings and attitudes that is *completely* unknown to them in the sense of never being consciously felt or thought. On the other hand, we think people can be quite adept at mislabeling inchoate feelings and dismissing undesirable cognitions, and we think it is likely that people often cannot recognize why they feel the way they do or why they hold particular beliefs and attitudes with such tenacity, intensity, or fervor (McGregor & Jordan, 2007; Wilson, 2004).

In Study 5, we discuss defenses that involve diminishing threatening affect and those that involve displacing such material. A person can attempt to suppress a thought or feeling (“I’m just not going to think about X”). But if suppression happens enough and

the person is sufficiently good at suppressing, that suppression may become a habit and more or less automatic. Similarly, a person can try to displace a negative thought or feeling onto another object or convince themselves that their feeling has another source. But if this happens enough and the person becomes sufficiently practiced, such displacement may become a habit or more or less automatic. In Study 5, we show that our Protestant participants who can neither diminish nor displace threatening affect tend to *reverse* it. And likewise, we think if this reversal happens enough, this process too will become a habit and be more or less automatically deployed.

We think it is not terribly fruitful to classify something as definitively conscious or unconscious, though it does make sense to talk about processes as more conscious or less conscious (Kim et al., 2013). In the studies here, we think various levels of consciousness were represented. Study 1’s attitude survey probably reflected a more conscious attempt at reversing one’s inclinations. Study 2—using judgments of psychiatrists from the 1950s—likely illustrated phenomena more in line with classic conceptions of reaction formation (i.e., the process is toward the extremely unconscious end). Studies 3 and 4 involved experiments that lie in the middle. They involved diffuse and inchoate affect (Study 4 probably more than Study 3), and it seems unlikely that participants understood the nature of their reversals—and their completely private responses were unlikely to be attempts to impress anyone else. Study 5 used two defense mechanism inventories to measure reaction formation. Because of the vehemence with which most of the statements in the inventories are expressed, we think it likely the reaction formation here occurred toward the less-conscious end of the spectrum. However, in all studies, we recognize that the phenomena may have occurred along a range of consciousness, with it being unlikely that effects are either completely conscious or unconscious.

### Study 1: Turning the Other Cheek

In searching databases of the Interuniversity Consortium for Political and Social Research (ICPSR), Association of Religion Data Archives, and Roper Center for Public Opinion, it was surprisingly difficult to find items reflecting Sermon-on-the-Mount beliefs about turning the other cheek and helping and loving one’s enemy. However, we did find one older dataset (that we had worked with in Cohen & Nisbett, 1994) that had items close to this.

The survey also had demographic items on participants’ religion and religiosity, as well as some general questions about ideology and worldview that might help us understand the basis for respondents’ beliefs about turning the other cheek and helping one’s enemies. On one hand, “Turn the other cheek /Love thy enemy” beliefs might derive from an essentially benevolent and humanistic worldview. The seeming inconsistency between helping those we dislike, disapprove of, or who have hurt us might be resolved if it was embedded in a worldview that people are basically good and trustworthy and should be treated with equal respect and dignity. On the other hand, “Love thy enemy” beliefs might also derive more from a religious mandate about how one *should* act and feel. That is, one might love one’s enemies because the Bible says this is the right attitude, not because this attitude genuinely flows from a deeper humanism or belief in human beneficence. Loving one’s

enemies when there are no actual feelings or beliefs to foster doing so does not necessarily imply reaction formation; but it is consistent with hypotheses that counterintuitive and counterinstinctive attitudes about loving one's enemies derive from a belief that one *should* do so rather than an underlying faith in humanity that might genuinely give rise to it.

Note that a "Turn the other cheek /Love thy enemy" attitude could be correlated with *both* a religious mandate and a benevolent worldview. Note also the possible mediation: that the religious mandate could lead to a benevolent worldview that in turn gives rise to "Turn the other cheek" attitudes. It is only when the religious mandate leads to a "Turn the other cheek" attitude and is unrelated to any ideology of benevolence that it seems as if people hold this attitude largely because they *should* do so.

## Method

The studies in this paper were ruled exempt from human subjects review by the University of Illinois Institutional Review Board or were run under IRB 13115 "Cultural Influence on Reaction Formation."

Participants were 1,374 American men (76% white, 22% African American, 2% other) interviewed in 1969 about their attitudes toward violence (Blumenthal et al., 1972).

**Dependent variable of "Turn the other cheek/ Love thy enemy" mentality.** To operationalize the "Turn the other cheek/ Love thy enemy" mentality, we used the 3 items comprising the "Kindness Index" (which were in turn based on Scott, 1965): "When a person harms you, you should turn the other cheek and forgive him," "Even if you don't like a person, you should still try to help him," and "It's important to be kind to people even if they do things you don't believe in." As shorthand, we call this the "Other cheek" index.

### Independent variables.

**Religious group.** Protestants were coded as 0 ( $n = 982$ ), non-Protestants as 1 ( $n = 259$  Catholics and 33 Jews). As elsewhere (Hudson & Cohen, 2016; Kim et al., 2013; Kim & Cohen, 2017), we used Catholics and Jews as our non-Protestant group for a type of "minimal difference" sampling (Cohen, 2007), as Catholicism and Judaism are Protestantism's closest cultural siblings.

The other independent variables were used to explore the basis for the "Other cheek" mentality. As noted, the basis for such a mentality might derive from an essentially benevolent and humanistic worldview. To tap into this, we used the following:

**Humanism.** Blumenthal and colleagues' Humanism Index gave respondents a list of six values and examined the extent to which they valued human dignity, equality and freedom over respect for property, respect for law, and financial security.

**Benevolent worldview.** For this variable, we combined Blumenthal et al.'s Trust Index (with items, such as "Do you think most people would try to take advantage of you if they got a chance or would they try to be fair?"), Blumenthal and colleagues' (reverse scored) Suspicion/Resentment Index (with items such as "I commonly wonder what hidden reason another person may have for doing something nice for me"), and a 3-item composite we created by averaging questions about the trustworthiness of policemen, Black protesters, and white student demonstrators (alpha for the composite from the 3 indices = .60).

On the other hand, an "Other cheek" mentality might essentially be a mentality by edict. That is, it might be a mentality not developed out of a humanistic and benevolent worldview but a mentality held because one *should* hold that mentality. If so, a respondent's "Other cheek" sentiments might be predicted by his religiosity.

**Religiosity.** We used Blumenthal and colleagues' Religiosity Index, derived from religious service attendance (never to several times a week) and participants' response to the question, "In general, how religious minded would you say you are" (not at all to very). Answers to attendance and religiosity items were standardized and averaged. This is the measure of religiosity used throughout the paper (except Study 2, in which religiosity was not measured).<sup>3</sup>

## Results

As seen in Table 1, Protestants were more likely than Catholics and Jews to endorse the "Other cheek" items and were in fact twice as likely to receive the top score on the index created by Blumenthal and colleagues (1978) from these items.

Next, we explored which variables predicted the "other cheek" index among Protestants versus non-Protestants. We regressed the "other cheek" index simultaneously onto humanistic orientation, benevolent worldview, religiosity, as well as religion (Protestant vs. non-Protestant) and the interactions between religion and these three other variables. The endorsement of the "Other cheek" items was predicted by different variables for Protestants versus non-Protestants. For non-Protestants, an "Other cheek" mentality was associated with a humanistic orientation ( $b = .24$ , partial  $r = .09$ ,  $t = 3.16$ ,  $p = .002$ ) and a benevolent worldview ( $b = .23$ , partial  $r = .08$ ,  $t = 2.96$ ,  $p = .003$ ). For Protestants, however, the "Other cheek" mentality was not associated with either humanism ( $b = 0$ ,  $p = .99$ ) or a benevolent worldview ( $b = .03$ ,  $p = .61$ ). Thus, the Religious Group  $\times$  Humanistic Orientation interaction was significant ( $b = .24$ , partial  $r = .08$ ,  $t = 2.76$ ,  $p = .006$ ) as was the Religious Group  $\times$  Benevolent worldview interaction ( $b = .21$ , partial  $r = .07$ ,  $t = 2.30$ ,  $p = .02$ ).

As seen in Figure 1, Protestants and non-Protestants who were 1 *SD* above the mean on both a benevolent worldview and humanistic orientation were equally likely to endorse the "Other cheek" mentality (predicted values = 3.21 for Protestants, 3.18 for non-Protestants). At 1 *SD* below the mean on both these measures, however, differences were quite pronounced. Non-Protestants who were suspicious, resentful, and distrusted others and who did not place much relative importance on others' dignity, equality, and freedom were—as one might expect—relatively low scorers on the "Other cheek" measure (predicted value = 2.24). However, whereas such relatively misanthropic non-Protestants were almost a full point lower on the 5-point "Other cheek" measure, relatively misanthropic Protestants hardly differed from more trusting, more humanistic people of either faith (predicted value = 3.15).

<sup>3</sup> There were small differences across studies. Studies 3 through 5 asked the attendance questions in an open-ended format and we converted these to scale scores. Studies 3 and 5 asked "How religious are you?" from *not at all* to *extremely religious* (scored 1–5). Study 4 asked "Overall, how important is your religion to you and/or how religious or spiritual do you consider yourself?" (from *not at all* to *very*, scored 1–5).

Table 1  
Endorsement of “Turn the Other Cheek” Sentiments by Protestants vs. Catholics and Jews

Group	“Other cheek” index <i>strongly disagree</i> (1) to <i>strongly agree</i> (4)			Percent of respondents with the top score on Blumenthal et al.’s Kindness Index (%)
	“When a person harms you, you should turn the other cheek and forgive him.”	“Even if you don’t like a person, you should still try to help him.”	“It’s important to be kind to people even if they do things you don’t believe in.”	
Protestants	2.42	3.42	3.35	25
Catholics and Jews	2.11	3.32	3.21	12
Group comparison	$F(1, 1266) = 25.86, p = .001, r = .14$	$F(1, 1269) = 6.35, p = .01, r = .07$	$F(1, 1266) = 7.93, p = .005, r = .08$	$\chi^2(1,1273) = 20.07, p = .001$

In contrast, as seen in Figure 2, among Protestants, religiosity strongly predicted the turn the Other cheek mentality ( $b = .38$ , partial  $r = .24$ ,  $t = 8.78$ ,  $p = .001$ ). Among non-Protestants, religiosity had no relation to this mentality ( $b = .015$ ,  $p = .85$ ). The Religious Group  $\times$  Religiosity interaction was thus significant ( $b = -.37$ , partial  $r = -.12$ ,  $t = -4.15$ ,  $p = .001$ ).

Such data do *not* prove that Protestants who placed relatively little emphasis on others’ dignity and equality, were resentful of people, and believed others were trying to take advantage of them were insincere in their professions of loving their enemies. But one wonders why such suspicious and distrustful people would endorse loving their opponents or turning the other cheek. The data imply that among Protestants, holding a love-thy-enemy mentality was a matter of religious mandate—because one *should* hold this mentality—regardless of whether others were untrustworthy and deserving of resentment and regardless of how one felt about their dignity, equality, or freedom.

The data also do *not* imply that people who endorse “Other cheek” items will actually follow this path in their behavior. Offering the other cheek after one has been slapped and loving one’s enemy are hard to do in practice. However, the data show that Protestants are more likely than non-Protestants to endorse this seemingly counterinstinctual mentality, believing it to be a mindset that they should try to achieve.

Such endorsement is not by itself evidence of full-blown reaction formation. It is, however, a sort of proto-reaction formation, a step in the direction of reaction formation, just as the counterinstinctual Antitheses are a step toward “transforming . . . hearts and

becoming more like Jesus not only in action but also in thought and motivation” (Huntsman, 2010, p. 93). For evidence that Protestants show more fully developed reaction formation, we next turn to clinicians’ judgments, experimental procedures, and validated defense mechanism scales.

## Study 2: Clinicians’ Judgments

Study 1 showed that Protestants were more likely than non-Protestants to endorse a mentality of turning the other cheek and loving their enemies. Protestant respondents’ counterinstinctual stance did not derive from a benevolent, humanistic orientation (as it did for non-Protestants endorsing this mentality), but rather seemed to derive from a religious sense that this is what they should do. Of course, like all studies using self-reports from mass surveys, Study 1 faces difficulties of interpretation, due to social desirability, reference group effects (Heine, Buchtel, & Norenzayan, 2008; cf. Oishi & Roth, 2009), and issues of whether people are telling more or less than they know (Nisbett & Wilson, 1977; Shedler et al., 1993; Wilson, 2004).

At the other extreme from mass surveys are studies that involve more probing interviews, using skilled clinicians to carefully observe and make judgments about what is going on under the surface. There are arguments against using clinicians’ assessments, because they can be subjective and affected by the biases of the therapist (and as will be seen, there is some truth to this criticism). Nevertheless, there are also data that skilled clinicians can detect issues that respondents can (intentionally or unintentionally) fake their way past on “objective” instruments (Shedler et al., 1993). Given the methodology of Study 1, it would seem helpful to also

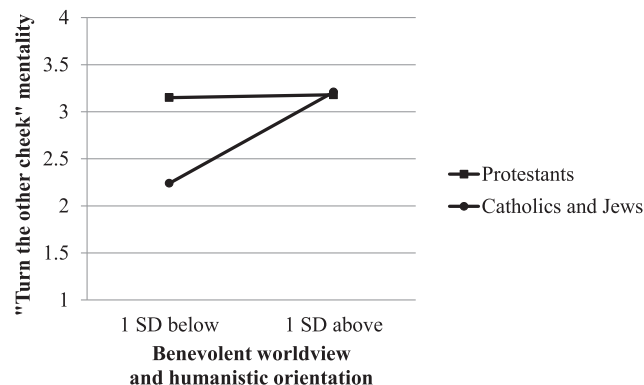


Figure 1. Endorsement of “Turn the other cheek” sentiments as a function of the participants’ religion and whether they had a benevolent worldview and humanistic orientation.

