2012

Dropping the Metaphorical Rod of Blame to Save Humanity: Testing Multiple Means of Breaking the Link between External Explanations for Transgressions and Increased Cynicism Toward Humanity

Phillip D. Getty
Lehigh University

Dropping the Metaphorical Rod of Blame to *Save* Humanity: Testing Multiple Means of Breaking the Link between External Explanations for Transgressions and Increased Cynicism Toward Humanity

by

Phillip D. Getty

A Thesis

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Master of Science

In

Psychology

Lehigh University

09/01/2011

Thesis is accepted and approved in partial fulfillment of the requirements for the Master of Science in Psychology

Dropping the Metaphorical Rod of Blame to Save Humanity: Testing Multiple Means of Breaking the Link between External Explanations for Transgressions and Increased Cynicism Toward Humanity

Phillip D. Getty

_____

Approval Date

_____
Mike Gill, Ph.D.
Advisor and Committee Chair

_____
Gordon Moskowitz, Ph.D.
Committee Member

_____
Dominic Packer, Ph.D.
Committee Member

_____
Ageliki Nicolopoulou, Ph.D.
Department Chair

Table of Content

List of Figures

Abstract

We are interested in the psychology of social explanations and the ironic possibility that external explanations might foster cynicism toward humanity even as they increase compassion toward specific transgressors. We proposed several studies testing means of breaking this external explanations/cynicism link by disrupting the biased retrieval process with subtle reminders of positive instances of humanity (e.g., Gandhi, the self) (Study 1), varying the presence of exculpatory information about the bad actors composing external explanations (Study 2), and inducing lower-level construal (Study 3). While we did not find evidence for all of the mechanisms we examined, we did find that a subtle reminder of one's own humanity could break the link between external explanations and cynicism. We provide additional information suggesting that the link between social explanation and cynicism is more complex than previously thought along with ideas for future research aimed at better understanding the relationship between cynicism and social explanation.

Dropping the Metaphorical Rod of Blame to _Save_ Humanity: Testing Multiple Means of Breaking the Link between External Explanations for Transgressions and Increased Cynicism Toward Humanity

Social explanations are answers to the question of why an individual or group acts a particular way or experiences a given outcome (Andreychik & Gill, 2009; Gilbert, 1998; Gill & Andreychik, 2007, 2009, 2011; Heider, 1958; Jones & Davis 1965; Kelley, 1973; Wiener, 1985). Such explanations often point to either *internal causes* (something inherent to the target caused the behavior or outcome; "*Tommy punched Sam because he's a violent and depraved little boy"*) or *external causes* (something outside of the target caused them to behave or experience an outcome: "*Tommy punched Sam because he comes from a violent home where he has learned an inappropriate behavior"*). Across several studies now, researchers have found that prosocial responding to both victim and villain vary as a function of these causal dimensions. In particular, when people infer that external forces have caused the plight or the misguided behavior of an actor, they respond more compassionately and less punitively toward the actor (Gill & Andreychik, 2009, 2011; Zucker & Weiner, 1993).

Gill and Getty (2010), however, have presented evidence that even as external explanations foster prosocial responses to the actor who is being explained, they can have a not-so-prosocial side effect occurring simultaneously. Specifically, they suggested that frequently relying on external explanations might increase one's cynical feelings towards humanity because external explanations typically point to how the target individual was *harmed or hindered in some way by other people*. External explanations may lead one to spare transgressors the rod, but to damn humanity in the process. The present proposal

will describe studies that investigate mechanisms for shutting down this tendency for external explanations to foster negative feelings about humanity. First, however, we will review some of the pertinent literature to position our work within the broader person-perception literature.

The Commencement: Attribution Theory and Social Explanations

Person perception is the study of how people reason about and make sense of other people's behavior (see Gilbert, 1998; Moskowitz & Gill, *in press*; for reviews). Although work on person perception has taken many forms in the last 50 years, one dominant stream of activity has centered on Attribution Theory, which was spawned from Fritz Heider's (1958) influential work. He coined the metaphor of the "naïve scientist" to describe how lay people, like scientists, are on a quest for accurate understanding.

Heider suggested that the naïve scientist scrutinizes behavior with a form of causal analysis. Heider proposed that lay interpretation of behavior reflected the common-sense understanding that others are *affected by* and also *cause changes in* their environment. This common sense understanding is evident in the types of explanations people use to describe the causes of others' behavior. Thus, some explanations suggest that the behavior was caused by a "force" that resides within the actor (i.e., *internal* or *dispositional* cause; "*Tommy punched Susie because he's a violent and depraved little boy"*) or by a "force" that impinged upon the actor from "outside," and which influenced or pushed the actor to act (i.e., *external* or *situational* cause; "*Tommy punched Susie because he was forced to defend himself against Susie's menacing advance"*). In

Heider's view, observers interpret the meaning of a single act by weighing the relative influence of these two causal factors.

The goal of this causal analysis is to infer, logically, whether a given behavior is likely to reflect a general disposition of the actor. Heider asserted that inferring dispositions is crucial to the effective social functioning of an individual, because we need to predict others' behaviors in order to optimize future interactions. For example, if I see Tommy in a fit of rage, yelling and shoving strangers after eating ice cream, I may immediately infer that (a) this guy is a jerk, (b) he has a rare disorder that causes aggression after ingesting cold dairy products (both internal causes), or (c) someone slipped a jerk pill instead of M&Ms in his hard-serve (external cause). Since the instances of jerk pills being slipped are rare at best and because ice cream rarely makes people angry, the most likely option is to infer that Tommy is a jerk (dispositional inference) and to anticipate jerk-like behavior in future interactions. Therefore, if I see Tommy in line at the local Cold Stone Creamery, I know that I should seek services elsewhere, or be prepared to defend myself.

Heider set the stage for other researchers to incorporate his ideas into a more complete understanding of the "logical rules" people *should* use when explaining behavior, assuming that their goal is accurate understanding. Heider may have begun the study of lay behavioral explanations, but it wasn't until later that his ideas were tested and expanded to describe a more complete rule system used by observers as they engage in causal analysis of others' actions.

The Logical Rules of Attribution

Jones and Davis (1965) sought to describe the inferential rules that people use to come to a judgment about an actor. They suggested that lay people first infer an actor's intentions, or what it was that the actor was trying to accomplish. Then, under certain informational circumstances, they may go further to make a *correspondent inference* about the actor's underlying disposition, or distinguishing character traits. In short, Jones and Davis addressed two questions: (1) how do people infer an actor's intentions? (2) Under what conditions does the observer assume that the inferred intention is informative about the actor's distinctive, stable character traits? To answer the first question, Jones and Davis suggested that people infer intentions through an analysis of the *non-common effects* of an act. That is, they consider what effects were produced by an act that would not have been produced by alternative acts.

