Emotions and Identity Nonverification

Jan E. Stets[1] and Peter J. Burke[1]

Abstract

When individuals' identities are not verified, most theories and research suggest that they feel bad when others evaluate them more negatively than how they see themselves. It is less clear whether they feel good or bad when others evaluate them more positively than how they see themselves. We examine people's emotional reactions to nonverifying feedback across seven studies that include both a survey and a laboratory component. In the survey, individuals feel a little better when others slightly overrate them (an enhancement effect). *In both the survey and laboratory, individuals also feel bad for being highly overrated* (a consistency effect), *and this consistency effect overpowers the enhancement effect. The consistency effect emerges when we measure: (1) individuals' identity meanings, (2) their responses to how they think others see them* (reflected appraisals), *and (3) the meanings in the situation that are relevant to an individual's identity* (identity-relevant situational meanings).

Keywords

identity, nonverification, emotion, enhancement, consistency

INTRODUCTION

How do individuals emotionally react when they receive feedback from others that is more negative or more positive than how they see themselves? For feedback that is more negative than people's self-views, almost all social psychological theories predict that people will report negative feelings. For example, self-enhancement theory (Sedikides and Gregg 2008), self-verification theory (Swann and Buhrmester 2011), affect control theory (Robinson, Smith-Lovin, and Wisecup 2006), and identity theory (Stets 2006) all maintain that individuals will feel sad or angry that they fail to live up to their self-views. However, when individuals receive feedback that others see them more positively than how they see themselves, different predictions emerge about what individuals will affectively experience. Self-enhancement theory and affect control theory predict positive emotions; people will feel good about receiving a favorable evaluation. In contrast, self-verification theory and identity theory predict negative emotions to overly positive views; an overly positive evaluation is distressing because it is inconsistent with the way individuals see themselves.

In the current study, we examine individuals' emotional reactions to feedback

[1]University of California, Riverside, CA, USA

Corresponding Author:
Jan E. Stets, Department of Sociology, University of California, Watkins Hall, Riverside, CA 92521-0419, USA.
Email: jan.stets@ucr.edu

that is more positive than how they see themselves using data from a series of studies that had both a survey and laboratory component. Our purpose is not to test which theory works best, but rather to discover the underlying processes that influence the effects that we find. We begin with a review of previous research on two general processes that explain individuals' emotional reactions to evaluations from others: a self-enhancement process (people seek positive evaluations and avoid negative evaluations) and a cognitive consistency process (people seek evaluations that match their self-views and avoid evaluations that do not match their self-views) (Stets and Asencio 2008).

PAST RESEARCH AND CURRENT CONCERNS

Early research suggested that the enhancement process predicted individuals' affective reactions to others' evaluations of them, while the consistency process predicted people's cognitive reactions to others' evaluations (Shrauger 1975). People were more likely to feel good rather than bad when others evaluated them positively, even when these positive evaluations did not match their self-views. However, upon recall of these evaluations, they were more likely to be discredited, distrusted, and distorted. Essentially, people did not believe evaluations that were inconsistent with their self-views even though they felt _good_ following receipt of them.

A recent meta-analysis supports Shrauger's (1975) earlier findings (Kwang and Swann 2010). In undertaking this review, Kwang and Swann (2010) indicated that valid results can only be obtained if researchers measure individuals' self-views when determining whether their emotional reaction to feedback is due to enhancement or verification. Since most people hold positive self-views,

positive evaluations could lead to positive emotions either because the evaluations are positive (enhancement) or because they confirm the positive self-views (consistency). This could be circumvented if the self-views were measured.

Kwang and Swann's (2010) meta-analysis indicated that the enhancement effect was more likely to occur for affective responses to evaluations, and the consistency effect was more likely to occur for cognitive responses to evaluations. In other words, individuals had a stronger tendency to feel good following positive evaluations and bad following negative evaluations, and they had a stronger tendency to trust and accept evaluations that verified how they saw themselves and distrust and refute evaluations that did not verify how they saw themselves. To be clear, both self-verification and self-enhancement effects emerged for affective _and_ cognitive responses. It was just that the enhancement effects were stronger for affective responses than cognitive responses, and the verification effects were stronger for cognitive reactions than affective reactions.

We agree with Kwang and Swann's (2010) assessment that it is important to actually measure the self-view, but we would add that the measurements should be more precise than measuring only whether individuals hold a positive or negative self-view. Researchers should measure the meanings held in one's identity standard given a particular identity that is claimed. Essentially, degrees of positive or negative meanings in one's self-view rather than whether a positive or negative self-view is held would be important in assessing consistency. So, this is our first point: a more accurate measurement of the meanings in the identity should be undertaken to assess the effects of feedback from others relative to an identity.

Additionally, identity theory suggests two other factors that are important to

measure when determining whether people's emotional responses to evaluations from others is due to enhancement or consistency. First, the meanings of the feedback should be measured, and the measurement should be from the perspective of the person receiving the feedback. Within the context of identity theory, these meanings are *reflected appraisals*, and they figure importantly into identity verification and the emotional reactions that occur. The second factor is that *situational meanings* are relevant for the identity. These are important for understanding the activation of the identity processes involved in identity verification. Neither the reflected appraisals nor situational meanings were considered by Kwang and Swann (2010) in their meta-analysis.

Reflected appraisals are people's interpretations of the feedback they receive. It is what individuals think that others think of them. It is rooted in Cooley's ([1902] 1964) notion of the "looking-glass self" in which people see themselves reflected in the reactions of others to them. Cooley added that individuals want their self-views confirmed and shared by others in the community. When their self-views are not confirmed—that is, when people feel others do not see them in the way they see themselves—Cooley suggested people may become upset and attempt to change things.

We suggest that if prior research had measured the reflected appraisals rather than the "experimenter" assessing the feedback as positive or negative, the consistency dynamic might have had a stronger effect in how people emotionally reacted to the evaluations. If the reflected appraisals were measured on the same scales as the identity meanings themselves, the degree of consistency of the two could be directly assessed, and there may have been more negative feelings when they thought they were evaluated

more positively than how they saw themselves.

Identity theory also suggests that the definitions or *meanings in the situation* an individual takes into account have importance. The degree to which a situation carries meanings that are relevant to one's self-views (the meanings in the situation and the meanings in one's self-view share the same dimensions of meaning) should influence the likelihood those self-views are activated in the situation (Stets and Carter 2012). When the relevant identity is activated, the motivation to have that self-view verified in the situation is also activated. Thus, when self-views are activated, individuals should react positively to evaluations from others that are consistent rather than inconsistent with their self-views. In Kwang and Swann's (2010) meta-analysis, if researchers had taken into account the correspondence between the content of people's self-meanings and the relevance of the meanings in the situation, a greater correspondence between the two might have resulted in a stronger consistency effect on affective responses. People may have reported more negative feelings even when the evaluations were more positive than how they saw themselves when relevant situational meanings are present.

To recapitulate, any assessment of the relative impact of consistency and enhancement effects of feedback on the self must measure three things: the meanings of the identity, the meanings of the reflected appraisals, and the meanings in the situation. We do this in the present study of people's emotional responses to others' evaluation of them. Identity meanings capture more information as to who one is than simply whether individuals see themselves positively or negatively. Reflected appraisals capture more meaning than whether the feedback is positive or negative, and that

information is from the perspective of the person receiving the feedback. We use identity theory to guide our analysis because the theory considers identity meanings, reflected appraisals, and situation meanings as important influences on people's emotional responses in situations. Data are examined from a series of studies conducted over several years in which the moral identity, moral behavior, reflected appraisals, situational meanings, and emotional reactions were obtained. Our goal is to examine the extent of both enhancement effects and verification/consistency effects of perceived feedback from others.