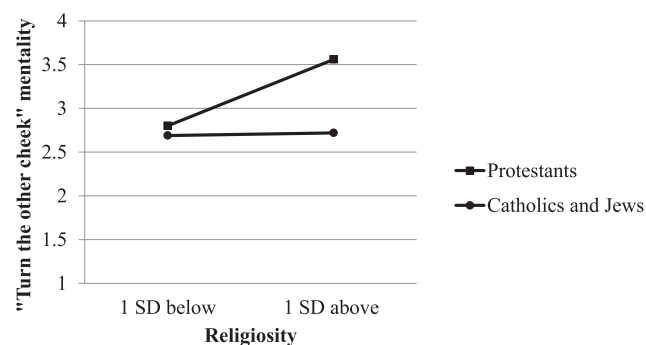


Figure 2. Endorsement of “Turn the other cheek” sentiments as a function of Protestants’ and non-Protestants’ religiosity.

use a method oriented toward more probing evaluations, while still remaining vigilant against problems that can derive from clinicians' subjectivity.

## Method

**Participants and procedures.** Using various keywords, we searched the ICPSR database for studies assessing defense mechanisms (more particularly, reaction formation) and religion. The only publicly available dataset we could find was the Human Aging study (Birren, 1963; Granick & Kleban, 2002), an interdisciplinary project conducted by 22 researchers on healthy men who spent 2 weeks living at the Clinical Center of the National Institutes of Health.

Participants were 47 men over age 65, recruited for a study involving a battery of medical and psychological tests. Included were 3 interviews conducted by 2 psychiatrists, who rated each participant on 10 defenses. Each of the 3 interviews was 2–3 hours long. Topics included: personal history (ex., early family life, parents, sex, marriage, child rearing, work, and psychiatric symptoms), old age, planning for the future, fears, death, grief, anger and guilt over others' deaths, self-concepts, how one would be described by a spouse or one's enemies, age-related changes in one's physical and mental abilities, sex life, ideals, influences, regrets, feelings toward oneself, best and worst things that happened in one's life, life motto, explanations for success, advice one would give others. Interviews ended with a "free period: to pursue areas developed in previous interviews, check out psychodynamic hypotheses, etc." (Birren, 1963, p. 198). Thus, topics covered a number of potentially sensitive areas, giving psychiatrists a good opportunity to assess defenses over the 6–9 hours of interview and conversation.

The two psychiatrists took turns as interviewer and as an "observer-auditor" watching behind a one-way mirror. In addition, three other psychiatrists served as "independent consultants" (Birren, 1963, pp. 161, 175). They each reviewed about 1/3 of the transcribed interviews and contributed their impressions to the two psychiatrists, who made the final ratings.

Of the 47 men, 23 were Protestants, 4 Catholics, and 20 Jews, as noted in the write-up (Birren, 1963, p. 8). The codebook only lists codes for whether the participant was Christian or Jewish. However, based on other information in the codebook, we can make a reasonably good guess about who the Catholics were. Of the Christian participants, there were only 4 foreign-born and—based on the dates of their immigration—approximately three of four were probably Catholic.<sup>4</sup> For the sake of "cleaner" categories, we included only native-born Christians in the Protestant group and compared them to the Jewish group.

Clinicians rated participants on their use of the following defense mechanisms<sup>5</sup>: reaction formation, denial, displacement, isolation, undoing, projection, rationalization, regression, repression, and suppression. The correlation between the 2 clinicians was .18 for reaction formation and an average of .3 for the other 9 defenses, lending credence to the critique that clinicians' assessments have a fair amount of subjectivity. However, there is no particular reason to expect bias; and these correlations are not dramatically different from, say, interobserver correlations for easy-to-observe personality traits such as the Big 5, even when observers know the targets well (average  $r$  of .32 when observers are cohabitators, "long

term acquaintances," work colleagues, friends, and family members, according to meta-analyses from Connelly & Ones, 2010, and Kenny et al., 1994).

## Results

As seen by ratings for the 10 defenses in Table 2, Protestants were significantly greater than Jews on only 1 defense—reaction formation (Protestant  $M = 5.46$ , Jewish  $M = 4.65$ ,  $F(1, 41) = 5.05$ ,  $p = .03$ ,  $r = .33$ , 95% CI [.03, .57]). There were no other significant differences in this (relatively small) sample (all  $r$ s between  $-.23$  and  $.20^6$ ), with the exception that Jews were rated as more likely to employ regression than Protestants were ( $p = .03$ ,  $r = -.34$ , 95% CI [.04, .58]), whereas Protestants were marginally more likely to employ isolation ( $p = .09$ ,  $r = .26$ , 95% CI [-.04, .52]). As seen by the raw numbers, reaction formation (along with repression) was the defense on which Protestants scored highest, whereas it was 6th highest among the Jews.

**Further analyses.** The low interrater reliabilities of the defenses led us to explore other ways of looking at the reaction formation ratings. For example, if we examine only cases that raters could agree on (that is, both raters scored the person as above their median or both scored the person as below their median on reaction formation), then more Protestants than Jews would be scored as high in reaction formation,  $\chi^2(1, 25) = 6.58$ ,  $p = .01$ ,  $\phi = .58$ . If we adopted a looser criterion and considered a participant as showing reaction formation if *either* rater gave him a score above the median, more Protestants than Jews would again be characterized as showing reaction formation,  $\chi^2(1, 43) = 8.58$ ,  $p = .003$ ,  $\phi = .45$ . As a compromise approach, one can also score respondents as 1 (both judges agreed the person showed reaction formation),  $-1$  (both agreed the person did not show reaction formation), or 0 (judges disagreed). Protestant also would be rated as higher on reaction formation than Jews on this measure,  $t(44) = 2.92$ ,  $p = .006$ . Ultimately, though, none of these solutions replace the better, but in this case, impossible solution of getting more ratings on these participants. Lacking such data, we suspect that the simple average of the two clinicians' ratings we have is likely the best indicator.

**Summary.** The only defense Protestants used significantly more than Jews was reaction formation, and it (along with repres-

<sup>4</sup> For the decades when these men immigrated (between 1900 and 1930), about 80% of Christian immigrants to the U.S. were likely Catholic or Orthodox (based on country-of-origin data from Department of Homeland Security, 2004). In this dataset, these 4 "probably Catholic" immigrants had reaction formation scores ( $M = 4.75$ ) similar to those of Jews (4.65). Thus, the  $p$  level for the Protestant versus non-Protestant difference decreased slightly if one puts the "probably Catholic" participants with the Jews ( $p = .02$ ). Of the Jewish seniors, 15 were foreign born. This was expected, because the great Jewish immigration to the U.S. occurred primarily between 1881 and 1924.

<sup>5</sup> At 5- and 11-year follow-ups, defenses were again rated, though sample sizes dropped dramatically and only one of the clinicians was able to make ratings for all 3 periods. Results are virtually unchanged if one averages in later ratings for participants who were in the follow-up (Protestant  $M = 5.40$ , Jewish  $M = 4.67$ ,  $F(1, 41) = 4.91$ ,  $p = .03$ ).

<sup>6</sup> Given the small sample, it's hard to know whether groups differ with respect to the other defenses, and this study was underpowered to detect such associations. Nevertheless, even with a small sample, we did find evidence of group differences in reaction formation.

Table 2  
*Defense Mechanisms Used by Protestant and Jewish Men, as Rated by Clinicians*

Defense	Jewish men	Protestant men	<i>p</i> value
Reaction formation	4.65	5.46	.030*
Denial	3.98	3.83	.694
Displacement	4.1	4.22	.741
Isolation	3.95	4.7	.087 <sup>^</sup>
Projection	4.8	5.13	.395
Rationalization	4.93	5.43	.197
Regression	4.7	3.3	.028*
Repression	5.2	5.46	.483
Suppression	5.75	5.04	.132
Undoing	3.93	4.04	.795

sion) was the most commonly used defense among Protestants.<sup>7</sup> The study, however, does little to dispel doubts of skeptics who regard clinical ratings as too subjective. It further gives little insight into the content of the reaction formation: we do not know whether attraction was replaced by repulsion, repulsion was replaced by attraction, or both. To examine these issues and to induce reaction formation in the laboratory, we conducted experimental studies in two different areas: (a) unconventional sexual desires and (b) more “hostile” feelings (of repulsion, negativity, and disgust) toward African Americans. In the first experiment, we explored whether participants induced to be attracted to taboo sexual behaviors instead react with repulsion; in the second, we explored whether participants induced to hold taboo repulsions (toward African Americans) instead react with attraction.

### Study 3a. Taboo Attractions: Male Participants

Sexual desires are subject to a number of taboos, with unconventional attractions judged as somehow perverted, immoral, disgusting, or depraved. In this study, we tried to induce participants to temporarily feel attracted to various types of unconventional sexual activities. Though it *might* be possible to do this chemically (Hiller, 2004), it is not necessary to do so. Participants can be induced to feel drawn toward various sexual activities that they are not normally attracted to by putting them in a state of sexual arousal. This can be done rather easily in the short term for male participants (Ariely & Loewenstein, 2006).

The present study is modeled most directly after Ariely and Loewenstein (2006) and, to a lesser degree, Loewenstein, Nagin, and Paternoster (1997). In the latter, men recruited from criminology classes and around University of Maryland’s campus became more likely to say they would “verbally coax” a date into letting them remove her clothes, after first having looked at *Playboy* nudes (vs. control photographs). In the former study, men at University of California-Berkeley became more likely to say they would enjoy a variety of conventional and unconventional acts when they answered while masturbating themselves into a “high but sub-orgasmic” state, as compared with while they were not doing so (Ariely & Loewenstein, 2006, p. 89). Thus, for example, unaroused men gave a mean of 19 (on a 0 = *no* to 100 = *yes* scale) to having a threesome with a woman and another man. For aroused men, the mean was 34.

In the present study, men were run in the lab (rather than at home as in Ariely & Loewenstein, 2006), and there was no

masturbation. However, under the guise of a study on advertisements and their effectiveness for people of different personality types, participants were exposed to a series of either erotic or nonerotic perfume commercials and answered a number of questions during and after the video commercials. These questions included items measuring participants’ attitudes and inclinations regarding conventional and unconventional/“deviant” behavior, as well as some individual differences noted below.

Given that the erotic video should put participants in an aroused, appetitive state, we would expect that, *all other things equal*, the erotic video would lead to greater attraction to both conventional and unconventional sexual behaviors (as in Ariely & Loewenstein, 2006). Evidence of reaction formation occurring would be shown to the extent that the erotic video induced not attraction but repulsion to the unconventional behaviors. We expected Protestant men to be more likely to show such reaction formation, as compared with their Catholic and Jewish counterparts. (Conventional behaviors are not taboo and thus we would expect the erotic video to increase the attractiveness of conventional sexual behavior for participants from all groups).<sup>8</sup>

Finally, we also have an exploratory measure of self-control (a handgrip task) often used to assess “ego-depletion” (Baumeister & Tierney, 2012). To the extent that reaction formation requires effortful use of ego resources, we would expect that Protestants in

<sup>7</sup> Though the Healthy Aging study was the only dataset we could find and analyze, we did find one other reference to possible ethnic/religious differences in defense mechanisms. Vaillant (1993) analyzed data from a Boston sample of 500 “nondelinquent” White male youths, who had been matched as controls to 500 “delinquent” youths. Of the 500 nondelinquent youths, 307 appeared to have been rated on their defenses at age 47. Vaillant then compared Italians ( $n = 74$ ) with WASPs ( $n = 100$ ), defining the latter as “Men of old American or English or Anglophone Canadian descent” (p. 139). Though religion data were collected, religion was not included in the analysis of defenses and perhaps not in the definition of WASPs. Defined as above, it is highly unlikely that the WASPs were exclusively Protestant. Of the original sample of 500, 33% were of “old American or English or Anglophone Canadian descent” (Glueck & Glueck, 1950). However, of this same 500, only 23.6% were Protestant. Even assuming there were *no* other Protestants besides those of American/English/Canadian descent, this implies that *at most* 72% of the “WASP” sample was Protestant. Additionally, of the original sample, 71.6% were Catholic, though Italians made up only 24% of the sample, implying that non-Italian Catholics were twice as prevalent as Italian Catholics in the sample; data from these non-Italian Catholics did not seem to be analyzed. For the Boston study, interviewers created a 20–30 page summary of a 2-hr semistructured interview, later scored by 2 raters for 15 defenses. The interviewer was not identified, but the 2 raters were “a medical social worker and a recent college graduate in psychology,” who were not psychoanalytically trained (Vaillant, 1993, p. 130). Table 8 of Vaillant (1993) provides percentages of respondents using a given defense “as a major style.” We hesitate to include these analyses primarily because the comparison does not seem to be based on religion but rather ethnicity. Nevertheless, the Vaillant study found little in the way of defensive differences between the 2 groups. “WASPs” had higher rates of reaction formation (14%) than Italians did (8%) but the difference was not significant,  $z = 1.28$ ,  $p = .20$  (contrast on proportions). Meta-analyzing Vaillant’s results with those of Study 2, we found that Protestants were more likely to be rated as engaging in reaction formation, as compared with their non-Protestant counterparts,  $Z = 2.50$ ,  $p = .012$  (unweighted by study  $n$ ) or  $Z = 1.78$ ,  $p = .075$  (weighted by study  $n$ ).