For example, Tommy is seen tipping Jane, the attractive employee at the cash register, prior to his aforementioned fit of rage. To figure out Tommy's intention with this act, we would have to determine the non-common effects of tipping (i.e., effects which are absent with non-tipping). At Cold Stone, this could be a difficult task—tipping sets off a chain reaction of effects. If Tommy throws fifty cents in the tip jar, Jane and the other employees are prompted to break out in song. An observer would probably have to ask herself if Tommy's desired effect was sticking Jane and her fellow employees with the humiliation of singing in public for a pittance. If so, she would probably make a *correspondent inference* that Tommy had a sick, twisted and depraved disposition, and that she should avoid eye contact. Similarly, she might infer that Tommy genuinely wanted to express appreciation for speedy service, which might spark the *correspondent inference* that Tommy is a "swell guy." On the other hand, it could be that Tommy was

simply trying to display customary behavior, or the expected social norms, which closely resembles a *situation* or *external cause*. If the observer infers this intention, she may not take the time to make a *correspondent inference*, because the inferred intention does not distinguish Tommy from any other customer (he is just an "ordinary guy"). These and other effects could be likely, but only if Tommy tips. Nevertheless, this example illustrates a problem: a single behavior could have numerous non-common effects.

How does an observer decide which is the intended non-common effect? Jones and Davis suggest that it is the *social desirability* of a given effect, or the degree to which other people would seek the same effect, that guides the observer's decision: Actors are assumed to be aiming for the effect that most people would aim for. Because most people desire to avoid public ridicule, the most logical answer might be that people tip to avoid public scrutiny or embarrassment. However, this inference does little to inform us about any distinguishing characteristics of Tommy, and certainly does nothing to clue us in to his violent post-ice-cream-eating behavior.

Using Jones and Davis' method, the observer could only make a meaningful correspondent inference about Tommy's distinguishing disposition if she infers or observes that the ostensible intended effects of his tipping behavior are out of the ordinary. In other words, it would be difficult or impossible for the observer to make a correspondent inference that might predict Tommy's violent outburst unless the observed effect is abnormal or undesired. In this case, the observer needs more data.

If the ultimate goal of person perception is predicting future behavior, Kelley (1973) noted that this goal is unlikely to be achievable via information extrapolated from a single act. In light of this observation, Kelley proposed the *covariation model*, which

emphasized the need to examine the distinctiveness, consistency and consensus of an act across multiple observations to make a reasonable causal inference about it. In our example of Tommy's fit, the observer would have to determine whether this was his first violent outburst after eating ice cream (*consistency*). She would next have to consider whether Tommy acts violently across different contexts (*distinctiveness*). Finally, she would have to consider and compare other customers' post-ice-cream-eating habits with Tommy's to determine if "whooping some ass" after finishing an ice cream cone is what people typically do (*consensus*).

For example, if in the past, the observer had never witnessed Tommy act violently in any setting (e.g., his behavior is *distinct*), the other customers are calm and benevolent (e.g., there is no *consensus* for violence), and this is the first time Tommy has been observed raging after eating ice cream (e.g., his behavior is *inconsistent*), then the observer might conclude that some *temporary* factor influenced Tommy's rage. Similarly, if Tommy's violence is *distinct*, *inconsistent*, and other people are acting similarly, the observer can be quite confident that the ice cream (external cause) is causing people to rage. On the other hand, if Tommy is known for having fits of rage in public (*non-distinct*) and he is the only one doing it (*no consensus*), and Tommy has been known to be violent after eating ice cream, then she will infer that Tommy has an inherent violent disposition. Even if Tommy is known for having fits of rage only after eating ice cream, she may still infer a violent disposition. However, if his behavior is *non-distinct*, *inconsistent*, and others are acting similarly, then the observer might be quite confused about the true cause, demonstrating a phenomenon Kelley (1972) coin as the *discounting principle* (discussed below): The true cause remains ambiguous.

As illustrated above, *the covariation model* requires multiple observations to come to a logical conclusion about an actor's internal disposition. Be that as it may, Kelley (1973) did not abandon the notion that observers sometimes make attributions from a single act. He observed that people possess preexisting knowledge about how both dispositions and situations have the power to bring about effects in the environment. Therefore, if an observer had never seen Tommy before and witnessed his violent outbursts, she could easily conjure to mind the notion that "only hooligans act out in public like this." On the other hand, if she has prior knowledge of a crime spree involving the random slipping of jerk-pills, she may conclude that some external force caused Tommy to act up. However, the observer may consider each cause as sufficient to produce a violent act (i.e., there are *multiple sufficient causes)*. If so, she will again employ the *discounting principle*: she cannot be confident in the attribution of causality to either potential cause. Although, if she believes that Tommy's display required *multiple necessary causes* (i.e., the effects of jerk-pills are most pronounced in people who are already prone to hooliganism) she may become more confident in her assessment of Tommy's character (Kelley, 1972). Moreover, if there were other factors that normally prevent public violence (two police officers standing at the back of the line), she would certainly gain even more confidence in her assessment of Tommy's disposition (i.e., the *augmentation principle*; Kelley, 1973).

As just illustrated with these classic attribution models, the observer is depicted as a logical thinker and rule follower with the time and the will to deliberate and scrutinize the potential causes of Tommy's behavior—just as a scientist would do—to come to an accurate conclusion about his disposition. However, the social cognition approach moved

beyond this notion of lay people acting like scientists and acknowledged people's tendency to act like quick-and-dirty information-processing units, looking for fast answers that work (i.e., the "cognitive miser" view of the social perceiver).

For example, Quattrone (1982) argued that people rarely come to an accurate dispositional inference about an actor's intentions by hashing out the evidence. Instead, people spontaneously infer personality traits ("Jerk:" internal cause). Only when motivated to do so will observers sacrifice cognitive resources to adjust the initial inference by considering other causal factors, such as the actor's situation (external cause). Nevertheless, any adjustment remains biased by observers' spontaneous dispositional attributions (*"Well, he may have been slipped jerk-pills, but he must have already been a jerk to react so strongly to them,").* Quattrone (1982) suggested that this processes was evidence of an anchoring-adjustment mechanism of person perception.