THEORY

In identity theory, whether individuals feel good or bad in a situation depends upon the degree of correspondence between how individuals think they are perceived in a situation and their identity standard meanings (Burke and Stets 2009). Increasing correspondence or identity verification produces positive emotions, and noncorrespondence or identity nonverification produces negative emotions. The verification process as it relates to emotions involves three important aspects of an identity (Burke and Stets 2009). First is the *identity standard*, which contains the meanings that define the identity: what it means to be who one is. Second is the *input* or perceptions of meanings in the situation that indicate how the person is coming across in the situation given the identity the person claims. These perceptions are based on *reflected appraisals* or how persons *think* others see them in the situation. Third is a *comparator* function that serves to compare the perceived meanings of the self in the situation (reflected appraisals) with the self-defining meanings in the identity standard. This comparison is the difference in magnitude between the reflected

appraisals and the identity standard meanings (reflected appraisals minus identity standard), which may be positive, negative, or zero.

The reflected appraisal process is important in understanding individuals' emotional responses. When individuals are asked to think about how others see them, it not only encourages them to interpret the cues of others in the situation and the meanings implied by those cues, but it also encourages them to call up their identity standard so that they can compare how the meanings implied by others' responses compare with the meanings in their identity standard. Essentially, the meanings in people's identity standards come to consciousness so that they can evaluate whether their identity is being verified in the situation.

Within identity theory, when the meanings of the reflected appraisals do not match the meanings in individuals' identity standard, people experience negative emotions, and they attempt to change their behavior to counteract the discrepancy (Burke and Stets 2009). If the reflected appraisals are too high (relative to the identity standard) people act to reduce the meanings in the reflected appraisals by decreasing the strength of their behavior, for example, they may act "less kind" if they think that others see them as "more kind" than how they see themselves. If the reflected appraisals are too low, people act to increase the meanings in the reflected appraisals by increasing the strength of their behavior, for example, they may act "more kind" if they think that others see them as "less kind" than how they see themselves. Overall, the behavioral response is a function of the simple discrepancy (reflected appraisal meanings minus the identity standard meanings). Positive discrepancies result in behavior that reduces the meanings of the identity, and negative discrepancies result in behavior that

increases the meanings of the identity. Essentially, behavioral responses attempt to counteract the discrepancy, irrespective of the direction (positive or negative). This is the negative feedback loop in the identity control process (Burke and Stets 2009).

The emotional response, however, is a function of the squared discrepancy. Greater negative emotions arise from both positive discrepancies (others evaluating the person more positively than the person) and negative discrepancies (others evaluating the person more negatively than the person). Furthermore, there is a stronger emotional response as the magnitude of the discrepancy increases.

When the reflected appraisal process is captured in a situation, self-verification strivings will be more likely to operate because individuals will access their self-views in order to see how they compare with others view of them. This response follows the *accessibility principle* (Swann and Buhrmester 2011). The idea is that when individuals have the opportunity to retrieve from memory their own self-views in light of their perceptions of others' views, any discrepancy between the two is distressing. When individuals have no reason or opportunity to access their self-views in a situation, nonverifying information may simply be categorized as good or bad, and they may respond positively to the positive information and negatively to the negative information. This is the enhancement effect, and it does not involve the retrieval of one's own self-views or a comparison of self- and other-views. It only involves the reward or punishment value of the feedback itself.

With the discrepancy defined in terms of the measured reflected appraisals and with a more refined measure of the identity standard than it being simply a positive or negative self-view, we hypothesize that:

> *Hypothesis 1:* An identity discrepancy in either a positive or negative direction will be associated with negative emotions.

Implicit in the discussion of the verification process is the fact that the meanings in the situation are relevant for the identity in question. For others' feedback or the reflected appraisals to be consequential, the meanings in the situation should correspond to the meanings in the identity standard. To the extent that this correspondence occurs, we can refer to the meanings in the situation as *relevant*, or what Stets and Carter (2012) label "potent" for an identity. This means that the pertinent identity is more likely to guide behavior in the situation compared to an alternative identity. Further, when the situation has relevant meanings for an identity and the identity discrepancy that emerges is positive or negative (rather than zero), individuals should feel bad because their identity is not being verified, and we should see a consistency effect. Lacking these relevant situational meanings, a positive discrepancy may have an enhancement effect. While the feedback is positive, it is not relevant to the identity.

Recent research supports the role of situational meanings in guiding identity-relevant behavior and emotions. Stets and Carter (2012) not only found that the moral identity was activated in a moral situation, but a high moral identity was positively associated with acting in a moral manner rather than an immoral manner. When individuals behaved immorally, they were more likely to report the negative emotions of guilt and shame when the situation meanings were more rather than less morally relevant. Thus, internal identity meanings and external situational meanings coalesced to influence moral action and feelings. Therefore, we expect that:

Hypothesis 2: The more relevant the meanings in the situation to the identity, the greater the effect of identity discrepancy in either a positive or negative direction on negative emotions.

In summary, we are interested in the conditions under which the consistency dynamic rather than the enhancement dynamic predicts people's emotional reactions to feedback from others. We think this dynamic will be more likely to emerge when meanings in the situation are very relevant for people's identities. All of these meanings need to be measured including the meanings contained in people's identity standard, the meanings regarding how they think others see them in the situation (reflected appraisals), and the meanings that they think are revealed in the situation in terms of how they should behave. Prior research has tended to neglect the alignment of these meanings when examining people's emotional reactions to others' evaluations of them. The current studies that we analyze in this article include these relevant features. We measure individuals' identity standard meanings and reflected appraisal meanings, and we calculate a discrepancy or difference. In addition, we measure the relevance to the identity of the meanings in the situation since the strength of the response to the discrepancy should vary by the degree to which there are relevant meanings to be perceived and to which a response is needed.

METHOD

Overview

The data for this research were obtained from a series of seven studies with the same basic design conducted over several years, each having two parts and each part facilitating an examination of whether identity nonverification is

associated with positive or negative emotions.[1] The first part of each study required participants to respond to a survey. In the second part, individuals participated in a laboratory study several weeks later. The survey measured the moral identity of the respondents, and this measure was used in both parts of the study. The separation of several weeks between the two parts (the survey study and the lab study) was to reduce the likelihood that participants in the lab study would make a direct connection between the two parts.

In addition to the measurement of the moral identity standard, the surveys provided eight different moral situations that the respondents were likely to have experienced (see appendix). For each situation, participants indicated what action they took in the situation, how they thought others would rate them in the situation (reflected appraisals), and how they felt emotionally. If participants had never experienced the situation, they were to imagine the situation, how they thought they would respond, how they would assess others' views of them in that situation, and how they thought they would feel. The first column in the appendix provides the percentage of individuals who actually did experience each of the moral situations.

While prior research investigating moral decisions typically has relied on a class of hypothetical moral dilemmas known as the trolley car problems

---

[1]Each of the studies examined the effects of the moral identity on cheating behavior either through commission (changing one's answers) or omission (failing to report being overscored). In addition, studies varied pretesting the measurement of the moral identity and some varied the posttesting of the moral identity. Aside from these basic design differences, the studies are the same. Some of the results from the first two studies are reported elsewhere (Stets 2011; Stets and Carter 2011, 2012).

(Greene et al. 2001; Hauser 2006), this research relies on real-life situations that individuals have experienced (Walker et al. 1995). The items were obtained from a 2006 study in which 150 undergraduates were asked to identify three recent situations where they had the choice between doing the right or the wrong thing. The most frequently listed situations were identified and incorporated into the present survey. Thus, we have situations that are relevant to the population from which our sample is drawn. These situations do not represent all of the moral situations they could encounter; rather, they represent the situations that they typically encounter in their everyday lives.

In the laboratory, participants were placed in a testing situation in which they had an opportunity to cheat without clear detection to obtain a higher score. Following the testing situation, participants were asked to reflect on how they thought others would see them in the current situation and how they felt. Thus, while the survey in each study operationalized participants' reactions to their behavior in a retrospective manner, the laboratory setting operationalized their reactions to their behavior immediately following the opportunity to cheat. By using the same participants in both the survey and the laboratory, we remove the possibility of differences between the two parts as due to differences in the participants.