<sup>8</sup> Ariely and Loewenstein (2006) did not report reaction formation in response to their unconventional items. However, there were likely self-selection pressures that screened out the more prudish participants. The 35 Berkeley undergraduates were “recruited with ads placed around campus . . . Before the experimental session, subjects were informed about the experiment, including the fact that it would involve masturbation” (p. 90).



the erotic video condition should show the greatest ego depletion, indexed by shorter handgrip times. To the extent that reaction formation was completely “automatic,” it should not lead to ego depletion.

A note on participants: For this first study, we used men rather than women for a few reasons. First, men were used in previous research with similar paradigms (Ariely & Loewenstein, 2006; Loewenstein et al., 1997). Second, it is much easier to induce sexual arousal in men using visual stimuli (Ellis & Symons, 1990). Third, the psychology of arousal is much simpler in men; the connections between attraction, arousal, and physiology are thought to be more complicated in women (as noted in the introduction to Study 3b, which involved women).

## Method

**Procedure.** One hundred eighty-four students (51 Protestant, 133 Catholic and Jewish men) were run individually. As elsewhere (Hudson & Cohen, 2016; Kim et al., 2013; Kim & Cohen, 2017) and as in Studies 4 and 5, students were prescreened and had to identify their current religion and the religion they grew up in as Protestant, Catholic, or Jewish.

The experimenter explained that the study explored the following:

personality factors relevant to effective advertising. . . . Why does an advertisement work for one group of people while the same advertisement does nothing to another group of people? The effectiveness of the advertisement probably depends on our personality and the kinds of thoughts and feelings we typically have. In today’s study, we are going to ask you about your attitudes towards certain consumer products before and after showing you some video advertisements related to them. In the later part of the study, we are also going to give you various measures of personality and attitudes to find out which aspects of your personality and attitudes influenced how you reacted to the advertisements that were shown to you.

As part of the cover story, participants were given several rolls of paper towels and asked to rate them. They were then shown commercials for paper towels, with attitude questions for participants to answer appearing at the top of the screen as the ads played. Participants watched the ads alone in the dark on a computer, listening to the videos through headphones the experimenter provided. The headphones, darkness, and privacy were to help participants immerse themselves in the videos, which we thought would be crucial for the erotic/nonerotic perfume videos that would come next. After the video, subjects made new paper towel ratings.

When the experimenter returned, she told participants they would next rate women’s perfumes. She gave them perfume samples and two minutes to make an initial set of perfume ratings. After the ratings, the experimenter told participants that, the series of “ads for this set of products are more sensitive than the ones you saw at the beginning of the study. For this reason, the questions that will be asked alongside the ads will touch on more sensitive topics as well.” She reassured them “that all of your information is completely confidential, so please try to be as open and honest as possible, reflecting how you are feeling at the moment as you answer these questions.” She then started the series of perfume ads, turned the lights out, and left the room.

**Erotic versus nonerotic video manipulation.** The main experimental manipulation was introduced with the video, which we created to be either extremely erotic or not erotic. (Perfume ads, particularly on the Internet, can have strong sexual content, so erotic ads do not look strange for the genre. Conveniently, one of the perfume makers also has a lingerie brand, so some of their lingerie ads could be spliced in and presented as if they were perfume ads). Both the erotic and nonerotic videos were edited to be about 9.5 min long.

As the video played, questions again appeared at the top of the screen. (Putting questions there made sure participants were always looking at the screen). Some questions concerned rather conventional, “vanilla” preferences: ex., “Would it be exciting to have a romantic partner who was extremely physically fit?” “Would it be exciting to have intercourse with your sexual partner?” “Would it be exciting to passionately kiss your partner?” and so on. Most of the questions, however, were about unconventional desires, many adapted from Ariely and Loewenstein (2006) but some added by us, with a few of Ariely and Loewenstein’s (2006) date rape questions as well: ex., “Would it be fun to lick your sexual partner’s feet?”, “Would it be fun to have your sexual partner lick your feet?”, “Would it be exciting to have your partner dress you up as a woman during a role play fantasy?”, “Would you keep trying to have sex after your date says ‘no,’” “Would it be exciting to watch someone else have sexual intercourse with your partner?”, “Would it be fun to tie up your sexual partner?” and so on.<sup>9</sup> There was a pool of 10 conventional questions and 32 unconventional ones, with half randomly assigned to be asked as the video played and half randomly assigned to be asked later in the study.<sup>10</sup>

If the erotic video generated arousal but no reaction formation, we would expect the appeal of both conventional and unconventional practices would be enhanced by the video (as in Ariely & Loewenstein, 2006). To the extent that participants experienced reaction formation, whereas we would expect to see increased endorsement of “normal” sexual practices, we would expect to see increased *repulsion* from unconventional practices in the erotic video condition.

After the video, the experimenter returned, and participants rerated the perfumes (continuing the cover story). They were then given 15 min to complete a set of personality scales (see below), and after that, proceeded to two tasks whose order was randomly assigned.

One of the tasks was merely the other half of the questions about interest in conventional/unconventional sexual behavior, along with some questions asking about respondents’ general attitude

<sup>9</sup> Conventional and unconventional items were chosen a priori; confirming our classification, conventional items were much more likely to be endorsed, compared with unconventional ones,  $F(1, 175) = 2126.93, p = .001$ . Performing a factor analysis and forcing a 2-factor solution gives a solution extremely similar to our a priori classification. Dropping items that did not load correctly according to the factor analysis gives results similar to that in the text (Religious Group  $\times$  Erotic/nonerotic video interaction  $p = .006$ ).

<sup>10</sup> We were unsure how long it would take for the reaction formation to occur, thinking that some delay might be required. It turned out that the key effect was similar regardless of whether questions were asked during the video or after ( $p = .50$  for Religious Group  $\times$  Erotic/nonerotic  $\times$  During/after video interaction predicting unconventional desires).

toward vices and their opinion about how various deviant acts should be punished. The attitudes questions included items about sex (ex. "There is way too much pornography in today's society"), smoking (ex. "I generally think less of people who smoke"), drinking (ex. "Society would be better off if everybody drank less"), and general moral decline (ex. "We would be better off if more people were God fearing"). The punishment questions asked participants to indicate the length of time they would punish offenders in 5 vignettes that included crimes such as third-degree homicide (killing a pedestrian while drunk driving), statutory rape between a gay man and his boyfriend, public indecency for anonymous sex between 2 gay men at a highway reststop, grand theft auto of 2 cars, and statutory rape with a 21-year-old having sex with his awestruck 17-year old girlfriend as her "birthday present."<sup>11</sup> At the end of the questionnaires, participants completed a few demographic items, including one asking about their sexual preference from (1) *completely homosexual* to (5) *completely heterosexual*. Analyses below only include participants scoring 3 or higher, because the erotic video used stimuli aimed at those with a heterosexual preference.

The other task (randomly assigned to proceed or follow the questions about unconventional/deviant behavior) was a handgrip task in which participants were told to squeeze a handgrip as long as possible so that a small collection of paper did not slip out. The task was randomly assigned to be described as a test of self-control or altruism, with participants in the altruism condition informed that for every 15 seconds they held the handgrip, the research team would donate \$1 to a charity of their choice (from a list provided).<sup>12</sup> Participants had a practice trial and then two actual handgrip trials that we averaged.

**Summary of dependent measures.** The primary measure was the participant's approval or condemnation of unconventional and deviant acts. This was a composite of the four indicators described above: participants' ratings for unconventional sexual behaviors during the video (16-item  $\alpha = .80$ ), ratings for such behaviors after the video (16-item  $\alpha = .80$ ), attitudes toward vice (15-item  $\alpha = .85$ ), and punishments for criminal acts (5-item  $\alpha = .58$ ).<sup>13</sup> Given their length, each of the 4 indicators had acceptable reliability and the overall measure comprised of the 4 indices did as well (4-item  $\alpha = .58$ ). The index was itself standardized.

Again, for this primary measure, if there were no countervailing defensive reaction occurring, we would expect the erotic video to make participants find both conventional and unconventional sexual behaviors more appealing. To the extent that participants find attraction to unconventional acts discomfiting and need to reverse these feelings of attraction, we would expect to find participants in the erotic video condition repulsed by the unconventional behaviors (even as they acknowledged increased attraction to conventional behaviors). The prediction that Protestants induced by the erotic video to feel unacceptable desires would exhibit reaction formation and show repulsion should lead to a Religious Group  $\times$  Erotic versus nonerotic video interaction in condemnation of unconventional behavior. For Catholics and Jews, the prediction was that reaction formation would not be exhibited, and that the erotic video would lead to greater attraction to all sexual activities, increasing their appetite for sexual behavior regardless of its unconventionality.

**Handgrip Task.** The handgrip task was used as a measure of ego depletion.

**Religiosity as a potential moderator.** As in Study 1, religiosity was a composite of religious service attendance and self-described religiousness.

**Individual difference measures.** We explored whether reaction formation might be predicted by conventional personality measures. As noted, the personality measures were given after the second set of perfume ratings and included: 3 of the big 5 (agreeableness, neuroticism, openness), the BIS-BAS scale, cognitive rigidity (measured by a sampling of items from dogmatism and F-scales), Need for Closure, social desirability, and anxiety (Taylor Manifest Anxiety short form). If reaction formation is "just" a manifestation of normal personality processes already tapped by these individual differences, we would expect to see Individual difference  $\times$  Erotic/nonerotic video interactions, with those scoring high on some dimension becoming more condemning toward unconventional practices in the erotic video condition.

Disapproval or a lack of attraction to unconventional practices is not *by itself* evidence of defensiveness, because people can vary in their taste for unconventional acts just as they vary in their taste for stock car racing, hummus, and folk dancing. Without some *other* validated measure of defensiveness, preferences by themselves are not evidence of defensive reactions. The key is the interaction involving the Erotic/nonerotic video condition, in which the manipulation that should make acts more appealing instead makes them more repellent.

## Results

We begin with the conventional sexual practices; the erotic video should increase the desire for conventional sexual behavior among participants of *all* groups. As expected, participants in the Erotic video condition did indeed become (marginally) more desirous of conventional sexual activity ( $M$ s for the standardized conventional sexual behavior index in the erotic condition = .14,  $SD = .99$  vs. nonerotic condition =  $-.14$ ,  $SD = .99$ ,  $F(1, 156) =$

<sup>11</sup> Respondents indicated the number of months of prison they would give in open-ended format, and some of the answers were plainly not real answers (ex. a sentence of 1,000,000 months). Answers were considered missing if they were outside Tukey's "inner fences" – that is, if they were more than 1.5\*interquartile range above the 75th percentile or below the 25th percentile. If one includes the observations but transforms sentences by creating the reciprocal, the Religious Group  $\times$  Erotic/nonerotic video interaction predicting unconventional desires remains similar to that in the text ( $p = .016$ ). In addition to the questions about vices and moral decline, there were also questions about respondents' attitude toward fat people, as some have argued that fat is becoming a moralized issue (Crandall & Biernat, 1990; Rozin, Kurzer, & Cohen, 2002). However, the fat items lowered reliability of the scale and thus we restricted our vices to traditional ones of smoking, drinking, and sex (Vandello & Cohen, 2016).

<sup>12</sup> Though participants did not know this, we cut them off after a certain amount of time. Initially, the cut-off was 90 seconds, but too many participants were reaching this limit. We increased the limit to 120 seconds, but still had several participants reach this ceiling. Handgrip data from participants with the 90-s limit were dropped; additionally, there were some experimental glitches that led to missing data for about 10% of participants.

<sup>13</sup> The indicators for deviant/unconventional acts thus included some nonsexual behaviors that involved vices or were criminally deviant but not sexually deviant. If one separates the sexual from the nonsexual items, the Religious Group  $\times$  Erotic/nonerotic interaction for the sexual items was  $p = .025$ ; for the nonsexual items, the effect was in the same direction but weaker (interaction  $p = .37$ ).

2.83,  $p = .095$ ). There was no Religious Group  $\times$  Erotic versus nonerotic video interaction as both groups became more interested in conventional sexual activity to an equal degree (interaction  $F = .002$ ,  $p = .97$ ). There were no other meaningful effects, except for an Erotic versus nonerotic  $\times$  self-control versus altruism  $\times$  Handgrip first versus last interaction ( $p = .02$ ) that we had no a priori prediction for and hesitate to interpret (all other  $ps > .27$  in Religious Group  $\times$  Erotic vs. nonerotic  $\times$  Self-control vs. altruism  $\times$  Handgrip before vs. after ANOVA).

The more interesting analysis involves the endorsement or rejection of unconventional behavior. Here we expected a Religious Group  $\times$  Erotic versus nonerotic video interaction, with the erotic video eliciting attraction to unconventional sexual behavior among non-Protestants but inducing a defensive, repulsive reaction among Protestant participants. The interaction occurred (Figure 3,  $p = .008$ ,  $F(1, 160) = 7.29$ ): non-Protestants who watched the erotic video showed a relatively greater endorsement of unconventional/deviant behavior, whereas Protestants who watched the erotic video showed a relatively greater repulsion from such behavior.

Decomposing this interaction, we found that seeing the erotic video, non-Protestants were significantly more positive about the unconventional/deviant behavior than Protestants were,  $t(89) = 2.81$ ,  $p = .006$  (Protestant  $M = -.43$  vs. non-Protestant  $M = .22$ ). However, there was no significant difference between the groups when the nonerotic video was shown ( $t = -1.29$ ; Protestant  $M = .12$  vs. non-Protestant  $M = -.10$ ).