As illustrated by the anchoring adjustment mechanism, Quattrone did not abandon the notion of conscious deliberation in his model; it just took a back seat. Similarly, Trope (1986) and Gilbert and his colleagues (see Gilbert, 1998, for a complete review) presented models incorporating aspects of automaticity in the process of inferring dispositions that also allowed for deliberate adjustment. Bringing this all back to the example of Tommy, these researchers would predict that the observer would automatically infer Tommy's disposition as the cause of the violent outburst and then, if so inclined, follow-up the assessment by considering possible environmental causes (infusion of jerk pills).

The Socio-Emotional Consequences of Causal Inferences

The study of person perception developed further when, rather than examining how perceivers select an internal or external cause to explain an act, researchers began to examine how explanations shape socio-emotional responses—such as compassion versus anger—to the actor. Weiner (1985; also see Weiner, 2006 for a complete review) has been a major proponent of and contributor to this approach. Wiener proposed a theory acknowledging a number of mediating determinates between an eliciting stimulus (the behavior of an actor or a group) and the response it elicits in the observer; and he emphasized the proximal role of emotions in this process. Weiner proposed that observing a behavior or outcome leads an observer to make a causal inference (this stage of *explanation selection* is what was emphasized in the theories described above). Once that inference is made, the observer scrutinizes the properties of the chosen cause—especially responsibility and controllability implications—which brings about an affective response (i.e., sympathy, anger, indifference). Finally, after an emotion is elicited, the observer responds behaviorally (i.e., approach, avoid, etc.). He generalized this sequence as one of thinking → feeling → action. Lots of evidence has been taken as support for this model.

For example, Wiener and Kukla (1970) presented early evidence supporting Weiner's model. Participants were placed in the role of an educator and asked to react to the performance of a fictitious student who failed to achieve academically. The participants were informed that the student either (a) had low intelligence or (b) did not exert effort. The authors found that participants experienced more negative affect and responded more punitively in the low effort condition than in the low intelligence condition. The authors concluded that the key difference between the two conditions was

that the student-target in the low effort condition had more *control* over his academic outcome than the student-target with low intelligence. Indeed, additional studies by a number of researchers have concluded that in response to a negative outcome, the degree to which an actor *could have done otherwise* is directly related to the extent to which an observer will respond with anger and punishment (see also Meyer & Mulherin, 1980; Reisenzein, 1986; Weiner, Perry, & Magnusson, 1988). One notable feature of Weiner's work is that he focuses on the distinction between *internal*-controllable (lazy) and *internal*-uncontrollable (unintelligent), rather than on the internal vs. external distinction that pervades work on explanation selection (reviewed in the previous section).

Perceived controllability has been a guiding topic of much research and is often framed as the most crucial mediator of people's emotional reactions to the acts and outcomes of others (see Weiner, 2006). Recent work, however, is beginning to question whether perceived controllability deserves such a dominant role. For example, Mullen and Skitka (2009) showed that perceived controllability is weighed less by members of collectivist societies—who pay more attention to past social contributions of the actor—than by members of individualist societies. Furthermore, Gill and Andreychik (2011) suggest that perceived controllability is not the sole mediator of the effect of explanations on emotional responses. They propose that perceived suffering also acts as a mediating mechanism by which external explanations evoke compassion and foster positive responses to both victims and villains. Indeed, Gill and Andreychik demonstrate that even when a target's negative acts are perceived as highly controllable, perceived suffering mitigates the anger and blame elicited by inferred controllability. We will elaborate on their model below.

The Psychology of External Explanations

The specific phenomenon of interest to us here focuses on external explanations (which are emphasized in the literature on explanation selection; reviewed by Gilbert, 1998) and socio-emotional responses (which are emphasized in the literature on explanations and socio-emotional responses; reviewed by Weiner, 2006). A handful of studies document a link between external explanations and prosocial emotions, although they sometimes differ in their views of the mechanisms responsible for this link. As discussed above, one camp emphasizes the role of perceived controllability and suggests that external explanations imply that a target lacks control over a given behavior or outcome, and that reductions in perceived controllability that lead observers to feel compassion for the target (Alicke, 2000; Zucker & Weiner, 1993). For example, Zucker and Weiner (1993) conducted path analyses supporting the notion that when people believe poverty is caused by external forces (e.g., limited availability of jobs and education) they tend to respond more prosocially to the poor, and this link is mediated by decreased perceptions of how much control the poor have over their economic outcomes.

Additional work further supports a link between external explanations and compassion, but argues that such effects are mediated by perceived suffering in addition to and independently of controllability perceptions (Gill & Andreychik, 2011). Gill and Andreychik (2011) sought to dissociate a mechanism involving perceived suffering from the mechanism involving perceived controllability. They presented three studies in which African Americans (i.e., victims of social injustice) were the target group, and two studies in which Chechen militants (i.e., villains) were the target group.

The results of Study 1 suggested that the link between external explanations and compassion for African Americans was only partially mediated by controllability perceptions, which left room for additional mediators to carry the effects of external explanations to compassion. Therefore, in Study 2, Gill and Andreychik tested and found that perceived suffering was an additional mediating mechanism. Finally, in Study 3 they included both perceived controllability and perceived suffering in a model predicting compassion and found that, indeed, together they carried the entire effect of external explanations to compassion. In Studies 4 and 5, using an experimental design, the authors asked participants to respond to questions about a Chechen militant group who committed grotesque and entirely controllable terrorist acts at a Russian primary school. All participants learned general information about long-standing conflict between Chechens and Russians. In the external explanations condition, however, additional information was given regarding the viciousness and brutality of the Russian military and the suffering it caused to Chechens; participants in the control condition were given no additional information. The authors found that the external-explanations group was less likely to support vengeful retaliation against the Chechens by the Russians. Meditational analysis suggested that the reason was that perceptions of Chechen suffering at the hands of the Russians lead participants to feel compassion for them despite their appalling acts. In sum, Gill and Andreychik provide strong evidence for their proposed model in which external explanations increase perceptions of suffering, which elicit compassion independent of perceived controllability. In fact, as the authors note, these findings provide strong evidence for the independence of the external explanations/suffering mechanism from perceptions of uncontrollability.

In addition to perceived suffering, Gill and Andreychik (2007) linked external explanations to prosocial behavior via prejudice-related compunction, a negative self-directed emotion that motivates people to become less prejudiced. In Study 1, the authors measured internal and external explanations for the socio-economic situation of Blacks, and feelings of compunction within Devine et al.'s (1991) "*should/would*" paradigm. Specifically, White participants indicated how they *should* and *would* react when encountering African Americans. They found that participants experienced more compunction when they imagined that they *would* show biases that they felt they *should not* possess. Interestingly, endorsement of external explanations moderated this relationship, such that participants who strongly endorsed external explanations felt greater compunction in response to should-would discrepancies than participants who weakly endorsed external explanations. This relationship was independent of the effects of internal motivation to be less prejudiced (Plant & Devine, 1998), and was mediated by the fact that those who endorsed external explanations perceived biases as being especially unjust and undeserved. In Study 2, the authors sought to replicate their findings in a novel paradigm. The new task allowed the authors to assess the extent to which external explanations contributed to the experience of compunction upon one's admitting that in some circumstances he or she would discriminate against young African American males (i.e., by avoiding them in a dangerous neighborhood). Again, those that strongly endorsed external explanations felt more compunction in response to such an admission than did those who weakly endorsed external explanations.