The design of the laboratory studies had two different types of moral behavior: cheating as an act of *commission* (changing answers on an exam to improve one's score) and cheating as an act of *omission* (failing to report being overscored on an exam). As we discuss in the following, committing a bad act may introduce moral meanings into the situation more than failing to commit a good act because doing something bad in the situation likely draws more attention than failing to do a good act, which involves no action (Stets 2011). Before outlining these two conditions, we review the protocol that was common to both.

Study Procedure

A lab assistant ushered participants one at a time into a room with a computer. Participants were told that they would be taking a test for a new version of a college entrance exam similar to the Scholastic Aptitude Test (SAT). The company sponsoring the test was interested in how well participants could perform on this new test. The test had 12 questions (six verbal questions and six math questions). Each question was worth 10 points. After answering each question, a feedback page appeared on the computer screen that reported the participant's answer, the correct answer, and the participant's points to date. Participants were told that everyone's scores would be tabulated at the end of the study.

Before the test began, the participants were reminded that they could earn $100 in a lottery for their participation. However, a cover story provided an incentive for them to be dishonest. They were told that the company sponsoring the study wanted to extend their appreciation to the participants for their help by entering four additional tickets per person into the $100 lottery. This increased each person's odds of winning the lottery. However, if a person's total points on the exam were not in the top 25 percent of all tabulated scores, the person would lose two of the extra tickets. If a person's points were not in the top 50 percent, all four extra tickets would be forfeited. Thus, it was in each person's interest to score high on the test.

Immediately before the test began, the lab assistant wanted to ensure that the

Table 1. Distribution of Sample by Gender and Laboratory Study

Lab Study	N	Percentage Male	Condition
1	546	42	Commission
2	473	42	Omission
3	367	41	Omission
4	287	33	Commission
5	340	43	Omission
6	466	26	Commission
7	540	37	Omission
Total	3,019	38	

participants understood how many tickets would be taken away if their total points were not in the top 25 percent or 50 percent. Participants were instructed to fill out a short questionnaire that reviewed this information. If they did not answer each question correctly, the assistant returned the review sheet, explained the ticket situation again, and asked them to answer the questions again.

After finishing the "test" and answering some final questions, the participants were debriefed. They were informed that their test score would not be tabulated with all other scores, and they would not earn additional lottery tickets. Rather, they earned one lottery ticket in exchange for their participation. Before exiting the study, they were asked if they knew what the study was about. A small percentage (12 percent) identified the true purpose of the study. Analysis of these data showed that those who identified the true purpose of the study did not produce statistically different results than those who did not identify the study's true purpose.[2]

Condition 1: commission. The commission condition occurred in three of the seven studies as indicated in Table 1. It

follows a protocol used by others (Kalkhoff and Willer 2008). Before the research assistant reviewed the incentive structure regarding test scores and the extra raffle tickets, the assistant indicated that the computerized test needed to be checked to see if it was working properly. There was really nothing wrong with the test. This simply provided the opportunity to show participants how to change their answers. The assistant went through a "sample question" with participants that was already posted on the computer screen. The assistant read the question aloud and then deliberately gave participants a wrong answer to enter on the computer. On the feedback page, the assistant engaged participants to "play along" and "discovered" that the "escape" key could be used to go back and change one's response. Upon this discovery, the assistant acted shocked and informed them not to use the "escape" key for this purpose. The assistant went through two more sample questions with them. For one of them, the assistant repeated the use of the "escape" key to change an answer. The assistant acknowledged the ability of the program to allow them to change their answers but indicated that they should not do that. The assistant next reviewed the incentive structure to obtain extra raffle tickets and then left the room for participants to begin the exam.

Condition 2: omission. The omission condition occurred in four of the seven studies as indicated in Table 1. Before the assistant reviewed the incentive structure, participants were alerted to a worksheet that was placed at their workstation. They were told that it was sometimes difficult for people to keep track of their score because they were concentrating too much on the exam questions. Therefore, to keep track of their score, they were instructed to use

_____
[2]These results are available upon request.

the worksheet, and all of them did. The worksheet provided a line for each question, whether they got that question correct, the previous points they earned, and their current point total. This information was provided on the computer screen, and participants used this to update their worksheets as they went through their exam.

There were also two additional lines at the end of the worksheet that were filled out after the test. The first line asked them to review (from the 12 questions on the worksheet) how many correct answers they received and then multiply this number by 10 for their total number of points. The second line asked them to provide the total points that the computer provided. Comparison of the two lines would show the discrepancy in the scores. Essentially, this worksheet helped ensure that participants noticed their over-scoring. Next, the assistant reviewed the incentive structure. Then the assistant left the room so the participant could take the exam.

All participants in this condition were overscored 20 points three times: after Questions 3, 8, and 11. After each of these questions, the computer indicated that their answer was correct irrespective of how they answered the question, and they received 30 points rather than 10 points. Thus, they earned 90 points on these three questions. This was an extra 60 points for those who answered all questions correctly and up to an extra 90 points for those who answered some of the questions incorrectly. After the test, participants were reminded that the company sponsoring the study was still working on the computer program, and they wanted to know if they noticed errors in the program such as grammatical errors, answers that were not correct, scoring problems, unclear questions, confusing formatting, or anything else. This is where participants could report that they were overscored.

Commission versus omission. In committing a bad act as in changing one's answer, individuals are acting upon the environment. The action draws their attention given the informational value of an action compared to the omission of an action (Allison and Messick 1988; Fazio, Sherman, and Herr 1982; Ross 1977). Omitting a good act as in not reporting being overscored is a form of decision avoidance. Rather than making a choice to act or engage in a bad behavior, omission postpones action by doing nothing (Anderson 2003). The reduced attention to doing nothing means that individuals will be less inclined to interpret, store, and retrieve their non-action from memory. It's as if nothing happened. Thus, failing to report being overscored in the omission condition carries less meaning of having done something bad compared to actively cheating in the commission condition.

Another difference between the commission and omission conditions is that because people in the commission condition have 12 opportunities to cheat, once for each question in the "test," there is a heightened level of moral meanings that may be present in this situation. It would be hard to ignore the moral implications of cheating or resisting the desire to cheat on each question. In the omission condition, however, there is only one opportunity to cheat by not reporting the over-scoring at the end of the test. With only one decision, which was buried in other potential problems with the test, participants might find it easier to not attend to the moral meanings in their behavior.

Finally, because commission compared to omission is seen as more intentional (Spranca, Minsk, and Baron 1991), the self is more likely to be implicated as the cause of the outcome. In contrast, if a person does nothing, the situation is ambiguous as to the cause of the outcome, so it is

easier to attribute the cause to an alternative source rather than implicate the self. Consequently, if the act of cheating (commission) carries more moral meaning than not reporting being overscored (omission), and commission implicates the self to a greater degree than omission, the greater moral meanings in the commission situation will be more relevant to the moral identity than the fewer moral meanings in the omission condition. Affective responses to consistency in the commission condition should therefore be stronger that in the omission condition. This is consistent with Hypothesis 2 presented earlier.

Sample

Participants were recruited from undergraduate classes at a large southwestern university from 2007 to 2009. For their participation, students were offered extra credit in their classes and a chance to win $100 in a lottery. Table 1 shows the number of participants for each study and the gender distribution. A total of 3,019 individuals participated in the seven studies. As shown in Table 1, 38 percent of the participants were men and 62 percent were women. The average age was 21 years. Respondents were ethnically diverse with 15 percent white, 28 percent Latino, 34 percent Asian, 7 percent black, 6 percent multiracial, and 10 percent other. The average parental income was $35,000 to $49,000. The analysis of the survey data is based only on the 3,003 respondents who indicated that they had experienced a situation in the survey (see percentage for each situation in the appendix).[3] Only responses to situations the respondent actually experienced

---

[3]However, an analysis of all respondents showed no significant difference between those who had experienced any of the situations and those who had not experienced any of the situations. These results are available upon request.

were included. Hence, these data, though retrospective, are not hypothetical.