Decomposing the interaction another way, we found that Protestants shown the erotic video became repulsed by unconventional/deviant behavior, as compared with Protestants shown the nonerotic video,  $t(47) = 1.93$ ,  $p = .06$ . In contrast, non-Protestants shown the erotic video became more attracted to the unconventional/deviant behavior, as compared with those shown the nonerotic video,  $t(127) = 1.84$ ,  $p = .068$ , replicating findings of Ariely & Loewenstein.

In the Religious Group  $\times$  Erotic versus nonerotic  $\times$  Self-control versus altruism  $\times$  Handgrip before versus after ANOVA, there were no other significant effects predicting reaction to unconventional/deviant behavior (all other  $ps > .123$ ).

**Further analyses.** Conducting a Religious Group  $\times$  Erotic versus nonerotic  $\times$  Conventional versus unconventional practices

(within-subjects) ANOVA confirms the difference in patterning for conventional versus unconventional behavior (3-way interaction,  $F(1, 168) = 4.37$ ,  $p = .038$ ). Decomposing this interaction, the erotic video makes all participants more interested in conventional sexual behavior, but has divergent effects on Protestants versus non-Protestants for unconventional behavior (simple interaction of Religious Group  $\times$  Conventional vs. unconventional behavior,  $t(86) = 2.03$ ,  $p = .045$  in the erotic video condition). In the nonerotic video, no such interaction occurred,  $t(84) = -.92$ .

An alternate way to decompose this 3-way interaction is to say that, for Catholics and Jews, there was no 2-way interaction of Erotic versus nonerotic  $\times$  Conventional versus unconventional behaviors,  $t(123) = .27$  – only an effect of erotic versus nonerotic video (simple effect  $t(123) = 3.22$ ,  $p = .002$ ) as the erotic video piqued their interest in sexual activity, regardless of how conventional or unconventional it was. On the other hand, for Protestants, the 2-way interaction emerged,  $t(44) = 2.30$ ,  $p = .026$ , with the erotic video making conventional behavior more attractive to them and unconventional behavior more repellent to them.

**Self-control resources depleted.** On the exploratory measure of the handgrip task, the 2-way interaction of Religious Group  $\times$  Erotic versus nonerotic video was not significant,  $p = .23$ .

**Moderation by religiosity.** Religiosity did not moderate the interaction of interest ( $t = -.40$ ,  $p = .69$  for 3-way interaction of Religious Group  $\times$  Religiosity  $\times$  Erotic vs. nonerotic video).

**Other individual difference variables.** We explored whether agreeableness, neuroticism, openness, BIS, BAS, authoritarianism, need for closure, social desirability, anxiety, and being a repressor (high social desirability, low anxiety) interacted with Erotic versus nonerotic video condition. None did. There were zero-order effects: those low on need for closure ( $r = -.18$ ,  $p = .015$ ), authoritarianism ( $r = -.13$ ,  $p = .088$ ), and social desirability ( $r = -.15$ ,  $p = .044$ ); those high on BAS ( $r = .16$ ,  $p = .03$ ) and openness ( $r = .21$ ,  $p = .006$ ); and nonrepressors ( $r = -.13$ ,  $p = .08$ ) were more favorable toward unconventional/deviant acts. If this had not been true, it would likely signal that there was a problem. However, none of these variables significantly interacted with whether participants watched an Erotic versus nonerotic video and none moderated the Religious Group  $\times$  Erotic versus nonerotic video interaction. Overall, there was little evidence that traditional individual difference measures predicted reaction formation, in which sexual arousal caused participants to become more repelled by (as opposed to attracted to) unconventional behaviors.

**Reaffirming “normality.”** As noted, only bisexual or heterosexual men were included in analyses above. We asked participants about their sexuality using a scale of 1 (*completely homosexual*) to 5 (*completely heterosexual*), following up with a question asking “how sure are you of your answer to the previous question (regarding sexual orientation)” (1 = *not at all sure*, 5 = *completely sure*). If Protestant participants in the erotic video condition were trying to reaffirm their “normality,” we would expect them to be particularly likely to say they were completely sure that they were completely heterosexual in this condition. The Religious Group  $\times$  Erotic versus nonerotic interaction was not significant, however ( $p = .38$ ).

**Summary.** Reaction formation seemed to be induced among Protestant male participants. The erotic video made all participants more endorsing of conventional sexual behaviors. However, when

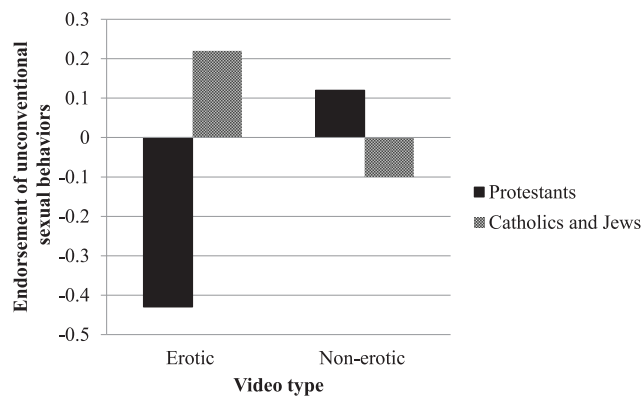


Figure 3. Participants' endorsement of unconventional sexual behaviors as a function of their religion and whether they watched an erotic versus nonerotic video.

it came to unconventional behaviors, Protestants versus Catholics and Jews diverged. Consistent with Ariely and Loewenstein (2006), Catholic and Jewish males in the erotic video condition also became more endorsing of unconventional sexual behavior. For Protestant participants, the opposite effect occurred: Whereas the erotic video “should” have put them in a state that made unconventional sexual behaviors attractive to them, the erotic video actually made Protestants more repulsed by such behaviors.

### Study 3b. Taboo Attractions: Female Participants

As noted in Study 3a, men are the simplest place to start. It is relatively easy to arouse them, a paradigm similar to the present one has been shown to work with men (Ariely & Loewenstein, 2006), and the connections between physiological arousal, subjective desire, liking, and wanting are relatively straightforward. Things are less simple with women. Relative to men, women are less easily aroused by visual stimuli. Physiological arousal is also relatively nonspecific; and, as measured by “objective” instruments, it is not *strongly* correlated with either self-perceptions of physiological arousal or with subjective feelings of arousal in women (Chivers, Rieger, Latty, & Bailey, 2004). And when physiology and subjective arousal are correlated, there is some suggestion that causal processes are more cognitively mediated in women—“the data suggest that it is the desire or subjective arousal that cues them to recognize the physical responses rather than the other way around” (Meana, 2010, p. 106). Further, Meana (2010, p. 116) argues that the relation between liking something and wanting it may be “calibrated” differently for men and women, with even relatively low amounts of liking triggering wanting more easily in men. Nevertheless, for Study 3b, we created different stimuli for (heterosexual) women and examined whether our findings would replicate with a female sample.

### Method

Methods were similar to Study 3a. We needed to alter stimuli to create a video we thought would be more appealing to women. The majority of questions comprising the dependent variable were retained, though we replaced some of the more male-centric items (ex. “Are women’s shoes erotic?”) with other items. There were 166 participants (53 Protestant, 113 Catholic or Jewish) run individually.

### Results

Again, we begin with conventional sexual practices. If the erotic video had the intended effect, it should increase desire for conventional sexual behavior among participants of *all* groups. There was, however, no significant effect of the erotic video ( $M$  for erotic video = .07,  $M$  for nonerotic video = -.07,  $F(1, 132) = .90$ ,  $p = .345$ ). In the Religious Group  $\times$  Erotic versus nonerotic  $\times$  Self-control versus altruism  $\times$  Handgrip before versus after ANOVA, there were in fact no significant effects of any kind (all  $ps > .125$ ). For the analysis of unconventional behavior, there were also no significant effects, including no main effect of Erotic versus nonerotic video ( $p = .145$ ) and no Religious Group  $\times$  Erotic versus nonerotic interaction ( $p = .411$ ; all other  $ps > .17$ ). On the handgrip task, there were also no effects (all  $ps > .16$ ).

### Discussion

We did not replicate the reaction formation effect among female Protestant participants. It is unclear whether this was attributable to (a) Protestant women being less susceptible to reaction formation, (b) Protestant women being less susceptible *in this particular paradigm*, or (c) some failure of our manipulation to produce the intended psychological state that has to be defended against. Given that the erotic video did not produce evidence of greater attraction to either conventional or unconventional behaviors, we suspect the last explanation may parsimoniously account for the failure.

We turn next to a different paradigm where we again use female and male participants. However, in this paradigm, we do not examine whether participants turn forbidden attraction into repulsion but instead whether they turn forbidden repulsions into attraction.

### Study 4: Taboo Feelings: Changing Repulsion Into Attraction

Study 4 examined a feeling that is taboo among many students on relatively liberal college campuses. In this study, we attempted to temporarily induce revulsion toward African American targets, predicting that these inductions would lead Protestants to engage in reaction formation against such taboo feelings. Thus, evaluatively conditioning Protestants to feel negatively toward African American targets should not produce repulsion (as it should among non-Protestants), but instead should produce positive feelings toward African Americans.

### Method

Participants were 170 Protestants and non-Protestants (83 women). Because of a shortage of eligible Judeo-Christian participants, we expanded our sample of non-Protestants to include not just Catholic and Jewish participants ( $n = 65$ ) but non-Protestants of other religions ( $n = 25$ ) as well as atheists, agnostics, and people with no religious affiliation ( $n = 35$ ). We collapse all non-Protestant groups together, but as will be seen, results are similar if the sample is limited to Catholic and Jewish non-Protestants. Because of a shortage of participants willing to sign up for lab studies, as the semester wore on, we also opened the study to online participants ( $n = 110$ ) from the psychology pool. By ethnicity, 119 participants were white, 45 Asian American, 10 Latino, 1 other (participants could check more than 1 ethnicity). African Americans were excluded for obvious reasons.

**Procedure.** Participants were given the cover-story that this experiment was about “psychology and economics,” and that they would “complete a number of tasks involving the subject of money, economic decision making, and financial behavior.”

**Evaluative conditioning manipulation.** The conditioning manipulation (Olson & Fazio, 2006) was embedded in what was described as an “attention task” designed to measure participants’ attentional and rapid-responding abilities. Participants were told they would view a “slideshow” containing wealth-related targets (e.g., a picture of a diamond, the word “diamond”). Their goal was to click target images/words as quickly as possible, while not clicking any nontargets. Each screen had either a pair of images/words or a single image/word.

The slideshow had 5 blocks of 33 slides each. Twenty-five blank “slides” were randomly mixed in each block to reduce any sense of rhythm in the presentation. Each slide was shown for 1.50 seconds. Of 33 slides, 8 were wealth-related target images/words that participants were to click, sometimes paired with a neutral image/word; 13 were neutral images/words, presented alone or in pairs; and 12 were the critical unconditioned stimuli (US)/conditioned stimuli (CS) pairs. Following Olson and Fazio (2006), the CSs were 8 pictures of Whites and 8 pictures of African Americans performing various careers, roughly matched on prestige. The USs were a total of 30 positive images/words and 30 negative images/words.<sup>14</sup> Though each CS was shown a total of 4 times across the 5 blocks, each US was shown only once throughout the experiment.

In Olson and Fazio (2006), participants in the experimental group saw pictures of African Americans paired with positive stimuli to create positive associations with African Americans. In the present study, participants in the experimental group saw pictures of African Americans paired with negative stimuli. For the experimental group, there was a 100% contingency between Black CSs and negative USs (30 pairings); there was also a 100% contingency between White CSs and positive USs (30 pairings). For the control group, CSs were exclusively paired with neutral words and images, whereas USs were paired with each other or with neutral words/images not containing people.

**Dependent measures.** Our dependent measure allowed participants to express favor toward African Americans versus Whites in a variety of contexts on tasks not obviously about race.

**Index for favoring African Americans relative to Whites.** This index was computed based on the 4 tasks below:

**White versus African American financial planner (Between-persons).** Participants were told to imagine they had “a job that pays over \$100,000 per year and are searching for a Certified Financial Planner to help [them] manage [their] investment portfolio.” They then reviewed a professional-looking website for a financial planner with average credentials. As a between-subjects manipulation, the website’s main banner contained a picture of either an African American or White man, along with pictures of him and his wife at the bottom of the page. (The second pictures made the man always visible on screen).

Participants (a) rated how likely they would be to hire the financial planner on a scale from 0–100%, (b) indicated how much they would be willing to pay the planner per hour [free response in a field accepting numbers between 0 and 999], and (c) rated the planner on several personal qualities, using a slider from *not at all* (0) to *extremely* (100). Personal qualities included: caring, trustworthy, intelligent, hardworking, and so on, and were averaged together. All three measures (likelihood to hire, salary, and personal qualities) were standardized and averaged with high scores indicating positive feelings toward the planner (3-item  $\alpha = .63$ ).

**Budget-cutting task.** In a measure adapted from Haddock, Zanna, & Esses (1993; see also Niens, Cairns, & Bishop, 2004; Son Hing, Li, & Zanna, 2002), all participants were shown a list of 10 campus organizations (ex. YMCA on Campus, the Bike Project on Campus, etc.) that purportedly receive the most funding from the University each year. Participants were given each group’s current budget (totaling \$6,000 across all groups) and were told they would have to cut the budget to \$4,000. One of the organizations was the “African American Student Alliance.” We mea-

sured the amount of money allotted to the African American organization, with larger amounts representing more positivity toward African Americans.

**Harshness toward White versus African American bankrupt couple (Between-persons).** Participants were to imagine they were “a federal judge responsible for approving the settlement of a bankruptcy case.” They would “read about the bankruptcy case, and then be asked to make a few of the difficult decisions about how to handle the case.” Participants were given the story of a young family that included a custodian, his wife (a teacher’s aide in an elementary school), and their two small children. Participants were told that through a combination of typical bills, unfortunate circumstances (e.g., unemployment, necessary medical bills) and elective purchases (e.g., a Caribbean cruise), the couple had accrued \$27,300 in debt.