Additional work from a different literature—the literature on perspective taking—provides more evidence for prosocial effects of external explanations. Vescio, Sechrist,

and Paolucci (2003) examined whether external explanations mediate the effects of perspective taking—i.e., mental simulation of what a target must have been feeling in a given situation—on improved attitudes toward a target. To test this relationship, the authors had participants listen to a fictitious broadcast of an African-American student discussing his personal experiences with discrimination. Participants were given instructions to either take the student's perspective or remain objective. Results revealed that perspective taking improved attitudes and this effect was mediated by increased acceptance of external explanations regarding any suffering of the target.

Finally, because of differences in culture, family experience, education, and other individual experiences (Morris & Peng, 1994), it has been proposed that people develop *social explanatory styles*: Personal theories about why people generally act as they do or experience particular social outcomes (Andreychik, 2006, 2009; Faulkner, 2007; Gill & Andreychik, 2010). These explanatory styles may either lead one to chronically attribute causality to dispositional causes (i.e., *dispositionism*) or environmental/"interconnected" causes (i.e., *externality*). These styles are not fixed, of course, or mutually exclusive. Rather, both dispositionism and externality exist and interact within each of us, although people may tend to rely more heavily on one theory than another. Crucially for present purposes, evidence suggests that the socio-emotional effects of chronically employing a particular explanatory theory may be similar to the effect of "one shot" explanations (e.g., external leads to more compassion). Indeed, external explanatory style is positively associated with dispositional compassion (see Gill & Andreychik, 2009 for a review).

Do External Explanations Always Have Prosocial Consequences?

Although the studies cited above provide compelling evidence for prosocial consequences of external explanations, other work suggests that there is no *necessary* prosocial consequence. Gill and his colleagues have shown that explanations take on meaning in the context of particular social mindsets, with mindsets modulating—sometimes completely inverting—the meaning of the explanation (see Gill & Andreychik, 2009).

For example, Andreychik and Gill, citing Social Identity Theory (SIT; Tajfel & Turner, 1986), proposed that if an external explanation threatens a perceiver's group identity, the explanation would evoke angry defensiveness rather than prosocial feelings. They examined this relation in a series of correlational and experimental studies. Across four studies, they compared high- and low-identifying White participants' responses to explanations about the low status of Blacks. They assumed that low-identifiers would be relatively likely to view external explanations through a lens of justice and care motives, which would result in a prosocial response to such explanations. In contrast, they argued that high-identifiers would be relatively likely to view external explanations as a threat to positive White identity, which would result in angry and defensive responses to such explanations. Consistent with this analysis, they found that high identifying Whites responded to external explanations with an absence of sympathy (Study 1a), increased negative stereotyping of Blacks (Study 1b), an absence of a sense of injustice, (Study 2), and increased implicit prejudice (Study 3). This pattern of results, the authors suggest, is indicative of a pattern of defensive responding on the part of high-identifiers: a meager attempt to save face and preserve their ingroup's superior status. Low-identifying Whites, on the other hand, responded to external explanations with greater sympathy (Study 1a),

rejection of derogatory stereotypes of Blacks (Study 1b), increased perception of injustice (Study 2), and reduced implicit prejudice (Study 3). These findings clearly suggest that social explanations have no fixed or singular meaning, and that external explanations can have negative socio-emotional consequences from within certain mindsets.

External Explanations and the Impugning of Humanity

Now we address a different sense in which external explanations might fail to have prosocial effects. In particular, we are interested in the implications that frequent employment of external explanations to explain individual transgressions might have for one's general feelings about humanity. Why might there be a connection? The reason is that each time an individual transgression (Bob killed a pedestrian while driving intoxicated with a hooker in his lap) is explained via an external explanation (Bob spent his childhood being ignored within a family full of drunkards and pimps), one is bringing to mind the malevolence, cruelty, thoughtlessness, or unfairness of other human beings that presumably led the transgressor to his misdeed. Although bringing to mind such unsavory and despicable actors might facilitate the experience of compassion for the transgressor, it might simultaneously feed an image of humanity as being full of despicable actors. In contrast, if an explanation implicates the individual transgressor as solely responsible (e.g., an internal explanation), then one will be more likely to view the transgressor as "just an isolated bad apple" and thus not impugn humanity as a whole.

Is there any existing evidence for the proposed relation between external explanations and cynicism regarding humanity? Yes. Early work examining the benefits of intra-group support among police officers coping with stress is relevant here (Chandler & Jones, 1979). Chandler and Jones asserted that police officers are prone to develop less

then amiable impressions of humanity because of the nature of their work. Specifically, police officers are regularly exposed to human suffering at the hands of other humans (externally caused suffering). These conditions lead officers to develop an overall cynical view of humanity. Whereas this work did not examine whether police officers' cynicism differed as a function of their explanations for criminal behavior, nor was there a direct comparison of levels of cynicism between police and non-police, it still suggests that widespread exposure to the negative instances of human action (the types of action often inherent in external explanations) does negatively affect how people evaluate humanity as a whole.

Additionally, correlational data from our lab are consistent with the proposed relation. Our concept of social explanatory style was mentioned above. We recently created a Social Explanatory Style Questionnaire (SESQ) that focused exclusively on explanatory style for transgressions. Respondents are presented with six short vignettes describing, for example, deceitfulness, hypocrisy, authoritarian bullying, and criminality. For each vignette, respondents indicate the extent to which they believe the transgressor acts as he or she does because of life experiences and circumstances (i.e., external causal forces) that would cause another person to act similarly. In one study, participants completed this Transgression SESQ along with the Humanity Esteem Scale. In that study, external explanations did not have prosocial effects at all: In fact, participants had more cynical views of humanity—as tapped by the Humanity Esteem Scale—to the extent that they characteristically attributed transgressions to external causal forces. Also, consistent with our "just a few bad apples" reasoning described above, a tendency to

characteristically attribute transgressions to forces internal to the transgressor did not decrease humanity esteem.