Measures

Participants' moral identity standard meanings and the reflected appraisals (how individuals think others are evaluating them in moral terms) were obtained in the surveys. The identity discrepancy was calculated as the reflected appraisals measure minus the identity standard measure. For the eight moral situations in the survey, we measured the reflected appraisals (how individuals think others evaluated them in each situation) and their emotional responses following these evaluations. We also measured how relevant each of the eight situations were to issues of morality. In the laboratory studies, we measured the reflected appraisals and the emotional responses after the participants completed the "test." We also included a binary variable, *commit*, that was coded 1 for the lab studies in which the participant had an opportunity to commit an act of cheating and 0 for the lab studies in which the participant could cheat by failing to report the over-scoring.

Moral identity standard. The self-meanings of the moral identity along the dimensions of care and justice were operationalized using 12 bipolar characteristics that have been used in prior studies. Items measuring care included being caring/uncaring, unkind/kind, helpful/not helpful, stingy/generous, compassionate/hardhearted, friendly/unfriendly, and selfish/selfless; items measuring justice included being honest/dishonest, unfair/fair, untruthful/truthful, not hardworking/hardworking, and principled/unprincipled (Aquino and Reed 2002; Stets and Carter 2012; Walker and Hennig 2004). Respondents were to think about how they saw themselves with respect to each characteristic and identify where

they would place themselves between each bipolar characteristic. Responses ranged from 1 to 5, where 1 reflected agreement with one bipolar characteristic, 5 reflected agreement with the other bipolar characteristic, and 3 placed the respondent in between the two bipolar characteristics. A factor analysis of the items revealed a single factor. Negatively worded characteristics were reverse coded; the items were standardized and summed with a high score representing a high moral identity. The final scale was standardized (mean of 0 and variance of 1) with a high score representing a high level of morality. The omega reliability for the scale was .90.

Reflected appraisals. To evaluate how participants thought others saw them in moral terms, two slightly different measures were used. In the first five studies, after each moral situation in the survey and after they completed the test in the laboratory, participants were asked how they thought others saw them along a set of characteristics including being likeable, intelligent, moral, dominant, spiritual, attractive, reliable, loving, and competitive. The wording was "In this situation, how do you think others would rate you as being [adjective]?" The scale responses ranged from 0 = not at all to 6 = very much. We are interested in their responses to how they thought others saw them as a moral person (a single item). The other characteristics were filler items. This reflected appraisal measure also was standardized (mean of 0 and variance of 1), thus having the same scale of standard deviation units as the identity standard measure.

In the last two studies, the reflected appraisal measures in the surveys and in the laboratory were changed to parallel the measurement of the moral identity standard. We came to realize the importance of measuring the two meanings in

the same way so that the comparison was of two measures with the same meanings and scale characteristics. In order to obtain feedback on the same dimensions of meaning as the identity, following their behavior, participants were administered the same 12 bipolar items used to measure the moral identity standard.

Moving from a single item to 12 items also increased the reliability of the measure. They indicated on a scale of 1 to 5 how they thought others would rate them on each of the characteristics where 1 reflected agreement with one bipolar characteristic, 5 reflected agreement with the other bipolar characteristic, and 3 placed the respondent in between the two bipolar characteristics. Each of the 12 items was then standardized, aligned so that a high score represented a high level of morality, and summed. The resulting scale score was then standardized to a mean of 0 and variance of 1, again, matching the identity standard scale. The average omega reliability across the different scenarios for the reflected appraisals scale was .97 in the surveys and .92 in the laboratory studies. We think future research should move toward a measure such as this, which has a higher reliability and assures that the same dimensions of meaning are used for both the identity standard and the reflected appraisals.

Moral identity discrepancy. The moral identity discrepancy was measured in two different ways: as a simple (linear) difference between the reflected appraisals and the moral identity standard and as a squared difference. For both measures, the moral identity score was first subtracted from the reflected appraisals score. For the simple (linear) difference, a positive result meant that participants perceived that others were evaluating them more positively than they viewed themselves along the moral dimension.

A negative result meant that participants perceived that others were evaluating them more negatively than they viewed themselves in moral terms. The linear difference is used to test the enhancement effect, which predicts that the more positive the difference between reflected appraisals and the identity standard, the better people will feel. The negative emotions will be reduced and positive emotions will be increased.

The squared difference was the square of the linear difference. A departure of the difference from zero in either a positive or negative direction meant an increased discrepancy between the moral identity standard and the reflected appraisals. The squared difference is used to test the consistency effect, which predicts that differences in either the positive direction or the negative direction will increase negative emotions. In identity theory, the *emotional* response to an identity discrepancy is governed by the squared discrepancy. Putting both the linear difference and the squared difference in the prediction equations allows us to test the extent of both enhancement and consistency effects on participants' emotional responses.

Although the last two studies used a different measure of reflected appraisals that goes into the measure of moral identity discrepancy, because both procedures standardized the resulting measures, there were no differences in means or standard deviations between the two measurement procedures.

Moral relevance. The survey measured how relevant to the moral identity situational meanings were in each of the situations. For this, 380 respondents drawn from the same population as the seven studies were asked to indicate how moral they saw each of the eight situations, however they defined morality. Half of the situations were worded in terms of

doing a good behavior (e.g., one does not allow another student to copy from one's own exam), and the other half of the situations were worded in terms of doing a bad behavior (e.g., being drunk and driving oneself home). Responses ranged from –5 = extremely immoral, for bad behaviors through 0 = neutral) to 5 = extremely moral, for good behaviors. We took the absolute values of this variable to capture the degree of moral relevance with scores ranging from 0 (moral relevance is low, that is, the situation was seen as neither moral nor immoral) to 5 (moral relevance is high, that is, the situation was seen as extremely moral or extremely immoral). The mean rating of each situation then was obtained as the measure of the moral relevance of the situations.

The mean ratings ranged from a low of 2.12 for allowing another student to copy one's exam to a high of 3.87 for letting a friend drive home drunk. The relevance values for each situation are in the second column of the appendix. This variable was standardized for the analyses to have a mean of 0 and a variance of 1. Given the earlier discussion of commission and omission, we assume that having the opportunity to commit a bad act will signal stronger moral meanings in the situation and be more relevant for the moral identity. However, not all situations are clear along the dimension of commission and omission. Further, commission and omission are not the only determinants of moral relevance. The potential outcome may also be important with acts capable of generating more serious harm as more morally relevant. This is an avenue for future research.

Negative emotions. Participants' emotions were measured in two slightly different ways across the seven studies. In all the studies, following each situation (in the survey) and after completing the

"test" (in the laboratory), participants were to report how they were feeling. In the first study, participants were to select the emotion that best represented how they felt from a list of eight emotions: happiness, fear, disgust, anger, sadness, shame, guilt, and empathy. The first five emotions are primary emotions and the remaining three are secondary emotions (Turner and Stets 2005). If they felt more than one emotion, they were to report the one they felt was the strongest. The emotion they identified was coded 1 if it was a negative emotion (fear, disgust, anger, sadness, shame, or guilt) and −1 if it was a positive emotion (happiness or empathy). After identifying an emotion, they were to report the intensity of the feeling from not at all intense to very intense (coded 0–6). The response to this intensity item was then multiplied by plus or minus 1 for the valence of the emotion. The resulting scores could range from −6 (intense positive emotion) to 6 (intense negative emotion). The final negative emotion scale was standardized with a mean of 0 and variance of 1. In general, this emotion measure is adapted from ones used by others, for example, Heise and Calhan (1995).

In the six remaining studies, an improved and more reliable measure of emotions was obtained that used a multi-item scale. Following each moral situation in the survey and behavior in the laboratory, participants indicated how strongly they felt each of six emotions: happy, fearful, angry, sad, shameful, and guilty. Response categories ranged from not at all intense to very intense (coded 0–9). Factor analysis of the items showed a single dimension. Happiness was reverse coded. All the items were then summed to create a negative emotion scale. The scale had an omega reliability of .86. The scale was then standardized with a mean of 0 and variance of 1. Because the resulting measures were standardized, there were no differences in the means and standard deviations across the different procedures for measurement.