The critical manipulation was that for half of the participants, the couple had prototypically White names (Todd and Allison) and attended a predominantly White church (First United Methodist). For the other half of the participants, the couple’s names (Reggie and Latisha) and church (Bethel A.M.E.) were more prototypically African American (this task was adapted from Braucher, Cohen, & Lawless, 2012).

After reading the case, participants decided how much of the debt had to be repaid (from \$0 to \$27,300). Participants also used a slider to indicate whether the couple should be “trusted” to mail in their court-ordered payments (0), or whether their wages should be garnished from their paychecks (100). Garnishment gives the couple less freedom, and it also publicizes the bankruptcy to the person’s employer; thus, garnishment was considered harsher than trusting the couple to mail their payment in. Both variables were standardized and averaged, with higher scores representing greater punitiveness toward the couple.

Because the manipulation was between-subjects on this task, it was yoked to the financial planner task, such that the target in this task was opposite in race to the target in the financial planner task. Favoring African Americans over Whites would thus be shown by the difference between how favorably participants treated the African American financial planner and how punitive they were toward the White bankrupt couple (or the reverse, how lenient they were toward the bankrupt African American couple and how

<sup>14</sup> As in Olson and Fazio (2006), the CSs were as follows. Black: cashier, businesswoman, telephone repairwoman, nurse, minister, businessman, potter, and landscaper. White: police officer, painter, businesswoman, pharmacist, professor, garbage man, architect, and bricklayer. The USs followed Olson and Fazio (2006) as closely as possible, but there were a few differences. Examples of positive images: young couple hugging; young boy eating ice cream; a litter of puppies; cute baby seal. Examples of positive words: magnificent, amazing, exciting, incredible. Examples of negative images: an alley filled with trash; automobile exhaust; a young boy crying; a vicious wolverine. Examples of negative words: repulsive, disgusting, sickening, wicked. Specific word/image pairings were selected randomly for each participant.

Wealth-related targets that participants were to click were: First block: 4 pictures of coins, and the words “pennies,” “nickels,” “dimes,” and “quarters.” Second block: 4 pictures of dollar bills of different denominations, and words “ones,” “fives,” “twenties,” and “hundreds.” Third block: 4 pictures of precious stones, and words “diamonds,” “rubies,” “emeralds,” and “sapphires.” Fourth block: pictures of “things rich people buy” (a yacht, jet, villa, and mansion), and words “yacht,” “jet,” “villa,” and “mansion.” Fifth block: pictures of precious metals and words “gold,” “silver,” “platinum,” and “pure gold.”

unfavorable they were toward the White financial planner). Computing the difference in treatment let us convert these two between-subjects variables into a within-subject variable called *diff\_treatment*.

**Physical attractiveness ratings.** Participants were to imagine they were venture capitalists evaluating a start-up company. The company would run a dating website, with an algorithm suggesting matches after “learning” about viewers’ preferences from a face rating task. Under the pretense of evaluating a “demo” of the website, participants rated the attractiveness of 15 White and 5 African American faces using a slider from *not at all attractive* (0) to *extremely attractive* (100). Participants chose which gender they rated. Participants answered a few filler questions about their willingness to back the startup, but the real purpose of the task was to get their opinions about the attractiveness of the African American ( $\alpha = .87$ ) and White ( $\alpha = .90$ ) faces, indicated by the difference score between them.

**Index score.** A principal axis factor analysis suggested that the *diff\_treatment*, budget cutting, and attractiveness ratings all tapped a single underlying construct (all  $\lambda_s \geq .30$ ). Thus, although the tasks were quite different from one another, they all shared a single common source of variance, which is likely positive or negative feelings toward people of a certain race/physical appearance. Therefore, these three measures were standardized and averaged to form a composite (high scores indicating greater favoritism toward African Americans over Whites). The index was itself standardized.<sup>15</sup>

**Suspicion probes.** The preceding tasks tried to get at participants’ racial attitudes without participants realizing we were measuring their racial attitudes. The evaluative conditioning paradigm also attempts to affect participants without them being aware of the manipulation. To assess this, after the measures above were collected, we asked participants: (a) “Think back to the first task you completed in this study (the slide show task) . . . were there any patterns that seemed to recur?” (b) “Was any sort of picture more likely to be paired with other pictures or words that were positive or negative? If so, what was the pairing?” (c) “Did you notice anything unusual about the words and pictures that were presented with pictures of people of different races? What specifically?” (d) “Relative to White people, were pictures of Black people: more likely to be paired with positive words; equally likely to be paired with positive/negative words; more likely to be paired with negative words; I don’t know” (e) “If you had to guess, what would you say the hypothesis of the study is?”

Even after the hints on the first 3 questions, only 13% seemed to be aware of what was happening. Eliminating those participants leaves results quite similar to those reported below (Religious Group  $\times$  Conditioned interaction  $p = .011$ ). The forced choice on the 4th question confirmed that participants were generally unaware that they were being negatively conditioned to African Americans. Conditioned respondents were 13% more likely than nonconditioned respondents to say that African Americans were paired with negative stimuli (33% of conditioned respondents vs. 20% of nonconditioned respondents). However, they were also 10% more likely to say that African Americans were paired with positive stimuli (14% vs. 4%). Overall, guesses about the valence of stimuli paired with African Americans did not differ between the conditioned versus nonconditioned,  $F(1, 161) = .11, p = .74$ . On the fifth question, only 7% correctly guessed the hypothesis,

likely because all preceding tasks did *not* seem to be about race. If one eliminates participants based on the first 3 and 5th questions, results remain similar to those below (Religious Group  $\times$  Conditioned interaction  $p = .020$ ).<sup>16</sup>

**Religiosity as a potential moderator.** As in Study 1, this was a composite of religious service attendance and self-described religiousness.

**Individual difference measures.** Respondents provided self-report ratings of: big 5 personality traits using a combination of the Big 5 Inventory (for agreeableness; John & Srivastava, 1999) and Ten Item Personality Inventory (for all other dimensions; Gosling, Rentfrow, & Swann, 2003); and attachment orientations using the Experiences in Close Relationships Short Form (Wei, Russell, Mallinckrodt, & Vogel, 2007).

Finally, to undo any potential harm caused by our manipulation, after participants completed all measures, they viewed another slideshow pairing both Whites and African Americans with positive images/words.

## Results

Our primary interest was a Religious Group  $\times$  Conditioned versus not conditioned interaction, with the Positivity toward African Americans over Whites Index as the dependent variable. The prediction was for an interaction, with conditioning leading non-Protestants to be relatively more negative toward African Americans (consistent with Olson & Fazio, 2006) and conditioning leading Protestant participants to be relatively more positive toward African Americans (consistent with reaction formation). We checked also for any effects of gender. Overall, women were more favorable toward African Americans than males were (female  $M = .10$ , male  $M = -.09, F(1, 162) = 4.75, p = .031$ , a pattern driven by the Protestants (Religious Group  $\times$  Gender interaction,  $F(1, 162) = 3.93, p = .049$ ). However, gender did *not* moderate the key

<sup>15</sup> One can analyze the data using somewhat different methods to test the omnibus effect across all dependent variables. Both a repeated-measures ANOVA and a MANOVA revealed that the omnibus Religious Group  $\times$  Conditioned interaction was significant across the 3 dependent variables, both  $ps \leq .016$ . Univariate tests showed Religious Group  $\times$  Conditioned interactions on the *diff\_treatment*,  $F(1, 166) = 8.09, p = .005$  and budget cutting variables,  $F(1, 166) = 3.63, p = .059$ , but only a nonsignificant trend for face attractiveness ratings ( $p = .292$ ). Differences between effect sizes for different measures may be random variation, as tests of the Dependent measure  $\times$  Religious Group  $\times$  conditioned interaction indicated no significant differences between the size of Religious Group  $\times$  Conditioned interactions among the measures (both  $ps \geq .187$ ).

<sup>16</sup> We also had some measures that were obviously about race (some modern racism questions, feeling thermometers, and internal motivation to suppress prejudice items). Early on, we were getting very high suspicion rates—much higher than one usually gets for the evaluative conditioning paradigm (Olson, person communication). We thought it likely that the questions obviously about race were creating suspicion. To address this, after the first 22 participants, we moved suspicion probes to after our unobtrusive racial measures (described in the text) and before our obvious racial measures. We did this to get a better indication of the extent to which participants were aware of the evaluative conditioning manipulation. Consequently, we suggest that measures after the suspicion probes be treated with some caution. If one combines postsuspicion measures with the measures described in the text (overall  $\alpha = .70$ ), the Religious Group  $\times$  Conditioned interaction remains significant,  $p = .005$ . Deleting the first 22 cases also would not alter conclusions (Religious Group  $\times$  Conditioned interaction  $p = .006$ ).

interaction of interest (Gender  $\times$  Religious Group  $\times$  Conditioned interaction,  $F(1, 162) = .20, p = .655$ ), so we collapsed over gender in results below.

The predicted Religious Group  $\times$  Conditioned versus not conditioned interaction occurred,  $F(1, 162) = 9.47, p = .002$ ; Figure 4. Decomposing the interaction, evaluative conditioning had the expected effect on non-Protestants, leading them to have more negative judgments about African Americans (conceptually replicating Olson & Fazio, 2006; conditioned vs. not conditioned  $M_s = -.17$  vs.  $.21, t(124) = 2.20, p = .030$ ). For Protestant participants, however, evaluative conditioning led to the reverse. That is, conditioning to induce negative feelings toward African Americans led Protestants to show more positivity toward African Americans (conditioned vs. not conditioned  $M_s = .22$  vs.  $-.42, t(42) = 2.18, p = .031$ ).<sup>17</sup>

#### Further analyses.

**Moderation by religiosity.** The Religious Group  $\times$  Conditioned interaction was not moderated by religiosity ( $t = .06, p = .95$ ).

**Other individual differences.** We examined whether any Big 5 factor, attachment anxiety, or attachment avoidance interacted with the Conditioned variable to predict higher scores on the Positivity toward African Americans over Whites Index. There were 2 significant or marginally significant interactions: Openness  $\times$  Conditioned ( $b = -.22, t = -2.19, p = .030$ ) and Avoidance  $\times$  Conditioned ( $b = .17, t = 1.77, p = .079$ ). Centering the individual difference variables at  $\pm 1$  SD from the mean, effects were that high openness people and low avoidant people could be conditioned to feel negatively toward African Americans (simple slope of Condition at  $+1$  SD on openness and  $-1$  SD on avoidance,  $b = -.45, t = -2.06, p = .041$  and  $b = -.39, t = -1.85, p = .066$ , respectively). Those low in openness and high in avoidance did not show a conditioning effect ( $p = .29$  and  $p = .66$ , respectively). Even at the highest observed avoidant value ( $+3.16$  SDs from the mean), the effect of conditioning would not be significant ( $p = .15$ ), though at the lowest observed openness value ( $-3.54$  SDs from the mean), the effect of conditioning would be marginally significant ( $p = .059$ ). Thus, for the 7 individual difference variables, there was evidence for greater susceptibility to conditioning among those high in openness and

low in avoidance, but only at the most extreme value ( $-3.54$  SDs from the mean) on openness was there any evidence for reaction formation. Overall, reaction formation does not appear to be a process well predicted by individual difference variables we usually talk about when we think of personality processes.

**Summary.** If Study 3 demonstrated that Protestants (vs. non-Protestants) were likely to turn taboo attractions into repulsions, Study 4 demonstrated that the process can also run the other way, with Protestants turning taboo repulsions into attractions. Thus, in Study 4, evaluative conditioning to induce repulsion successfully made non-Protestants harsher toward African Americans; but among Protestants, such negative evaluative conditioning actually produced greater positivity toward African Americans.

### Study 5: Defense Mechanism Questionnaire to Identify Who Shows Reaction Formation

Using established questionnaires that measure defenses, Study 5 examined whether Protestants were more likely to show the defense of reaction formation than non-Protestants were. It also connected the present work to other research showing that Protestants (vs. non-Protestants) were more likely to sublimate forbidden thoughts and desires into creative work (Hudson & Cohen, 2016; Kim & Cohen, 2017; Kim et al., 2013; Weber, 1905).

In terms of individual differences, one important issue is whether: (a) It is the same individuals who engage in both sublimation and reaction formation, and whether one defense or another is elicited depends on the situation and the opportunities it affords. If this were true, the crucial distinction would be between “defensive” and “nondefensive” individuals, with the particular defense they employed just depending on the situation. Or (b) Individuals have different signature defensive styles, such that forbidden thoughts elicit sublimation from one sort of person and elicit reaction formation from another sort of person.

In terms of the latter hypothesis, sublimation is considered one of the most mature defenses, whereas reaction formation is considered less mature. Reaction formation is still considered *relatively* mature, though, because it often leads to the person adopting conventional, socially appropriate attitudes (e.g., spurring them to “Love thy neighbor” [Study 1] or adopt nonracist attitudes [Study 4]). This sense of maturity draws on Vaillant’s distinction between defenses that alienate others versus those that do not alienate (and may even attract) others.