Finally, while developing the Humanity Esteem Scale (based on the Rosenberg Self-Esteem Scale), Luke and Maio (2009) found that humanity esteem increased when participants were exposed to media images that promoted social values ("Families are the building blocks of a strong nation"). They also found that humanity esteem decreased by exposing participants to media images that threaten societal values (Terrorist activities). Thus, to the extent that exposure to multiple external explanations for transgressions involves increased exposure to those who threaten societal values, such exposure should lower positive feelings towards humanity.

All in all, this body of research suggests that the relationship between reasoning about individuals and responses towards the larger group may be nothing short of a complex blame game in which refraining from blaming the individual fosters blaming of humanity in general. Nonetheless, these findings do not definitively tell us whether multiple exposures to external explanations *cause* people to express negative feelings towards humanity because none of the work above directly manipulated such exposure experimentally. Accordingly, Gill and Getty (2010), investigated this relationship using an experimental design.

Gill and Getty (2010) had participants read six vignettes describing transgressions (e.g., "Megan is an accountant for a large company. Over a period of years, she 'faked the books' and shifted company money into her personal bank account."). Some participants received no explanations for the transgressions. Other participants received internal explanations for each transgression (e.g., "Megan wants more excitement and

more material goods in her life. This motivates her to divert company money into her personal account."). And, finally, the remaining participants received external explanations for each transgression (e.g., "Megan's boss is always criticizing her and taunting her about being 'stuck' at her job due to constraints on her ability to move her family. Her boss gives her terrible raises because of her being 'stuck.' This motivates her to divert company money into her personal account.") The authors predicted that humanity esteem would be lower in the external explanations condition than in the other two conditions. They reasoned that this would happen in spite of the fact that the individual transgressors in that condition will be responded to more compassionately and less angrily than in the other conditions. This would indicate that while external explanations directly activate compassionate feelings towards individual transgressors, they do so at the expense of the larger group.

Study 1 revealed that, in fact, transgressors in the external explanation condition were responded to more compassionately and less angrily than transgressors in the internal explanations or control conditions, although no decrements were found in feelings toward humanity. Study 2 corrected for some methodological shortcomings of Study 1 (e.g., a poor measure of humanity esteem; the act of explicitly reporting compassion toward the transgressor primed positive feelings and prevented reductions in humanity esteem). In Study 2, as in Study 1, the primary manipulation was the type of explanation given for the transgressions. However, rather than having participants respond directly about the individual transgressors, they were lead to believe that they would be participating in a "social memory task." Specifically, they learned that they would read about six individuals (the same six as in Study 1), and after some intervening

"distracter tasks" they would try to remember as much about those six individuals as they could. In reality, one of the ostensible "distracter tasks" was a measure of cynicism toward humanity created by Gill and Getty (2010). Study 2 confirmed predictions: Those in the external explanations condition reported more cynical attitudes toward humanity than did those in either the internal explanations or control conditions.

Gill and Getty (2010), then, found evidence of a complex blame game. The result of this game at times seems to have damning consequences for individuals (when explanations focus on inner motives and traits of the transgressors) and, at other times, humanity seems to be the ultimate target of blame (when explanations focus on the external force of other people's malevolence, unkindness, and mindlessness as giving rise to the individual actor's transgression). The intended contribution of the studies proposed below is to examine factors that will allow the prosocial effects of external explanations (i.e., compassion for the individual transgressor) to persist while eradicating the undesirable tendency for external explanations to promote the impugning of humanity. Can we save humanity?

The Current Studies

The logic underlying the studies below was to identify several cognitive-processing features that potentially enable external explanations to increase cynicism toward humanity, and then to disrupt those features to test whether such disruption prevents the external explanations/cynicism link. Accordingly, we went about testing three cognitive-processing features that we believed were likely mechanisms enabling the external explanations/cynicism link: (a) *Biased retrieval*. When reporting their attitude toward humanity, the "bad actors" composing external explanations may form a

disproportionate share of what is brought to mind (i.e., an availability heuristic; Tversky & Kahneman, 1973); (b) *Failure to think "beyond the information given."* Participants are given no exculpatory information about the "bad actors" composing the external explanations, such as what their "meta-desires" might be with respect to their bad acts (e.g., have they come to regret their bad acts?; see Pizarro, Uhlmann & Salovey, 2003). Because of this, we suspect that participants likely blame and become angry toward those bad actors, which could contribute to a dim view of humanity; and (c) *Abstract level of construal*. When processing the external explanations, participants might be construing the information in abstract, high-level terms (see Trope, in press). At an abstract level of construal, the "bad actors" in the external explanations are likely construed as "representatives of humanity" behaving badly. At a lower, more concrete level of construal the "bad actors" might become just a limited set of bad actors. The three studies presented below were an attempt to identity three potential mechanism enabling the external explanations/cynicism link by disrupting and/or overriding the biased retrieval process (Study 1), varying the presence of exculpatory information about the bad actors composing external explanations (Study 2), and inducing lower-level construal (Study 3).

Study 1: Disrupting Biased Retrieval Processes

("Wait a minute…I'm Human, Too…and so is Gandhi!")

When observing objects, events or situations, the information that comes to mind quickly and easily is given priority when making a judgment about what was observed. It seems that we are wired to believe that easily retrievable, available information that comes to mind quickly is the most valid information and should be given the most weight

when making our judgments (Moskowitz, 2005). When these judgments occur, it is said that people are demonstrating the *availability heuristic* (Tversky & Kahneman, 1973).

Tversky and Kahneman (1973) were the first to experimentally demonstrate the availability heuristic. They asked participants to read and memorize lists of male and female names, some of which were better known and famous than others. The authors varied the presentation of famous names, such that in one condition, there were several well-known male names and in another there were several well-known female names. After reading the lists, participants were either asked to judge the number of gender-specific names in each list or to recall as many names as possible. They found that participants demonstrated a significant advantage for memorizing names that were of the same gender as the famous names. Likewise, while there were an equal number of female and male names, participants reported believing that a higher frequency of males or females existed depending on which gender included the famous names. In sum, because the well-known names were salient and easily available in mind, participants seemed to recall more of the gender-congruent names and thus overestimated their frequency, thus demonstrating how available information can bias our judgment of related events or objects.

Specifically in line with the notion of the availability heuristic (Tversky & Kahneman, 1973), we suspect that a similar biased retrieval process could be a mechanism that enables external explanations to increase cynicism about humanity. The idea is that the "bad actors" composing external explanations are disproportionately brought to mind when people report their feelings about humanity. Therefore, we suspect that providing additional, relevant information from which to judge humanity could either

prevent or override the biased judgments of humanity brought about by the external explanations. We propose to test these ideas by (a) providing a subtle reminder of participants' humanity, which could motivate participants to override the external explanations/cynicism link by anchoring the judgments on the valued self, and/or (b) by providing information about an extreme exemplar of human benevolence, that might prevent the external explanations/cynicism link from ever taking place. We will briefly develop these two ideas below.