Control Variables

Preexisting negative emotions. Emotions in situations are an outcome of the identity verification process, but they are also an outcome of an individual's earlier emotional state. People enter situations already in a particular feeling state given experiences from earlier in the day, yesterday, or perhaps longer (Stets and Osborn 2008). Thus, we control for emotions that participants may have brought into the laboratory from previous encounters. In all seven laboratory studies, when participants initially entered the laboratory, they were administered the same emotion scale that they received at the end of the study, after completion of the "test."[4] Additionally, at the beginning of the survey, the respondents filled out the same emotions scale they completed after each of the scenarios. These prior feelings were coded and scaled in the same way as the emotions discussed previously to have a mean of 0 and variance of 1.

Moral behavior (bad act). Emotions in situations may also be an outcome of participants' behavior, thus we control for this influence. Across all the studies, behavior was assessed by participants' responses to each moral situation in the survey in terms of how they behaved and whether they cheated or failed to report being overscored in the lab. In the survey, when participants reported

---

[4]In the first laboratory study, only half of the participants received the pre-emotion questionnaire as the impact of this questionnaire on subsequent aspects of the study was being investigated. This reduced the N in the laboratory studies from 3,003 to 2,743.

doing the "right" or "good" behavior, it was coded 0; "bad" or "wrong" behavior was coded 1. For example, responses were coded 0 when participants "donated to a charity" or "did not copy a student's answer." Alternatively, responses were coded 1 when participants "did not donate to charity" or "copied a student's answer." Participants' behavior in the laboratory was coded 1 if they did cheat on the exam or did not report being overscored and 0 otherwise.

Demographic characteristics. To investigate whether the findings varied by the characteristics of the sample, we control for several demographic characteristics: gender, race, and parental income. Gender was coded 0 for female and 1 for male. Race was measured with six binary variables that were coded either 0 (if the person was not a member of the category) or 1 (if the person was a member of the category). The race variables are white, Latino/Chicano, Asian, black, multiracial, and other. White was the omitted variable in the analyses. Income was based on participant's parents' total income. The values were coded at the midpoint of the categories with under $5,000 (coded $3,000), $5,000–$9,999 (coded $7,500), $10,000–$14,999 (coded $12,500), $15,000–$24,999 (coded $17,500), $25,000–$34,999 (coded $30,000), $35,000–$49,999 (coded $42,500), $50,000–$75,999 (coded $62,500), and $75,000 and over (coded $90,000).

ANALYSIS

To examine the effects of identity discrepancy (the difference between reflected appraisals and the identity standard) on emotion, we use ordinary least squares (OLS) regression in which participants' emotion is regressed on the linear discrepancy (difference), the squared discrepancy, and the control variables. In the survey, because each person responded to up to eight situations,

the error terms on these responses would not be independent. Therefore, the regression analysis assumed there were correlations among the error terms across situations within persons but not between persons.

A significant negative linear effect indicates that as the reflected appraisals exceed the identity standard, the less negative or more positive is the emotion that will be felt. Conversely, the more the reflected appraisals fall below the identity standard, the more negative emotion will be felt. This is the enhancement effect. A significant positive squared effect indicates that an identity discrepancy in either a positive or negative direction increases negative emotion. This is the consistency effect. Both the linear and the squared discrepancies are entered into each equation, and it is possible that both effects are present.

We also hypothesized that the effects of the identity/reflected appraisals discrepancy on emotion would be stronger in the surveys to the extent that the situation is high in moral relevance (that is, the extent that there were identity relevant meanings in the situation). To test this in the survey data, we included two interaction terms in the regression equations. The first is the product of relevance and the linear discrepancy; the second is the product of relevance and the squared discrepancy. Of course, relevance itself is included in these equations.

In the analysis of the laboratory data, we include a binary variable for the condition (commission vs. omission) as well as interactions of the discrepancy variables with the laboratory condition. This allows us to examine differences in results between the two conditions. We hypothesized that there would be more moral identity-relevant meanings in situations of commission of a "bad" act than in situations of omission of a "good" act and that

the effects of discrepancy would be stronger.

Because our analysis combines data from seven laboratory and survey studies, it is important to know whether the coefficients that are estimated from one study are the same as the coefficients that are estimated from the other studies or whether they differ by study. Tests of these equivalences of coefficients across studies are made, and they show no significant differences. Thus, the estimates of the effects can be pooled across studies.

RESULTS

The means, standard deviations, and ranges of all the variables are listed in Table 2. Table 3 provides the correlations among all of the variables in the surveys, and Table 4 provides the correlations among all of the variables in the laboratory studies.

We first turn to the analysis of the survey data. The results are presented in Table 5. Recall that in the surveys, participants responded to multiple moral situations, thus we allowed the error terms in the regression to be correlated across situations within persons. The analysis is based on 3,003 clusters (individuals) of 17,479 responses for which there is complete data. The analysis regressed negative emotions on each of the independent variables, the interactions of those variables with the linear and squared discrepancies, and the control variables.[5]

The findings reveal the expected positive effect of the squared discrepancy on negative emotion in accord with Hypothesis 1, that is, a discrepancy in either direction is positively associated with negative emotion. This is a consistency effect. Reflected appraisals that exceed or fall below one's own identity standard produce negative emotion. However, the linear component is also significant. Reflected appraisals that exceed one's own moral identity are associated with a reduction in negative emotion. This is an enhancement effect. Thus, it appears that both enhancement and consistency effects are present in the survey data.

However, when we combine the linear component with the squared component of the discrepancy to examine the overall effect, as shown in Figure 1, we see that the combined effect of both components is a U-shaped curve that is shifted slightly to the right on the x-axis.[6] The linear (enhancement) effect, while making people feel less negative for small over evaluations, is dominated by the squared (consistency) effect when the level of over-evaluation increases. A small amount of over-evaluation helps, but given the squared effect, beyond this small amount of over-evaluation, negative emotions predominate. At first glance, then, there appears to be some truth to both the enhancement and consistency hypotheses. But, a closer examination reveals that emotional reactions to feedback (reflected appraisals) are governed by a U-shaped curve: a consistency effect.

We expected the consistency effect to be stronger in situations that have greater moral relevance. This is Hypothesis 2. Table 5 shows support for this. The strength of the negative emotional

---

[5]Because we pooled data from seven different survey studies, we tested whether there were variations in results across the studies. The interactions by study were not significant, thus allowing us to pool the estimates across studies. Additionally, interactions of the main effects by the control variables were tested and were not significant. None of the control variables moderated the effects, which are therefore robust across groups. These results are available upon request.

[6]The curve in Figure 1 combines the linear and quadratic effects: $Y = -.09D + .10D^2$.

Table 2. Means, Standard Deviations, and Ranges of Variables

Variables	Mean	Standard Deviation	Minimum	Maximum
Surveys (N = 3,003 clusters; 17,479 responses)				
Negative emotion	.00	1.00	−1.05	5.80
Linear difference	.00	1.00	−4.34	3.78
Square difference	.00	1.00	−0.79	10.35
Moral relevance of situation	.00	1.00	−.81	1.57
Do bad act	.34	.48	.00	1.00
Laboratory studies (N = 2,743)				
Negative emotion	.00	1.00	−.63	5.96
Linear difference	.00	1.00	−3.51	2.93
Square difference	.00	1.00	−.65	7.41
Pre-emotion	.00	1.00	−.75	6.18
Cheat	.44	.50	.00	1.00
Commit condition	.37	.48	.00	1.00
Surveys and laboratory studies (N = 2,743)				
Gender (male)	.37	.48	.00	1.00
Race				
White	.15	.36	.00	1.00
Latino	.28	.45	.00	1.00
Asian	.35	.48	.00	1.00
Black	.07	.26	.00	1.00
Multiracial	.06	.23	.00	1.00
Other	.10	.29	.00	1.00
Family income	48.97	31.72	3.00	90.00