Sublimation may also be more mature than reaction formation in another sense. That is, as shown by Kim and Cohen (2017), sublimation seems to involve two psychological skills: the ability to tamp down troublesome affect and the ability to displace it into an acceptable channel. The importance of tamping down was shown in Study 2 of Kim and Cohen (2017), finding that Protestant repressors have high interest in creative activities, compared with Catholic repressors and Protestant nonrepressors. The importance

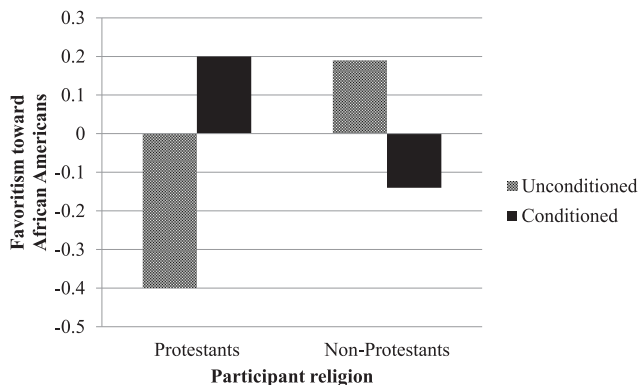


Figure 4. Positivity toward African Americans as a function of religion and whether the participant had been conditioned by having negative stimuli paired with African American photos.

<sup>17</sup> If one examines data with only Catholics and Jews as the non-Protestants, the interaction remains significant ( $p = .027$ ), though the simple effect for non-Protestants does not remain significant. Also, the distribution on the dependent variable showed skew. If one squares the dependent variable, this dramatically reduces skew and results remain similar to those in the text ( $p = .003$  for the Religious Group  $\times$  Conditioned interaction when all non-Protestants are included;  $p = .028$  when only Catholics and Jews are included as non-Protestants).

of both tamping down and displacing was shown in Study 3 of that paper, from which data for the current study are drawn. In that paper, it was found that Protestants who scored high on defenses that involved tamping down troublesome affect (ex., intellectualization, isolation, suppression, etc.) were more creative than Protestants who scored low on these defenses. Further, Protestants who scored high on displacing defenses (ex., displacement, fantasy, acting out, etc.) were also more creative than those who scored low on these defenses. None of the patterns above were shown by Catholics.

Relevant to the defense of reaction formation, we expect that Protestants who are unable to sublimate troublesome affect by diminishing it or displacing it will instead reverse it. Thus, our hypothesis—in addition to Protestants showing greater reaction formation overall than non-Protestants—is that those Protestants who do not or cannot engage in the diminishment and displacement processes necessary for successful sublimation will instead be prone to reaction formation.

## Method

Participants included 96 Catholic and 81 Protestant undergraduates (71% female). (The limited availability of Jews in the psychology participant pool that year prevented us from recruiting them for this study.)

**Materials.** As described in Kim and Cohen (2017), we had 2 well-established measures of defense mechanisms: the Defense Style Questionnaire (DSQ) and Life Style Index (LSI).

**Reaction formation.** The 8 LSI items mostly involved moral revulsion, often but not exclusively about sexual matters.<sup>18</sup> Sample items (answered yes vs. no): “I feel outraged at dirty movies,” “I would never go to a movie that was X-rated,” “Promiscuity is disgusting,” and “My moral standards are higher than those of most people I know.” The 2 DSQ items (answered on a 1-to-9 Likert scale) were: “If someone mugged me and stole my money, I’d rather he be helped than punished” and “I often find myself being very nice to people who by all rights I should be angry at.” We put all items on a 0–1 metric and averaged them (10-item  $\alpha = .75$ ).

**Affect diminishing and affect displacing defenses.** As described in Kim and Cohen (2017), we created an index of Affect Diminishing defenses and another of Affect Displacing defenses. Protestants scoring high on Diminishment and Displacement tended to be the most creative, suggesting that both these processes were operative in sublimation. These effects did not hold for Catholic participants.

A brief explanation of each defense and sample items are in Kim and Cohen (2017), so we do not repeat that information here. The Affect Diminishing defenses were: isolation, intellectualization, undoing, suppression, sublimation, anticipation, repression, dissociation, denial ( $\alpha = .79$ ). The Affect Displacing defenses were: displacement, fantasy, passive-aggressive behavior, acting out, help-rejecting complaining, pseudoaltruism, projection, projective identification ( $\alpha = .80$ ).

**Religiosity as a potential moderator.** As in Study 1, this was a composite of religious service attendance and self-described religiosity.

## Results

Primary analyses involved (a) a main effect of religious group and (b) an interaction between affect diminishing defenses, affect displacing defenses, and religious group, such that Protestant participants who could neither diminish nor displace troublesome affect were expected to show greater reaction formation.

As noted in Kim and Cohen (2017), the defense inventories were given after a manipulation in which participants filled out a control questionnaire or a questionnaire about their transgressions that used either a high-frequency or low-frequency scale (for a project on asceticism; Schwarz & Scheuring, 1988). We included this variable as well as gender in a Religious Group  $\times$  Gender  $\times$  Manipulation Condition ANOVA.

As seen in Figure 5, Protestant respondents did score higher on reaction formation ( $M = .40$ ) than Catholic respondents did ( $M = .32$ ),  $F(1, 160) = 6.35$ ,  $p = .013$ . Women scored higher than men on reaction formation,  $F(1, 160) = 9.49$ ,  $p = .002$ . However, whereas the Religious Group difference was greatest among participants who answered the transgression versus control questionnaire and among men, neither variable significantly moderated the Religious Group main effect (Religious Group  $\times$  Gender and Religious Group  $\times$  Questionnaire condition interactions were not significant,  $F(1, 160) = 1.65$ ,  $p = .20$  and  $F(2, 160) = .84$ ,  $p = .44$ , respectively.<sup>19</sup>

As seen in Figure 6, in terms of individual differences in defense use among Protestants, a regression revealed an Affect Diminishment  $\times$  Affect Displacement interaction predicting reaction formation ( $b = .06$ ,  $\beta = .24$ ,  $t = 2.26$ , partial  $r = .18$ ,  $p = .025$ ). Among Catholics, this interaction was not significant ( $b = -.01$ ,  $t = -.38$ ,  $p = .71$ ). The Religious Group  $\times$  Affect Diminishment  $\times$  Affect Displacement interaction was thus marginally significant,  $b = -.06$ ,  $\beta = -.18$ ,  $t = -1.76$ ,  $p = .081$ , partial  $r = -.14$ . Figure 6 plots the interactions at  $\pm 1$  SD from the mean on both defenses. When means for affect displacement and affect diminishment defenses are 1 SD below the average, Protestants are more likely than Catholics to show reaction formation and the difference is quite pronounced (simple slope for Religious group  $b = -.22$ ,  $t(162) = 3.29$ ,  $p = .001$ , partial  $r = .25$ ). Even when means for affect displacement and affect diminishment defenses are at their average value, Protestants are still more likely than Catholics to show reaction formation (simple slope for Religious group  $b = -.07$ ,  $t(162) = 1.93$ ,  $p = .055$ , partial  $r = .15$ ). When affect diminishment and affect displacement defenses are both 1 SD above the mean, there is no significant difference

<sup>18</sup> Two items lowered scale reliability and were removed. These items – “Touching anything slimy makes me feel nauseous” and “Using public bathrooms is very upsetting to me”—had to do with a general squeamishness rather than moral revulsion (“Promiscuity is disgusting.”). Including them changes slightly the Religious Group main effect predicting reaction formation ( $p = .076$ ) and hardly effects the Religious Group  $\times$  Affect Diminishment  $\times$  Affect Displacement interaction ( $p = .089$ ).

<sup>19</sup> The Religious Group difference in reaction formation was significant when participants first completed the transgression questionnaire,  $t(119) = 2.46$ ; the difference was smaller,  $t(51) = .52$  when participants first completed the control questionnaire. Though the LSI and DSQ are measures of chronic defenses, it is possible—and we think likely—that reflecting on one’s transgressions enhances religious group differences. However, as noted in the text, the interaction of Religious Group  $\times$  Transgression questionnaire manipulation condition was not significant.



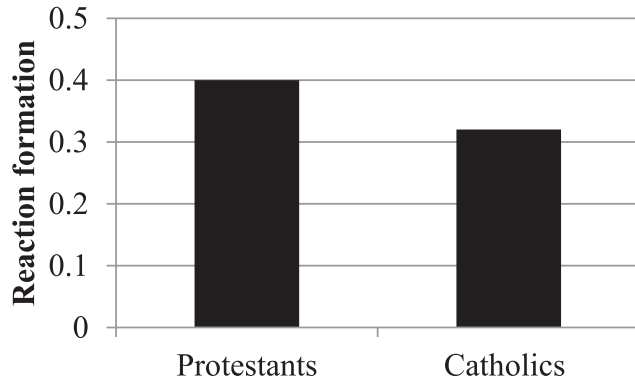


Figure 5. Reaction formation scores of Protestant and Catholic respondents.

between groups in reaction formation (simple slope =  $-.05$ ,  $t = -1.75$ ,  $p = .457$ ).

**Further analyses: Moderation by religiosity.** The main effect of Religion on reaction formation was marginally significantly moderated by religiosity (Religion  $\times$  Religiosity interaction  $b = -.06$ ,  $t = -1.77$ ,  $p = .079$ , partial  $r = -.14$ ), as highly religious Protestants scored much higher on reaction formation (simple slope of religiosity among Protestants,  $b = .14$ ,  $\beta = .61$ ,  $t = 6.32$ ,  $p = .001$ , partial  $r = .44$ ) whereas this tendency was significant but less pronounced among Catholics ( $b = .09$ ,  $\beta = .37$ ,  $t = 3.87$ ,  $p = .001$ , partial  $r = .29$ ).

The three-way interaction of Religion  $\times$  Affect Diminishment  $\times$  Affect Displacement was not moderated by religiosity (4-way interaction  $p = .39$ ).

Finally, given that affect diminishment and displacement were the processes driving Protestant sublimation (Kim & Cohen, 2017) and that an inability/unwillingness to use these defenses predicted Protestant reaction formation, we also examined whether reaction

formation scores correlated negatively with creativity scores among Protestants (creativity measure described in Kim & Cohen, 2017). This was not the case, however ( $r = -.04$ ,  $p = .73$ ; the correlation was also not significant among Catholics,  $r = .02$ ). Apparently, engaging in reaction formation does not preclude creativity among people of either religious group—a result that is sensible unless one believed that all creative activity was driven by sublimation.

## General Discussion

Across studies, Protestant participants seemed more likely than non-Protestants to show reaction formation. They were more likely to endorse “love thy enemy” sentiments, more likely to be repulsed by unconventional/deviant behavior when induced to be attracted to it, more likely to be attracted to African Americans when conditioned to feel negatively toward them, more likely to be judged by clinicians as engaging in reaction formation, and more likely to score high on reaction formation measures from commonly used defense scales. The last finding was notable, because that study also highlighted an individual difference predictor showing which Protestants were more likely to use reaction formation. A dozen conventional individual difference measures failed to predict reaction formation. However, what predicted Protestant reaction formation was a profile of individuals’ defense mechanism use, such that Protestants who neither diminished nor displaced troublesome affect engaged in reaction formation. Thus, if Protestants could not diminish or displace troublesome affect, they reversed it.

It is important to note also what results did *not* derive from. Results did not seem to derive from impression management concerns (participants’ responses in experiments were entered into a computer with no one else in the room; participants were not implicitly or explicitly accused of anything by an experimenter; participants’ subsequent behavior did not involve interacting with anyone). Results were not attributable to reactance (participants in

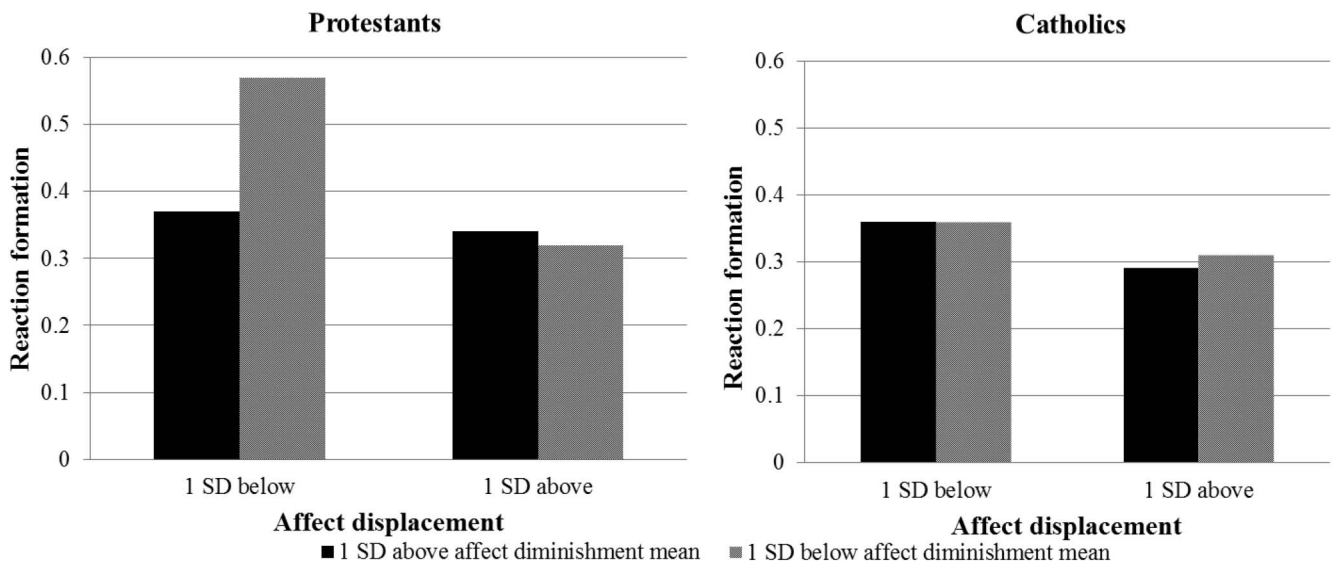


Figure 6. Reaction formation scores as a function of participants’ religion and their use of affect diminishing and affect displacing defenses.

the evaluative conditioning study were overwhelmingly unaware that they had been manipulated, and eliminating those few participants who may have been aware did not change results). Results were not attributable to manipulations simply failing to work on Protestants (in the study with the erotic vs. nonerotic video, the erotic video made them more likely to endorse conventional sexual behaviors even as it made them more condemning of unconventional/deviant ones). Results were not attributable to non-Protestants (or at least Jews) simply being better able to hide their reaction formation in *nonclinical* settings (clinicians also rated Protestants as higher in reaction formation than Jews). Results were not attributable to religion being confounded with conventional personality variables—there were no differences between Protestants and non-Protestants on any of the conventional measures, with the exception that Protestants were lower in attachment anxiety. However, attachment anxiety did not predict reaction formation; and in fact, reaction formation patterns were not predicted by individual difference variables that we usually think of when discussing “normal” personality processes involving self-control and emotion regulation (conscientiousness, agreeableness, neuroticism, anxiety), closed or open-mindedness (openness to experience, dogmatism, need for closure), sensitivity to reward and punishment (BIS, BAS, extroversion, neuroticism), social desirability and impression management (the Crowne Marlowe scale), or general or relationship insecurity (avoidant attachment, anxious attachment, anxiety). The processes described here as reaction formation are hard to dismiss as “simply” the product of either some other phenomena or some other personality process.