While we are hopelessly bound to humanity, people rarely, if ever, name humanity as their preferred ingroup. In fact, Lickel, Hamilton and Sherman (2001) studied lay theories of groups and found that the abstraction of group extended only as far as loose affiliations of interests (e.g., Coltrane fans). Not once did they suggest that species-level affiliation was seen as a viable ingroup. However, studies of de- and infra-humanization showed that strong identifiers from diverse groups report believing that their ingroup possesses more human-like qualities than do out-group members (Castano & Kafta, 2009). What does this mean? It could be that our species-level affiliation is merely a distant, abstract concept concealed in our allegiances to ingroups, but present nonetheless.

One implication of the idea that we *are* human but that "*being human"* is not a salient self-categorization is that subtly reminding people of their humanity might override the biased retrieval process underlying the external explanations/cynicism link. With such a reminder, the information that comes to mind when participants are asked about humanity should be less disproportionately influenced by the bad actors in the external explanations because it will include information about the (presumably valued)

self. We suspect that in the face of negative information that threatens their now "human ingroup" (i.e., the external explanations), participants should be motivated to use the self as a strong positive anchor in order to "defend" what has become the ingroup. On the other hand, because the negative information that threatens their now "human ingroup's" value would inherently reflect on the self, they could just as likely be motivated to preserve their own positive self image.

There is, however, an alternative possibility. It could be that providing any additional positive information deemed applicable to one's judgment of humanity might also disrupt the external explanations/cynicism link. For example, being reminded of an extreme exemplar of human benevolence (e.g., Gandhi), might prevent the link between exposure to multiple external explanations and cynicism toward humanity from ever forming. Study 1 included subtle manipulations of the sort just described to test whether such manipulations are capable of breaking the link between external explanations and cynicism.

Method

Participants.

Eighty-six (43 female) undergraduate students at Lehigh University participated in order to fulfill the requirements of an introductory psychology course. One participant was excluded from the final analysis because she indicated that she did not take her participation in the study seriously. Thus, our final sample consisted of 85 participants (42 females).

Procedure.

Participants arrived individually at the lab, provided informed consent, and were introduced to a "study of social memory" (as in Gill & Getty, 2010). They learned that they would first read about some behaviors, then do some filler tasks, and then take a memory test regarding the behaviors they read about. This cover study enables us to provide participants with an apparently sound reason why they would be reading about behaviors but not responding to them.

Following these instructions, regardless of condition, participants were presented with six vignettes describing the deceitfulness, hypocrisy, authoritarian bullying, and criminality of six different transgressors (taken from Gill & Getty, 2010). Explanations for the transgressions were manipulated such that in the *internal explanations condition* (N = 43) each transgression was followed by an internal-cause explanation. In contrast, in the *external explanations condition* (N = 42) each transgression was followed by an external-cause explanation. All the vignettes for each condition can be found in Appendix A.

After reading the vignettes, participants were quickly moved to a different room to complete the "filler task," which was presented as a survey packet. Our primary manipulation of interest—the *categorization task*—was the first "filler task."

Aron, Aron and Smollan's (1991) Inclusion of Other in the Self (IOS) Scale, was adapted and modified for this purpose. Traditionally, the IOS Scale consists of seven sets of circles overlapping at varying degrees. At one end of the scale, the circles are completely separate, representing no connection between the self and other. On the other end of the scale, the circles almost completely overlap, representing deep connection or

"oneness" between the self and other. Participants choose the set of circles that best represent their relationship with another person of interest.

For our purposes, the IOS was modified in two different ways. First, participants were instructed to strictly indicate the extent to which a presented exemplar (e.g., "fern"; "self") is an example of a given category (e.g., "living thing"; "human being") rather than indicating any emotional closeness. Second, rather than responding to the entire scale, participants were only allowed to respond to a scale representing approximately 70%, 85%, and 100% overlap. The reason is that we really wanted to make it clear that the participant *is* human (and a fern *is* a living thing) by providing no choices for claiming otherwise. In all, the IOS was modified into three versions: (1) the *control condition* (*N* = 27), in which participants indicated the extent to which a "fern" is an example of a living thing; (2) the *self condition* (*N* = 30), our primary test condition, in which participants indicted the extent to which the "self" (i.e., the respondent) is an example of the category "human being" (*N* = 30); (3) the *Gandhi condition* (N = 28), in which participants indicated the extent to which Gandhi is an example of the category "human being." Each version of the modified IOS scale can be found in Appendix B.

After participants completed the categorization task, they completed the Cynicism Towards Humanity Scale (CTHS; Gill & Getty, 2010; see Appendix C). This was used to measure feelings toward humanity, our primary dependent variable of interest. The CTHS consisted of 4 items to which participants responded on a 5-point scale with endpoints of 1 (*Strongly Disagree*) and 5 (*Strongly Agree*). For example, one statement read: *The way people act sometimes can really make me feel cynical and pessimistic about humanity*, whereas another read, *I think, for the most part, people are good and*

admirable (reversed). Following appropriate reverse scoring of one item, the items were averaged to form a single index on which high scores represented cynical attitudes toward humanity (M = 3.45, SD = .66, α = .69).

The CTHS is a new measure still being developed. Because of this, we thought it beneficial to include another measure that taps into feelings of cynicism. Therefore, we included an additional dependent measure, The Beliefs about Human Nature Scale (BaHN, Gill, 2010; see Appendix D). The scale consisted of 12 items to which participants responded on a 5-point scale with endpoints of 1 (*Strongly Disagree*) and 5 (*Strongly Agree*). The BaHN scale provides reliable and relatively orthogonal indices of the extent to which people believe humanity is *aggressive* (M = 3.16, SD = .74 α = .73), *prosocial* (M = 3.72, SD = .67, α = .80) and *selfish* (M = 3.65, SD = .67, α = .81). We reasoned that any increase in beliefs about the aggressive or selfish nature and/or decrease in beliefs about the prosocial nature of humanity could be construed as increased cynicism towards humanity. Indeed, CTHS cynicism correlated strongly with BaHN aggression scores (r = .46, p < .0001) and BaHN prosociality scores (r = -.383, p < .0001), but not with BaHN selfishness scores (r < .1, p > .5).

Finally, participants returned to their original testing room to report their memories of the transgression vignettes. Then, post-experimental probing was conducted. For this purpose, after reporting their recollections, participants were asked to (1) describe what they believe the purpose of the study may "really" have been, and (2) whether they took their participation seriously. As previously indicated, one participant was excluded from analysis for indicating that she did not take her participation seriously.