Table 3. Correlations Among the Variables for the Surveys (N = 3,003 clusters; 17,479 responses)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Negative emotion	1.00													
2. Pre-emotion	.17*	1.00												
3. Linear difference	−.32*	.06	1.00											
4. Square difference	.24*	−.02	−.27*	1.00										
5. Do bad act	.46*	.02	−.48*	.24*	1.00									
6. Moral rel.[a]	—	—	—	—	−.21*	1.00								
7. Gender (male)	−.01	−.01	.05*	.01	.06*	—	1.00							
8. White	−.07*	−.06	.03	−.01	−.03	—	.06*	1.00						
9. Latino	.03	.01	−.09*	.03	−.02	—	−.14*	−.27*	1.00					
10. Asian	.05*	.01	.06*	.00	.06*	—	.13*	−.31*	−.44*	1.00				
11. Black	−.03	.02	.01	−.01	−.01	—	−.08*	−.12*	−.17*	−.20*	1.00			
12. Multiracial	−.03	−.01	−.01	.00	−.01	—	−.01	−.11*	−.16*	−.18*	−.07*	1.00		
13. Other	.00	.03	.00	−.01	−.02	—	.00	−.14*	−.20*	−.23*	−.09*	−.08*	1.00	
14. Income	−.05*	−.04	.01	−.02	−.02	—	.10*	.20*	−.18*	−.06*	−.01	.06*	.08*	1.00

[a]Moral relevance was a characteristic of the situation and did not vary by participant.
*$p \leq .05$.

reaction to the squared discrepancy is much stronger for morally relevant situations (an increase of .07 for situations that have a moral relevance of 1 standard deviation above the mean, or a decrease of .07 for situations that have a moral

Table 4. Correlations Among the Variables for the Laboratory Studies (N = 2,743)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Negative emotion	1.00													
2. Pre-emotion	.68*	1.00												
3. Linear difference	.02	.07*	1.00											
4. Square difference	.07*	.03	−.17*	1.00										
5. Cheat	.08*	.03	−.06	.01	1.00									
6. Commit	.03	.00	−.01	.02	−.20*	1.00								
7. Gender (male)	.01	−.01	.07*	.04	−.04	−.09*	1.00							
8. White	−.08*	−.05	.03	−.02	−.08*	.00	.05	1.00						
9. Latino	.02	.00	−.05	.01	.00	.03	−.14*	−.26*	1.00					
10. Asian	.03	.01	.05	.03	.05	−.02	.13*	−.31*	−.45*	1.00				
11. Black	−.01	.02	.01	−.03	.05	−.01	−.08*	−.12*	−.17*	−.21*	1.00			
12. Multiracial	.00	−.01	−.03	.01	−.05	−.01	−.01	−.10*	−.15*	−.18*	−.07*	1.00		
13. Other	.02	.04	−.01	−.01	.01	.00	.01	−.14*	−.20*	−.24*	−.09*	−.08*	1.00	
14. Income	−.04	−.03	.02	−.03	−.07*	−.03	.10*	.20*	−.16*	−.07*	−.01	.06	.08*	1.00

*$p \leq .05$.

Table 5. Standardized Regression Coefficients for Negative Emotion in the Surveys (N = 3,003 clusters; 17,479 responses)

	Emotion
Independent variables[a]	β
Linear difference	−.09**
Square difference	.10**
Pre-emotion	.17**
Moral relevance	.12**
Bad act	.40**
Gender	−.06**
Race	
Latino	—
Asian	—
Black	—
Multiracial	—
Other	—
Income	−.03**
Linear × relevance	—
Linear × gender	—
Linear × race	—
Linear × income	—
Square × relevance	.07**
Square × gender	—
Square × race	—
Square × income	—
R^2	.28**

[a]Interactions with study were not significant.
**$p \leq .01$; —, not significant.

relevance 1 standard deviation below the mean), but there is no effect of relevance on the linear difference. Because the relevance of the moral meanings in the situation moderates the effect of the squared discrepancy but not the linear discrepancy, when there are few moral meanings in the situation, the enhancement effect remains strong while the consistency effect diminishes.

We can see this in the "Low Relevance" line of Figure 2 (relevance is 1 standard deviation less than average). It approaches a straight line with a negative slope (indicating a more enhancement effect). However as the relevance of the moral meanings in the situation increases, the squared discrepancy increasingly dominates the linear (enhancement) effect, and the effect approaches a consistency effect. We can see this in the "High Relevance" line in Figure 2 (relevance is 1 standard deviation higher than average). There is a clear and strong U-shaped effect indicating increased negative emotions for reflected appraisals that are above or below the identity standard. Therefore, moral meanings in the situation coalesce with moral identity meanings to produce a strong consistency effect on affective reactions to identity discrepancies.
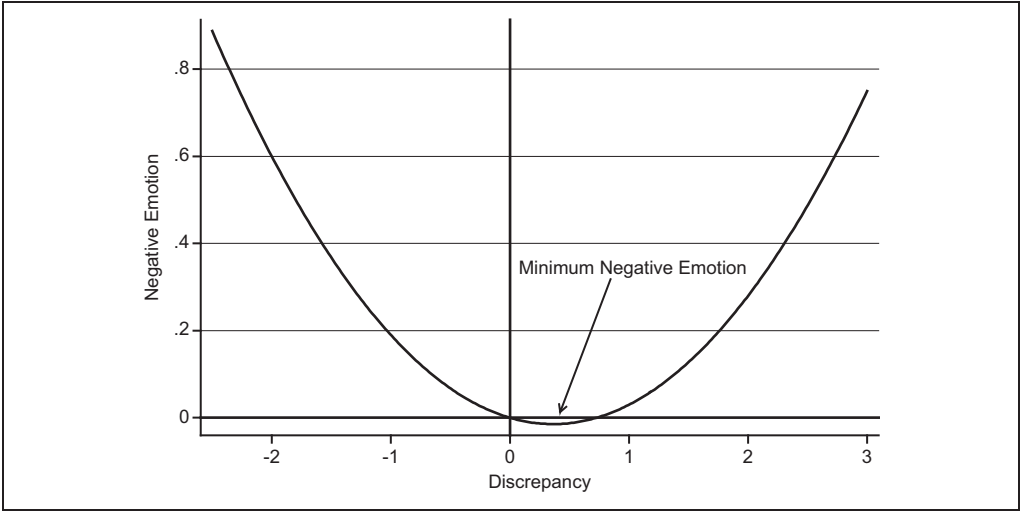
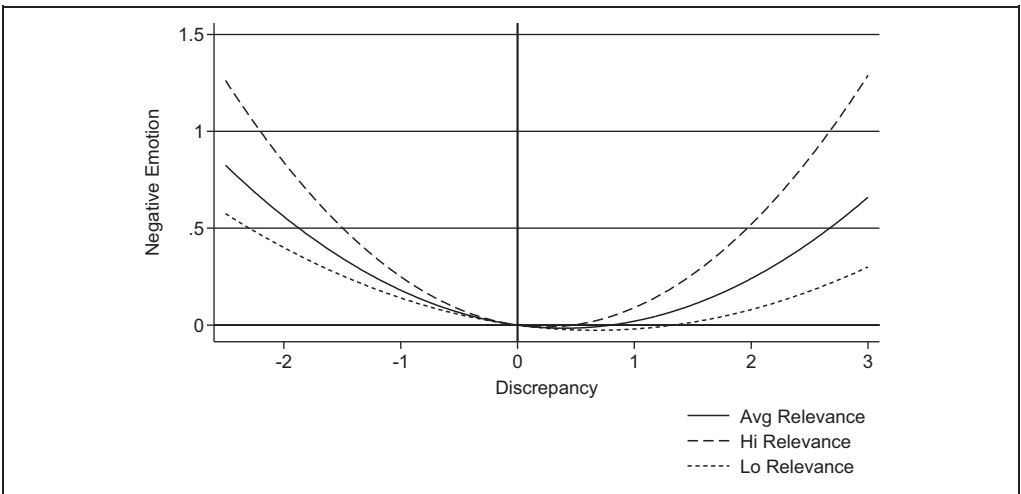**Figure 1.** Negative Emotion in Survey Studies for Combined (Linear + Squared) Discrepancy

Figure 2. Negative Emotion in Survey Studies for Combined (Linear + Squared) Discrepancy by Level of Relevance (Average $\pm$ 1 SD) of Meanings in Situation to the Moral Identity

Finally, Table 5 reveals effects for several of the control variables. Pre-emotions are positively related to the emotional response. A situation with high moral relevance and the enactment of a "bad" act in a situation both increase negative emotion overall. Women and those with a low income report more negative emotions than men and those with a high income. None of the effects of discrepancy were moderated by the control variables of gender, race, and income. Thus, the results are robust across these groups.