### Is It Really Reaction Formation?

One might object that the data do not indicate reaction formation. People can genuinely find certain sexual behaviors disgusting without feeling secretly attracted to them. People can genuinely turn the other cheek without wanting to strike back. Not everything that *might* be a defense is actually a defense.

We agree that “sometimes a cigar is just a cigar.” But one criterion is that we call a process defensive when it reacts to something in peculiar ways. In the case of detecting such reactivity, experiments seem particularly useful. If objects become more repellent when they should become more attractive or if they become more attractive when they should become more repellent, then it seems like these shifts may reflect a defensive process protecting the person against feelings he or she does not wish to have (see also Sarnoff, 1960).<sup>20</sup>

### Unresolved Issues

There are still several unresolved questions. First, whereas we have argued that Protestants should be more susceptible to reaction formation than non-Protestants, we would expect that the defense is not completely alien to non-Protestants. Likely there are some situational or individual-difference predictors of reaction formation among non-Protestants, though we have yet to identify such situations and have come up basically empty on conventional individual difference measures. Under what conditions non-Protestants might show reaction formation is thus an open question (see also Norenzayan & Heine’s, 2005, discussion of types of universality).

Second, it is unclear how much religiosity moderates group differences in reaction formation. Study 1 found clear evidence of moderation in predicting “love thy enemy” sentiments. Study 5 showed a marginally significant moderation effect. Studies 3 and 4 showed no significant moderation effect. Meta-analyzing results across studies, we found that the effect of religiosity moderating Religious Group differences was small but significant (meta-analytic  $Z = 3.13$ ,  $p = .0018$ , unweighted by study  $n$ ;  $Z = 4.31$ ,  $p = .000016$ , weighted by study  $n$ ; average effect size translates to about  $d = .14$ , unweighted,  $d = .2$ , weighted). Religiosity does not necessarily moderate all religious group differences in defense mechanism use (Kim & Cohen, 2017; see also Hayward & Kimmelmeier, 2011). On balance, the evidence here suggests it does moderate differences in reaction formation. To the extent that differences in the moderation effect across studies were not merely sampling error, exploring when and why religiosity acts as a moderator may be a useful avenue for future research.

Third, we speculated that Protestants are likely to engage in reaction formation when they *cannot* sublimate. But is this correct? We have argued that sublimation is the more complicated, more “mature” defense. The maturity claim aligns with traditional thinking in that sublimation typically affords more good for both the individual and society than reaction formation does. The claim that sublimation is more complicated is based on the idea that it involves processes of affect minimization and processes of displacement. Thus, we have argued that if Protestants cannot engage in these two processes, then they are more likely to engage in reaction formation. That is, if they cannot minimize or displace problematic affect, then they reverse it. Metaphorically, if they can neither quiet nor redirect the troubling voices inside their heads, they shout them down, vehemently asserting the opposite.

The argument is predicated on the thesis that people will generally choose the more mature defense, if they are capable of doing so. However, sometimes the issue may be less about *cannot* and more about *will not*. If people are “cognitive/effort misers” attempting to take the easy route, then it is possible that reaction formation is the default or the defense people will go to first. Further, to the extent that reaction formation leads one to conform to local norms, reaction formation may not drive people away; instead, it may be socially rewarded. Imagine, for example, a man

<sup>20</sup> This criterion does bring up some interesting philosophical issues. Suppose the person’s reaction formation becomes so chronic and hyper-vigilant that their extreme response to the stimuli becomes their “normal,” persistent attitude. In this case, for any given individual, experiments would be unable to detect any *heightened* reactivity in threatening situations. Thus, experiments would be unable to distinguish between chronic attitudes that originated as reaction formation and later became “normal” versus chronic attitudes of similar extremity that did not originate in reaction formation. In the former case, are we willing to call such attitudes “genuine” and ignore how they were originally produced, or should we consider such attitudes reaction formation *par excellence*, because the reaction formation has managed to commandeer the person’s everyday life and personality?

A seemingly analogous but very different issue emerges in cases where a person’s reaction formation is employed so effectively and so automatically in a given situation that one can see no trace of any defensive effort being exerted (e.g., no evidence of any ego depletion effects). In that case, an experiment *is* an effective way to detect defensiveness, because one can compare individuals’ attitudes in situations where threatening feelings and beliefs should arise versus attitudes in situations where no threat should be produced.

who shows reaction formation against his own gay inclinations by becoming homophobic. If he lives in an environment that is homophobic, he may “fit in” reasonably well; and if his prejudices are rewarded, he may even gain status for holding and acting on them. Or, for example, persons who show reaction formation by being dutiful and loyal toward parents that they actually have some resentment toward will draw the praise of family and community for being a “good son” or “good daughter.” In general, we think it likely that the type of defense people employ (a) depends not just on what people *can* do (in terms of their abilities to control and channel their thoughts and feelings) but on what they *will* do (in terms of attempting to minimize effort or draw social approval) and (b) depends not just on intrapsychic factors but on social/contextual factors such as how much the use of a defense is rewarded or punished by the person’s immediate social environment and reference groups. What is afforded and rewarded is not just a matter of individual psychology, but also the sociocultural milieu. Such factors go beyond our present scope but seem like promising topics for future work.

In the meantime, however, the present studies demonstrated the way religion plays a role in shaping people’s psychological defenses. With Protestantism’s emphasis on the “morality of mentality” (A. Cohen & Rozin, 2001), Protestant adherents seem to engage in greater use of reaction formation than their non-Protestant counterparts. If, as Darwin thought, controlling our mental life is the highest stage of moral culture, then at least among those belonging to the relatively late-developing religion of Protestantism, it may take some of our most repellent impulses to drive us to our purest thoughts.

## References

- Abramowitz, J. S., Deacon, B. J., Woods, C. M., & Tolin, D. F. (2004). Association between Protestant religiosity and obsessive-compulsive symptoms and cognitions. *Depression and Anxiety, 20*, 70–76. <http://dx.doi.org/10.1002/da.20021>
- Adams, H. E., Wright, L. W., Jr., & Lohr, B. A. (1996). Is homophobia associated with homosexual arousal? *Journal of Abnormal Psychology, 105*, 440–445. <http://dx.doi.org/10.1037/0021-843X.105.3.440>
- Ariely, D., & Loewenstein, G. (2006). The heat of the moment: The effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making, 19*, 87–98. <http://dx.doi.org/10.1002/bdm.501>
- Baumeister, R. F., Dale, K., & Sommer, K. L. (1998). Freudian defense mechanisms and empirical findings in modern social psychology: Reaction formation, projection, displacement, undoing, isolation, sublimation, and denial. *Journal of Personality, 66*, 1081–1124. <http://dx.doi.org/10.1111/1467-6494.00043>
- Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1993). When ego threats lead to self-regulation failure: Negative consequences of high self-esteem. *Journal of Personality and Social Psychology, 64*, 141–156. <http://dx.doi.org/10.1037/0022-3514.64.1.141>
- Baumeister, R., & Tierney, J. (2012). *Willpower*. New York, NY: Penguin.
- Berman, N. C., Abramowitz, J. S., Pardue, C. M., & Wheaton, M. G. (2010). The relationship between religion and thought-action fusion: Use of an in vivo paradigm. *Behaviour Research and Therapy, 48*, 670–674. <http://dx.doi.org/10.1016/j.brat.2010.03.021>
- Birren, J. (Ed.). (1963). *Human aging*. Washington, DC: Department of Health, Education, and Welfare.
- Blumenthal, M., Kahn, R., & Andrews, F. (1978). *Justifying violence*. Ann Arbor, MI: ICPSR.
- Blumenthal, M., Kahn, R., Andrews, F., & Head, K. (1972). *Justifying violence*. Ann Arbor, MI: Institute for Social Research.
- Braucher, J., Cohen, D., & Lawless, R. (2012). Race, attorney influence, and bankruptcy chapter choice. *Journal of Empirical Legal Studies, 9*, 393–429. <http://dx.doi.org/10.1111/j.1740-1461.2012.01264.x>
- Brehm, J. W. (1966). *A theory of psychological reactance*. New York, NY: Oxford University Press.
- Cheval, B., Radel, R., Grob, E., Ghisletta, P., Bianchi-Demicheli, F., & Chanal, J. (2016). Homophobia: An impulsive attraction to the same sex? Evidence from eye-tracking data in a picture-viewing task. *Journal of Sexual Medicine, 13*, 825–834. <http://dx.doi.org/10.1016/j.jsxm.2016.02.165>
- Chivers, M. L., Rieger, G., Latty, E., & Bailey, J. M. (2004). A sex difference in the specificity of sexual arousal. *Psychological Science, 15*, 736–744. <http://dx.doi.org/10.1111/j.0956-7976.2004.00750.x>
- Cohen, A. B. (2003). Religion, likelihood of action, and the morality of mentality. *International Journal for the Psychology of Religion, 13*, 273–285. [http://dx.doi.org/10.1207/S15327582IJPR1304\\_4](http://dx.doi.org/10.1207/S15327582IJPR1304_4)
- Cohen, A. B. (2009). Many forms of culture. *American Psychologist, 64*, 194–204. <http://dx.doi.org/10.1037/a0015308>
- Cohen, A. B. (2013). *Culture reexamined*. Washington: APA.
- Cohen, A. B. (2015). Religion’s profound influences on psychology: Morality, intergroup relations, self-construal, and enculturation. *Current Directions in Psychological Science, 24*, 77–82. <http://dx.doi.org/10.1177/0963721414553265>
- Cohen, A. B., Hall, D. E., Koenig, H. G., & Meador, K. G. (2005). Social versus individual motivation: Implications for normative definitions of religious orientation. *Personality and Social Psychology Review, 9*, 48–61. [http://dx.doi.org/10.1207/s15327957pspr0901\\_4](http://dx.doi.org/10.1207/s15327957pspr0901_4)
- Cohen, A. B., & Rankin, A. (2004). Religion and the morality of positive mentality. *Basic and Applied Social Psychology, 26*, 45–57. [http://dx.doi.org/10.1207/s15324834basp2601\\_5](http://dx.doi.org/10.1207/s15324834basp2601_5)
- Cohen, A. B., & Rozin, P. (2001). Religion and the morality of mentality. *Journal of Personality and Social Psychology, 81*, 697–710. <http://dx.doi.org/10.1037/0022-3514.81.4.697>
- Cohen, A. B., Siegel, J. I., & Rozin, P. (2003). Faith versus practice: Different bases for religiosity judgments by Jews and Protestants. *European Journal of Social Psychology, 33*, 287–295. <http://dx.doi.org/10.1002/ejsp.148>
- Cohen, A. B., & Varnum, M. (2016). Beyond East vs. West: Social class, regions, and religions as forms of culture. *Current Opinion in Psychology, 8*, 5–9. <http://dx.doi.org/10.1016/j.copsyc.2015.09.006>
- Cohen, D. (2007). Methods in cultural psychology. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 196–236). New York, NY: Guilford Press.
- Cohen, D., Kim, E., & Hudson, N. W. (2014). Religion, the forbidden, and sublimation. *Current Directions in Psychological Science, 23*, 208–214. <http://dx.doi.org/10.1177/0963721414531436>
- Cohen, D., & Nisbett, R. E. (1994). Self-protection and the culture of honor. *Personality and Social Psychology Bulletin, 20*, 551–567. <http://dx.doi.org/10.1177/0146167294205012>
- Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers’ accuracy and predictive validity. *Psychological Bulletin, 136*, 1092–1122. <http://dx.doi.org/10.1037/a0021212>
- Crandall, C. S., & Biernat, M. (1990). The ideology of anti-fat attitudes. *Journal of Applied Social Psychology, 20*, 227–243. <http://dx.doi.org/10.1111/j.1559-1816.1990.tb00408.x>
- Darwin, C. (2008). *The descent of man*. Princeton, NJ: Princeton University Press.
- Department of Homeland Security. (2004). *2003 yearbook of immigration statistics*. Washington, DC: Author.
- Dutton, D. G., & Lake, R. A. (1973). Threat of own prejudice and reverse discrimination in interracial situations. *Journal of Personality and Social Psychology, 28*, 94–100. <http://dx.doi.org/10.1037/h0035582>