Results

CTHS and BaHN scores acted as our dependent variables. We predicted an explanation (internal, external) by categorization task (control, self, Gandhi) interaction. We expected the pattern to be that participants in the *self* and *Gandhi* conditions would report equivalent, low levels of cynicism in response to the transgression vignettes regardless of explanations condition. In contrast, participants in the control condition (i.e., *fern-as-living thing*) would respond in line with the original Gill & Getty (2010) finding: Greater cynicism in the external explanations condition than in internal explanations condition.

To test this hypothesis, we began by analyzing our data in a 2 (sex: male, female) X 2 (explanations for transgressions: internal, external), X 3 (categorization task: control, Gandhi, self) mix model multivariate analysis of variance (MANOVA) with our four measures of cynicism (i.e., CTHS and the three subscales of the BaHN, respectively) entered as our within subjects factor. The test revealed a main effect of our cynicism measures (*Wilks' Λ* = .709, $F(3, 72) = 9.851$, $p < .001$). We next conducted pairwise comparisons between each of the dependent variables to tease apart this main effect (see Figure 1 for the means). The comparisons revealed that belief in humanity's inherent prosociality was significantly higher than CTHS cynicism ($t(85) = 2.219$, $p = .029$, $d = .397$) or belief in humanity's inherent aggressiveness ($t(85) = 4.727$, $p < .001$, $d = .779$), but not significantly different from belief in humanity's inherent selfishness ($t\ (85) < 1$, $p > .5$). Furthermore, belief in humanity's inherent aggressiveness was the lower that CTHS cynicism ($t\ (85) = 3.615$, $p = .001$, $d = .41$) and belief in humanity's inherent selfishness, ($t\ (85) = 4.819$, $p < .001$, $d = .668$). Finally, CTHS cynicism was also lower than belief in humanity's inherent selfishness ($t(85) = 2.02$, $p = .047$, $d = .298$).

The test of our between subject factors also revealed the predicted explanations X categorization interaction, although it was weak, F (2, 74) = 2.113, p = .128 (see Figure 2 for means). To better understand this interaction, we analyzed each indicator of cynicism toward humanity (i.e., CTHS, the three subscales of the BaHN scale) in a 2 (sex: male, female) X 2 (explanations for transgressions: internal, external), X 3 (categorization task: control, Gandhi, self) factorial analysis of variance (ANOVA).

We found clear support for our predictions in just one of the four analyses. That is, we found an explanations X categorization task interaction when we examined the aggression subscale of the BaHN, F(2, 74) = 5.25, p = .007, partial-η^2 = .13.[1] Figure 3 shows the means for each condition. We conducted follow-up planned comparisons to test the predictions noted above. These comparisons confirmed our prediction that the control condition would conceptually replicate the pattern of Gill and Getty (2010): Belief that aggression is fundamental to human nature was higher in the external explanations condition than in the internal explanations condition (t(79) = 3.13, p = .002, d = 1.2). In contrast, within the Gandhi condition there was no effect of explanation condition (t(79) < 1 p > .3). Finally, within the self condition, we found an unexpected reversal of the Gill and Getty (2010) effect. Specifically, participants in the external explanations condition reported significantly less belief that aggression is fundamental to human nature than did participants in the internal explanations condition (t(79) = 2.05, p < .05, d = .92). We did further analyses to better understand the pattern of means.

Specifically, we next examined the effects of the categorization task within each explanation condition. First, according to our idea that the self-as-human task would disrupt the external explanations/cynicism link, we tested whether belief in inherent

human aggression was less in the external/self condition than in the external/control condition. It was, $t(74) = 1.86$, $p = .066$, $d = .65$. In contrast, there was no difference between the external/Gandhi and external/control conditions, $t < 1$. Thus, these analyses suggest that the reminder of self-as-human broke the link between external explanations and cynicism whereas the reminder of Gandhi-as-human did not. We had predicted that cynicism in the internal explanations condition would be consistent across categorization conditions (see above). Was it? We found that whereas there was no difference between the internal/control and internal/Gandhi conditions, $t(74) = 1.34$, $p = .19$, there was a significant difference between the internal/control and internal/self conditions, $t(74) = 2.59$, $p = .012$ $d = 1.6$, with cynicism increasing following the reminder of self-as-human. Thus, unexpectedly, within the internal explanations condition a reminder of one's humanity actually increased cynicism.

These findings suggest that the relationship between explanations and negative beliefs about humanity is more dynamic than previously hypothesized. Whereas being reminded of one's humanity does seem to break the link between external explanations and cynicism, it also seems to *create* a link between internal explanations and cynicism! We will offer a speculative explanation of these findings in the discussion section.

Beyond the explanations X categorization task interaction, this analysis also revealed a significant sex X categorization task interaction, $F(2, 74) = 3.14$, $p < .05$, partial-$\eta^2 = .08$, and a marginal sex X explanations condition interaction, $F(1, 74) = 3.39$, $p = .07$, partial-$\eta^2 = .04$. We had no *a priori* predictions here so we performed post hoc simple effects tests to explore these interactions. First, we explored the sex X categorization task interaction. See Figure 4. Our follow-up tests revealed that this

interaction was driven by the fact that males were more cynical than females in the control condition (*LSD* = .617, $p < .05$, $d = .85$) and only in the control condition (See Figure 4). Moreover, males did not differ across the categorization conditions, whereas females in the Gandhi condition were significantly more cynical than females in the control condition (*LSD* = .661, $p = .021$, $d = .79$). This specific result might be explained by a contrast effect, which will be elaborated on in the discussion section.

Of greater interest were our follow-up contrasts exploring the sex X explanations interaction. This interaction is depicted in Figure 5. As can be seen there, collapsing across levels of the categorization task manipulation, females in the external explanations condition reported significantly greater belief in inherent human aggressiveness than did females in the internal explanations condition, *LSD* = .524, $p < .02$, $d = .69$. In contrast, males showed high and equivalent levels of cynicism across both explanation conditions (*LSD* = .18, ns). These findings suggest that, collapsing across the categorization task conditions, females showed the pattern uncovered by Gill and Getty (2010): Negative beliefs about humanity increase in response to external explanations. Males did not show this pattern but rather had negative beliefs about humanity regardless of our manipulations.