We now turn to the analysis for the laboratory studies. This analysis regressed

Table 6. Standardized Regression Coefficients for Negative Emotion in the Laboratory Studies (N = 2,743)

Independent Variables[a]	β
Linear difference[b]	—
Square difference[b]	—
Pre-emotion	.70**
Cheat[b]	—
Commit	.07**
Linear × commit	—
Square × commit	.10**
Pre-emotion × commit	−.07**
Cheat × commit	.10**
R^2	.49**

[a]Effects of control variables gender, race, and income are not significant.
[b]Interaction effects with study, gender, race, and income were not significant.
**$p \leq .01$; —, not significant.

negative emotions on each of the independent variables as well as interactions of those variables with the "commit" variable (coded 1 for those in the commission of a bad act and 0 for those in the omission of a good act). Also included in the regression equation were the control variables of race, gender, and income. The results are presented in Table 6.

The baseline findings show that there is no effect for either the linear discrepancy or the squared discrepancy. These baseline effects apply only to the omission condition because of the significant interactions with the commit variable. To see the effects in the commission condition, we add together the zero-order effect (which is not significantly different from zero) and the interaction effect. Thus, for the commission condition, we see that the linear discrepancy still has no effect, but the squared discrepancy does have a significant effect of about .10 ($p \leq .01$). This confirms the consistency effect but disconfirms the enhancement effect, thus supporting Hypothesis 1, but only in the commit condition. Comparing conditions shows that the effects of a discrepancy

only holds in the commission condition where we expected the relevance of the moral meanings to be greater and thus the discrepancy of the moral identity to have an influence in how individuals felt. This is consistent with Hypothesis 2.

With respect to the control variables, the pre-emotions are positively related to negative emotions. Negative feelings that are reported upon arrival at the laboratory are significantly related to negative feelings after completion of the test in both conditions. In this way, previous emotions continue to have an effect long after they have been originally activated. Cheating also is significantly associated with negative feelings, but only in the commission condition. Finally, the negative emotions do not significantly vary by the characteristics of the sample, thus the results are quite robust. The control variables do not influence negative emotions nor do they moderate the effects of a discrepancy (linear or squared) on negative emotions. In general, the results, pooled across the laboratory studies, show a strong consistency effect (for the squared difference) and no enhancement effect (for the linear difference) on negative emotion in the commission of a bad act (but not the omission of a good act).[7]

Taken together, the survey results and laboratory results across the seven studies reveal that consistency rather than the enhancement effects describes individuals' affective response. People responded negatively to feedback that is more (and less) positive than their identity. We think that what facilitates the consistency effect is giving people the opportunity to retrieve their own

----

[7]Because we pooled data from seven different laboratory studies, we tested whether there were variations in the results across the studies. The interactions by study were not significant, thus allowing us to pool the estimates across studies. These results are available upon request.

self-views (identity) in light of their per-
ceptions of others' views (reflected
appraisals) and comparing the two. A pos-
itive discrepancy, when more than mini-
mal, will be distressing. Additionally,
meanings in the situation that are rele-
vant to the identity that is implicated in
the situation will make the discrepancy
between self and others' views all the
more consequential. Thus, the consis-
tency effect on affective reactions is condi-
tioned on at least these factors being pres-
ent in the situation.

DISCUSSION

In this article, we used a large data set
derived from seven studies that included
both a survey and laboratory component
to address the emotional responses
that occur when identities are not veri-
fied. We examined whether individuals
showed an enhancement response (they
felt good) or consistency response (they felt
bad) to identity nonverification in a positive
direction (the meanings in the reflected
appraisals are more positive than the mean-
ings of the identity standard). We suggested
that the lack of consistency effects in prior
research may have been due to not rou-
tinely including identity standard mean-
ings, reflected appraisal meanings, and rel-
evant identity meanings in the situation.

By using both a linear discrepancy and
a squared discrepancy in our analysis, we
measured the degree of both the enhance-
ment effects (positive feelings for over
evaluation and negative feelings for
under evaluation) and consistency effects
(negative feelings for both over and under
evaluation) and tested the strength of
each of these responses. We also tested
the extent to which the meanings rele-
vant to the identity in the situation
moderated the possible enhancement
and consistency effects.

The results showed a small enhance-
ment effect in the surveys but not in the

laboratory studies. The findings also
showed a strong consistency effect in the
surveys and in the commission condition
of the laboratory studies. While enhance-
ment effects were found only in the
survey, even there they combined with
stronger consistency effects such that
the overall effect of non-verification in a
positive or negative direction in the sur-
vey was a strong U-shaped function show-
ing increased negative emotion when
reflected appraisals were above or below
the identity standard.

The consistency effects but not the
enhancement effects were moderated by
the relevance to the identity of the situa-
tional meanings in the surveys and in
the laboratory studies. In the surveys,
the greater the relevance of meanings in
the situation the greater the consistency
effect, and the more it dominated the
enhancement effect. In the laboratory
studies, we have similar results with the
consistency effect being very pronounced
in the commission condition where we
think more relevant meanings were being
produced by the action of the respondent
compared to the omission condition where
overt action did not take place. Thus,
Hypothesis 2 is generally supported in
the survey and laboratory studies.

Why did the enhancement effect, such
as it was, occur in the surveys but not in
the laboratory studies? We can only spec-
ulate as to what might have occurred.
First, in both the surveys and in the labo-
ratory, individuals had the opportunity to
retrieve from memory their own self-
views in light of their perceptions of
others' views. However, the retrieval of
these self-other views emerged using dif-
ferent time frames. In the survey, individ-
uals reflected on the past in terms of how
they thought others saw them given their
past behavior. In the laboratory, partici-
pants did or did not engage in immoral
behavior and then thought about how
others saw them. If individuals perceived

themselves as overrated in the surveys, it may not have any immediate implications for their behavior or self-views because the events had already transpired. Thus, the over-evaluations wouldn't be very troubling (unless they were large) or encourage people to raise their future identity standards, so individuals could reflectively enjoy the relatively positive meanings. This enjoyment may have been more difficult in the laboratory where, upon reflection, an over-evaluation of the current self may have immediately set up new expectations that participants found distressful. Future research may want to measure the time elapsed since the recalled events in the survey as our suggested explanation would imply more enhancement with greater elapsed time.

There are some limitations to this research. Future work may want to move beyond student participants to the adult population to test the robustness of the effects. In sampling adults, it is important to use moral situations that they have experienced as we did with the student population. We also need to examine other identities that individuals claim beyond the moral identity as well as how more salient and more important/central identities may strengthen the consistency/verification dynamic.

Overall, the results are consistent with the expectations of identity theory but are not fully consistent with the meta-analysis of Kwang and Swann (2010) who found evidence of both enhancement and consistency effects of positive feedback on affective responses, though enhancements effects were stronger on average. We suggested several factors that may alter their results if taken into account. First, it is important to use reflected appraisals. As suggested by Thomas and Thomas (1928), a situation that is perceived as real will be real in its consequences. In the symbolic interaction tradition, individuals' perceptions are the important factor to consider. Thus, it is not the feedback that is important, but the perception of that feedback, which in identity theory is taken to be the reflected appraisals or how persons *think* others view them.

By measuring the reflected appraisals, we do two things. First, we get at the important perceptions, but second, by asking people to reflect on what others think of them, we get them to think about themselves in relation to what others think. This is important because discrepancy effects can only take place when people retrieve from memory information about who they are and then compare the perceived feedback to that information. In the studies reviewed by Kwang and Swann (2010), it is not clear that the aforementioned process was captured. A second factor is the importance of having the meanings in the situation (and the feedback) to be relevant for the identity. If the meanings in the situation and the feedback (reflected appraisals) are not relevant to the self, or are not thought to apply to the self, there will be little effect on the self.