- Ellis, B., & Symons, D. (1990). Sex differences in sexual fantasy. *Journal of Sex Research, 27*, 527–555. <http://dx.doi.org/10.1080/00224499009551579>
- Glueck, S., & Glueck, E. (1950). *Unraveling juvenile delinquency*. New York, NY: Commonwealth.
- Gosling, S., Rentfrow, P., & Swann, W., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*, 504–528. [http://dx.doi.org/10.1016/S0092-6566\(03\)00046-1](http://dx.doi.org/10.1016/S0092-6566(03)00046-1)
- Granick, S., & Kleban, M. (2002). *Human aging*. Ann Arbor, MI: ICPSR.
- Haddock, G., Zanna, M., & Esses, V. (1993). Assessing the structure of prejudicial attitudes: The case of attitudes toward homosexuals. *Journal of Personality and Social Psychology, 65*, 1105–1118. <http://dx.doi.org/10.1037/0022-3514.65.6.1105>
- Heine, S. J., Buchtel, E. E., & Norenzayan, A. (2008). What do cross-national comparisons of personality traits tell us? The case of conscientiousness. *Psychological Science, 19*, 309–313. <http://dx.doi.org/10.1111/j.1467-9280.2008.02085.x>
- Hayward, R., & Kimmelmeier, M. (2011). Weber revisited. *Journal of Cross-Cultural Psychology, 42*, 1406–1420. <http://dx.doi.org/10.1177/0022022111412527>
- Hiller, J. (2004). Speculations on the links between feelings, emotions and sexual behaviour: Are vasopressin and oxytocin involved? *Sexual and Relationship Therapy, 19*, 393–412. <http://dx.doi.org/10.1080/14681990412331297974>
- Hudson, N. W., & Cohen, D. (2016). Taboo desires, creativity, and career choice. *Motivation and Emotion, 40*, 404–421. <http://dx.doi.org/10.1007/s11031-016-9540-3>
- Hughes, J., Grossmann, I., & Cohen, A. B. (2015). Tolerating the “doubting Thomas”: How centrality of religious beliefs vs. practices influences prejudice against atheists. *Frontiers in Psychology, 6*, 1352. <http://dx.doi.org/10.3389/fpsyg.2015.01352>
- Huntsman, E. (2010). The six antitheses: Attaining the purpose of the law through the teachings of Jesus. In G. Strathearn, T. Wayment, & D. Belnap (Eds.), *The sermon on the mount in latter-day scripture* (pp. 93–109). Provo, UT: Brigham Young.
- John, O., & Srivastava, S. (1999). The big five trait taxonomy. In L. Pervin & O. John (Eds.), *Handbook of personality* (pp. 102–138). Amsterdam, The Netherlands: Elsevier.
- Jordan, C. H., Spencer, S. J., & Zanna, M. P. (2005). Types of high self-esteem and prejudice: How implicit self-esteem relates to ethnic discrimination among high explicit self-esteem individuals. *Personality and Social Psychology Bulletin, 31*, 693–702. <http://dx.doi.org/10.1177/0146167204271580>
- Jordan, C. H., Spencer, S. J., Zanna, M. P., Hoshino-Browne, E., & Correll, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology, 85*, 969–978. <http://dx.doi.org/10.1037/0022-3514.85.5.969>
- Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the big five. *Psychological Bulletin, 116*, 245–258. <http://dx.doi.org/10.1037/0033-2909.116.2.245>
- Kim, E., & Cohen, D. (2017). Roads more and less traveled: Different emotional routes to creativity among Protestants and Catholics. *Journal of Personality and Social Psychology, 113*, 1000–1015. <http://dx.doi.org/10.1037/pspp0000105>
- Kim, E., Zeppenfeld, V., & Cohen, D. (2013). Sublimation, culture, and creativity. *Journal of Personality and Social Psychology, 105*, 639–666. <http://dx.doi.org/10.1037/a0033487>
- Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin, 129*, 522–544. <http://dx.doi.org/10.1037/0033-2909.129.4.522>
- Laurin, K., & Plaks, J. E. (2014). Religion and punishment. *Social Psychological and Personality Science, 5*, 835–843. <http://dx.doi.org/10.1177/1948550614534698>
- Li, Y. J., Johnson, K. A., Cohen, A. B., Williams, M. J., Knowles, E. D., & Chen, Z. (2012). Fundamental(ist) attribution error: Protestants are dispositionally focused. *Journal of Personality and Social Psychology, 102*, 281–290. <http://dx.doi.org/10.1037/a0026294>
- Loewenstein, G., Nagin, D., & Paternoster, R. (1997). The effect of sexual arousal on expectations of sexual forcefulness. *Journal of Research in Crime and Delinquency, 34*, 443–473. <http://dx.doi.org/10.1177/0022427897034004003>
- Mahaffey, A., Bryan, A., Ito, T., & Hutchison, K. (2011). In search of the defensive function of sexual prejudice. *Journal of Applied Social Psychology, 41*, 27–44. <http://dx.doi.org/10.1111/j.1559-1816.2010.00700.x>
- McFarlin, D. B., Baumeister, R. F., & Blascovich, J. (1984). On knowing when to quit: Task failure, self-esteem, advice, and nonproductive persistence. *Journal of Personality, 52*, 138–155. <http://dx.doi.org/10.1111/j.1467-6494.1984.tb00349.x>
- McGregor, I., & Jordan, C. H. (2007). The mask of zeal: Low implicit self-esteem, threat, and defensive extremism. *Self and Identity, 6*, 223–237. <http://dx.doi.org/10.1080/15298860601115351>
- McGregor, I., & Marigold, D. C. (2003). Defensive zeal and the uncertain self: What makes you so sure? *Journal of Personality and Social Psychology, 85*, 838–852. <http://dx.doi.org/10.1037/0022-3514.85.5.838>
- Mead, G. H. (1934). *Mind, self, and society*. Chicago, IL: University of Chicago Press.
- Meana, M. (2010). Elucidating women’s (hetero)sexual desire: Definitional challenges and content expansion. *Journal of Sex Research, 47*, 104–122. <http://dx.doi.org/10.1080/00224490903402546>
- Miller, W. (1997). *The anatomy of disgust*. Cambridge, UK: Harvard.
- Niens, U., Cairns, E., & Bishop, S. (2004). Prejudiced or not? Hidden sectarianism among students in Northern Ireland. *The Journal of Social Psychology, 144*, 163–180. <http://dx.doi.org/10.3200/SOCP.144.2.163-180>
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259. <http://dx.doi.org/10.1037/0033-295X.84.3.231>
- Norenzayan, A., & Heine, S. J. (2005). Psychological universals: What are they and how can we know? *Psychological Bulletin, 131*, 763–784. <http://dx.doi.org/10.1037/0033-2909.131.5.763>
- Oishi, S., & Roth, D. (2009). The role of self-reports in culture and personality research: It is too early to give up on self-reports. *Journal of Research in Personality, 43*, 107–109. <http://dx.doi.org/10.1016/j.jrp.2008.11.002>
- Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin, 32*, 421–433. <http://dx.doi.org/10.1177/0146167205284004>
- Ritter, R. S., & Preston, J. L. (2011). Gross gods and icky atheism: Disgust responses to rejected religious beliefs. *Journal of Experimental Social Psychology, 47*, 1225–1230. <http://dx.doi.org/10.1016/j.jesp.2011.05.006>
- Ritter, R. S., Preston, J. L., Salomon, E., & Relihan-Johnson, D. (2016). Imagine no religion: Heretical disgust, anger and the symbolic purity of mind. *Cognition and Emotion, 30*, 778–796. <http://dx.doi.org/10.1080/02699931.2015.1030334>
- Rozin, P., Kurzer, N., & Cohen, A. B. (2002). Free associations to “food”: The effects of gender, generation and culture. *Journal of Research in Personality, 36*, 419–441. [http://dx.doi.org/10.1016/S0092-6566\(02\)00002-8](http://dx.doi.org/10.1016/S0092-6566(02)00002-8)
- Sanchez-Burks, J. (2002). Protestant relational ideology and (in)attention to relational cues in work settings. *Journal of Personality and Social Psychology, 83*, 1000–1015. <http://dx.doi.org/10.1037/0022-3514.83.5.1000>

- Psychology*, 83, 919–929. <http://dx.doi.org/10.1037/0022-3514.83.4.919>
- Sanchez-Burks, J. (2004). Protestant relational ideology: The cognitive underpinnings and organizational implications of an American anomaly. *Research in Organizational Behavior*, 26, 265–305. [http://dx.doi.org/10.1016/S0191-3085\(04\)26007-5](http://dx.doi.org/10.1016/S0191-3085(04)26007-5)
- Sarnoff, I. (1960). Reaction formation and cynicism. *Journal of Personality*, 28, 129–143. <http://dx.doi.org/10.1111/j.1467-6494.1960.tb01607.x>
- Schwarz, N., & Scheuring, B. (1988). Judgments of relationship satisfaction: Inter- and intraindividual comparisons as a function of questionnaire structure. *European Journal of Social Psychology*, 18, 485–496. <http://dx.doi.org/10.1002/ejsp.2420180603>
- Scott, W. (1965). *Values and organizations*. Chicago, IL: Rand McNally.
- Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist*, 48, 1117–1131. <http://dx.doi.org/10.1037/0003-066X.48.11.1117>
- Sherman, S., & Gorkin, L. (1980). Attitude bolstering when behavior is inconsistent with central attitudes. *Journal of Experimental Social Psychology*, 16, 388–403. [http://dx.doi.org/10.1016/0022-1031\(80\)90030-X](http://dx.doi.org/10.1016/0022-1031(80)90030-X)
- Siev, J., Chambless, D. L., & Huppert, J. D. (2010). Moral thought-action fusion and OCD symptoms: The moderating role of religious affiliation. *Journal of Anxiety Disorders*, 24, 309–312. <http://dx.doi.org/10.1016/j.janxdis.2010.01.002>
- Siev, J., & Cohen, A. B. (2007). Is thought-action fusion related to religiosity? Differences between Christians and Jews. *Behaviour Research and Therapy*, 45, 829–837. <http://dx.doi.org/10.1016/j.brat.2006.05.001>
- Sommer, K. L., & Baumeister, R. F. (2002). Self-evaluation, persistence, and performance following implicit rejection. *Personality and Social Psychology Bulletin*, 28, 926–938. <http://dx.doi.org/10.1177/01467202028007006>
- Son Hing, L. S., Li, W., & Zanna, M. P. (2002). Inducing hypocrisy to reduce prejudicial responses among aversive racists. *Journal of Experimental Social Psychology*, 38, 71–78. <http://dx.doi.org/10.1006/jesp.2001.1484>
- Telushkin, J. (1994). *Jewish wisdom*. New York, NY: Morrow.
- Uhlmann, E. L., Poehlman, T. A., Tannenbaum, D., & Bargh, J. A. (2011). Implicit Puritanism in American moral cognition. *Journal of Experimental Social Psychology*, 47, 312–320. <http://dx.doi.org/10.1016/j.jesp.2010.10.013>
- Uhlmann, E. L., & Sanchez-Burks, J. (2014). The implicit legacy of American Protestantism. *Journal of Cross-Cultural Psychology*, 45, 992–1006. <http://dx.doi.org/10.1177/0022022114527344>
- Vaillant, G. (1993). *Wisdom of the ego*. Cambridge, MA: Harvard.
- Vaillant, G. (2012). *Triumphs of experience*. Cambridge, MA: Harvard. <http://dx.doi.org/10.4159/harvard.9780674067424>
- Vandello, J., & Cohen, D. (2016). *Domestic violence and purity*. Unpublished manuscript, University of Illinois, Chicago, IL.
- Vishkin, A., Bigman, Y. E., Porat, R., Solak, N., Halperin, E., & Tamir, M. (2016). God rest our hearts: Religiosity and cognitive reappraisal. *Emotion*, 16, 252–262. <http://dx.doi.org/10.1037/emo0000108>
- Vohs, K. D., & Heatherton, T. F. (2004). Ego threat elicits different social comparison process among high and low self-esteem people. *Social Cognition*, 22, 168–191. <http://dx.doi.org/10.1521/soco.22.1.168.30983>
- Weber, M. (1905). *The Protestant ethic and the spirit of capitalism*. New York, NY: Scribner.
- Wei, M., Russell, D. W., Mallinckrodt, B., & Vogel, D. L. (2007). The Experiences in Close Relationship Scale (ECR)-short form: Reliability, validity, and factor structure. *Journal of Personality Assessment*, 88, 187–204. <http://dx.doi.org/10.1080/00223890701268041>
- Weinstein, N., Ryan, W. S., Dehaan, C. R., Przybylski, A. K., Legate, N., & Ryan, R. M. (2012). Parental autonomy support and discrepancies between implicit and explicit sexual identities: Dynamics of self-acceptance and defense. *Journal of Personality and Social Psychology*, 102, 815–832. <http://dx.doi.org/10.1037/a0026854>
- Wilson, T. D. (2004). *Strangers to ourselves*. Cambridge, MA: Harvard.
- Wolpe, D. (2004, December 16). False promises in Berg's "Becoming." *Jewish Journal*. Retrieved October 7, 2016, from [http://www.jewishjournal.com/arts/article/false\\_promises\\_in\\_bergs\\_becoming\\_20041217](http://www.jewishjournal.com/arts/article/false_promises_in_bergs_becoming_20041217)

Received December 1, 2016  
 Revision received April 30, 2017  
 Accepted May 3, 2017 ■

### E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <https://my.apa.org/portal/alerts/> and you will be notified by e-mail when issues of interest to you become available!