In summary, the present study showed that the effect of a discrepancy between reflected appraisals and the identity standard (the nonverification of an identity) was associated with negative emotion whether the reflected appraisals were more positive or less positive than the identity standard. There was some evidence of an enhancement effect in which reflected appraisals that are slightly more positive than the identity standard make one feel good, but they are small and may hold only for the recall of past events. In any case, as the positive evaluations increase much beyond one's identity standard and persist, they produce negative emotions whether as past or current situations. The self plays a critical role in responding to feedback from

others. In predicting how people will respond, we must take their self-views into consideration. The meaning of the situation and the perceived responses of

others to the individual can only be understood relative to the individual's identity.

APPENDIX

Situations	Experience (percentage)	Moral Relevance
You are asked to donate to a charity. You did not provide a donation or you provided a donation.	92	2.43
You are drunk and need to drive home. You did not drive myself home or drove myself home.	44	3.67
You have the opportunity to copy another student's answers during an exam. You did not copy the student's answers or copied the student's answers.	73	2.59
You find an item that does not belong to you (for example, a cell phone, a wallet, a book and so forth). You did not return the lost item or returned the lost item.	85	3.21
You have the opportunity to take an item (for example, merchandise, money etc.) that doesn't belong to you. You did not take the item or took the item.	68	3.13
A friend of yours is drunk and wants to drive home. You did not let your friend drive home or let your friend drive home.	68	3.87
You have the opportunity to allow another student to copy your answers during an exam. You did not let the student copy your answers or let the student copy your answers.	72	2.12
A cashier returns more money to you than what is owed. You did not return the money or returned the money.	77	2.83

REFERENCES

Allison, Scott T. and David M. Messick. 1988. "The Feature-Positive Effect, Attitude Strength, and Degree of Perceived Consensus." *Personality and Social Psychology Bulletin* 14:231–41.

Anderson, Christopher J. 2003. "The Psychology of Doing Nothing: Forms of Decision Avoidance Result from Reason and Emotion." *Psychological Bulletin* 129:139–67.

Aquino, Karl and Americus II Reed. 2002. "The Self-Importance of Moral Identity." *Journal of Personality and Social Psychology* 83:1423–40.

Burke, Peter J. and Jan E. Stets. 2009. *Identity Theory*. New York: Oxford University Press.

Cooley, Charles H. [1902] 1964. *Human Nature and the Social Order*. New York: Schocken Books.

Fazio, Russell H., Steven J. Sherman, and Paul M. Herr. 1982. "The Feature-Positive Effect in the Self-Perception Process: Does Not Doing Matter as Much as Doing?" *Journal of Personality and Social Psychology* 42:404–11.

Greene, Joshua D., R. Brian Sommerville, Leigh E. Nystrom, John M. Darley, and Jonathan D. Cohen. 2001. "An Fmri Investigation of Emotional Engagement in Moral Judgment." *Science* 293:2105–8.

Hauser, Marc. 2006. *Moral Minds: How Nature Designed Our University Sense of Right and Wrong*. New York: Harper Collins.

Heise, David R. and Cassandra Calhan. 1995. "Emotion Norms in Interpersonal Events." *Social Psychology Quarterly* 58:223–40.

Kalkhoff, Will and Robb Willer. 2008. "Deterring Deviance: Rationality and Self-Control." Pp. 39–61 in *Experiments in Criminology and Law: A Research Revolution*, edited by C. Horne and M. J. Lovaglia. Lanham, MD: Rowman and Littlefield.

Kwang, Tracy and William B. Swann. 2010. "Do People Embrace Praise Even When They Feel Unworthy? A Review of Critical Tests of Self-Enhancement Versus Self-Verification." *Personality and Social Psychology Review* 14:263–80.

Robinson, Dawn T., Lynn Smith-Lovin, and Allison K. Wisecup. 2006. "Affect Control Theory." Pp. 179–202 in *Handbook of the Sociology of Emotions*, edited by J. E. Stets and J. H. Turner. New York: Springer.

Ross, Lee D. 1977. "The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process." Pp. 173–220 in *Advances in Experimental Social Psychology*, edited by L. Berkowitz. New York: Academic Press.

Sedikides, Constantine and Aiden P. Gregg. 2008. "Self-Enhancement: Food for Thought." *Perspectives in Psychological Science* 3:112–16.

Shrauger, J. S. 1975. "Responses to Evaluation as a Function of Initial Self-Perceptions." *Psychological Bulletin* 82:581–96.

Spranca, Mark, Elisa Minsk, and Jonathan Baron. 1991. "Omission and Commission in Judgment and Choice." *Journal of Experimental Social Psychology* 27:76–105.

Stets, Jan E. 2006. "Identity Theory and Emotions." Pp. 203–23 in *Handbook of the Sociology of Emotions*, edited by J. E. Stets and J. H. Turner. New York: Springer.

Stets, Jan E. 2011. "Applying Identity Theory to Moral Acts of Commission and Omission." *Advances in Group Processes* 28:97–124.

Stets, Jan E. and Emily K. Asencio. 2008. "Consistency and Enhancement Processes in Understanding Emotions." *Social Forces* 86:1055–58.

Stets, Jan E. and Michael J. Carter. 2011. "The Moral Self: Applying Identity Theory." *Social Psychology Quarterly* 74:192–215.

Stets, Jan E. and Michael J. Carter. 2012. "A Theory of the Self for the Sociology of Morality." *American Sociological Review* 77:120–40.

Stets, Jan E. and Shelley N. Osborn. 2008. "Injustice and Emotions Using Identity Theory." *Advances in Group Processes* 25:151–79.

Swann, William B. Jr. and Michael D. Buhrmester. 2011. "Self-Verification: The Search for Coherence." Pp. 405–24 in *Handbook of Self and Identity*, edited by M. R. Leary and J. P. Tangney. New York: The Guilford Press.

Thomas, William Isaac and Dorothy Swain Thomas. 1928. *The Child in America*. New York: Knopf.

Turner, Jonathan H. and Jan E. Stets. 2005. *The Sociology of Emotions*. New York: Cambridge University Press.

Walker, Lawrence J. and Karl H. Hennig. 2004. "Differing Conceptions of Moral Exemplarity: Just, Brave, and Caring." *Journal of Personality and Social Psychology* 86:629–47.

Walker, Lawrence J., Russell C. Pitts, Karl H. Hennig, and M. Kyle Matsuba. 1995. "Reasoning about Morality and Real-Life Moral Problems." Pp. 37–407 in *Morality in Everyday Life: Developmental Perspectives*, edited by M. Killen and D. Hart. New York: Cambridge University Press.

BIOS

Jan E. Stets is a professor of sociology and codirector of the Social Psychology Research Laboratory at the University of California, Riverside. She is an AAAS Fellow and the 2010 recipient of the ASA Section on Emotion's Lifetime Achievement Award. Recent publications include *Handbook of the Sociology of Emotions, Volume II* (with J. H. Turner), 2014; "Identity Theory" (with R. T. Serpe) in *Handbook of Social Psychology*, 2013; and "A Theory of the Self for the Sociology of Morality" (with M. J. Carter) in *American Sociological Review*, 2012. She is incoming editor (with R. T. Serpe) of *Social Psychology Quarterly*.

Peter J. Burke is distinguished professor of sociology and codirector of the Social Psychology Research Laboratory

at the University of California, Riverside. He is an AAAS Fellow and the 2003 recipient of the ASA Section on Social Psychology's Cooley-Mead Award. Recent publications include "The Past, Present, and Future of Identity Theory" (with J. E. Stets) in *Advances in Group Processes*, 2014; "Social Comparison in Identity Theory" (with J. E. Stets) in *Communal Functions of Social Comparison*, 2014; and "Does Incarceration Change the Criminal Identity? A Synthesis of the Labeling and Identity Theory Perspectives on Identity Change" (with E. Asencio) in *Sociological Perspectives*, 2011.