Discussion

Study 1 was based on the idea that biased retrieval processes represent one possible mechanism that enables external explanations to increase cynicism about humanity. Specifically, in line with the notion of the availability heuristic (Tversky & Kahneman, 1973), the idea is that the "bad actors" composing external explanations can—under certain conditions—be a disproportionate share of what is brought to mind

when people report their feelings about humanity. Study 1 tested the idea that a subtle manipulation of what is brought to mind when one thinks about humanity could disrupt this biased retrieval process and thereby prevent or override the external explanations/cynicism link. In particular, participants in one condition received a subtle reminder that they themselves are human, and participants in a different condition received a subtle reminder that Gandhi is human.

On one out of our four relevant dependent variables—when the dependent variable was belief in inherent human aggressiveness—we found evidence suggesting that the reminder of self-as-human did break the external explanations/cynicism link but the reminder of Gandhi-as-human did not. Specifically, whereas in the control condition, external explanations fostered increased belief in inherent human aggressiveness as compared to an internal explanations condition (conceptually replicating Gill & Getty, 2010), this pattern was reversed in a condition in which participants were subtly reminded of their own humanity. That is, following a self-as-human reminder, participants reported high levels of belief in inherent human aggressiveness in the internal explanations condition and low levels of such belief in the external explanations condition. Why might this reversal have occurred?

First, as respondents' positive self-concept becomes more relevant to their judgment of humanity (i.e., in our self-as-human categorization condition), they should have little trouble drawing positive parallels between the self and humanity. In other words, the self becomes the object by which humanity is judged. Even when one has received a subtle reminder of the self/humanity link, however, the precise manner in which the self colors one's judgment of humanity might be variable. It could be that

people are particularly motivated to draw on the self to judge humanity when faced by threatening information about humanity—as is presumed to be the case in the external explanations condition—because when that link between the self and humanity becomes salient, a sense of oneness with humanity is induced. Humanity becomes the ingroup. That is, for those in the external explanations condition whose human identity has been made salient, the vignettes might be perceived as a threat to a valued ingroup (i.e., humanity). Because respondents' humanity is available, powerful, and relevant for "defending" what has become "the ingroup," they are motivated to use the self as a strong positive anchor, thus overriding the negative effects of the external explanations—"Wait a minute…I'm human, too, and I'm pretty cool."

If in fact the reversal of cynicism in the external explanations condition was the result of judgments anchored on the self, judgments in the internal explanations condition could not be the result of the same anchoring, because the results were in the opposite direction. In contrast to the external explanations, the internal explanations might be less threatening to one's human ingroup. In the face of a more modest threat to "my human ingroup," participants might feel free to derogate the rest of humanity, or at least less motivated to defend humanity. The mechanism responsible for this effect might be the urge to see the self as unique and more valued than "the other humans." This is, of course, speculative and in need of testing.

Finally, in another condition in which participants received a subtle reminder that Gandhi is human, those in the external explanations condition showed no change in cynicism as compared to the external explanations/control condition. Thus, the reminder that Gandhi is human did not disrupt nor prevent the external explanations/cynicism link.

Why not? There could be several potential reasons. First, it might be that only when the valued self is tied to the human ingroup are people motivated to seek positive information to defend humanity and break the external explanations/cynicism link. In other words, even when reminded of Gandhi's humanity, respondents might not be motivated to take that information into consideration when judging humanity because it is not self relevant. Second, to our cohort of participants, Gandhi might not represent a salient example of extreme human benevolence. If that is indeed the case, perhaps reminding college-aged participants of the Dalai Lama's humanity, or even Bob Marley's humanity, might have the hypothesized effect. Third, if Gandhi did represent a salient exemplar of extreme human benevolence to this sample of participants, we suspect that this non-significant difference could represent a person schema confirming the already available information: When faced with external explanations for multiple transgressions (available information: humanity is full of bad actors) and/or information about Gandhi's humanity (schema: compared to Gandhi, the rest of humanity are a bunch of bad actors), respondents are likely to make the same snap judgment of humanity—humanity sucks.

A final pattern worthy of note is that males and females were differentially affected by our explanations manipulation (collapsing across the categorization task manipulation). Looking at beliefs about inherent human aggressiveness, female responses to our explanations manipulation replicated the findings of Gill and Getty (2010), with external explanations increasing negative beliefs. In contrast, male responses were not affected by our explanations manipulations. Why might this happen? One possibility is that males entered the study with pre-existing high levels of negative beliefs about humanity. Thus, they processed our vignettes—regardless of the explanations

manipulation—as further evidence for their negative beliefs or perhaps as "unsurprising." Thus, their beliefs are unaffected. In contrast, perhaps females entered the present study with more positive pre-existing beliefs about humanity. If that is the case, then as females processed the internal explanations vignettes, the transgressors might have been viewed as a strong contrast to the rest of humanity ("Most people are not nearly so bad!"), and thus their positive beliefs about humanity were preserved.

But what happens among females in the external explanations condition? Why do they become cynical? There are two possibilities. The first is that, as we suggest above, the sheer number of people behaving badly in the external explanations condition might prompt even those with a positive view of humanity to reconsider their view. On the other hand, it is possible that there is something more subtle going on. Those with a positive view of humanity presumably expect that the default for human development is that people turn out "good" and "nice." In the external explanations condition, from the vantage point of someone with positive expectations for humanity, this default is disrupted and targets who would "normally" be good people are corrupted by the "bad actors" who compose the external explanations. This might especially give rise to moral outrage in those with positive expectations of humanity because they perceive not only "lots of people doing bad things" but also the destruction of the human potential to become a decent person. This is merely speculative, of course, even if intriguing. Future research should measure pre-existing beliefs about humanity and some appropriate mediating variables (e.g., moral outrage) to collect evidence relevant to this proposed process.

Finally, why did we fail to find our predicted results on three of our four indicators of beliefs about humanity? One reason could be the low reliability of the measures. With the exception of the beliefs about human aggression measure, our measures had only marginal reliability ($\alpha(s) < .7$); therefore, we might not have been able to detect a clear signal of cynicism. Second, it might be that cynicism is an inherently elusive concept taking many forms. Cynicism might manifest as increased belief in an aggressive human nature, as seems to be the case in this study. But it could also be revealed as a strong belief in the "sinful nature" of humanity—a sense that humanity was tainted by original sin. In a less spiritual sense, it could be that cynicism is simply feeling a "negative vibe" about human nature. Of course, cynicism might be exposed as a decreased tendency to see the benevolent side of humanity. Cynicism could rear its ugly head as any one of these phenomena, and thus any particular measure might "miss the mark." Thus, future studies should explore additional ways of gauging cynical attitudes toward humanity. Also, we should note that Study 1 should not be taken as indicating that the aggression subscale of the BaHN is the best measure to tap our phenomenon of interest. Indeed, Gill & Getty (2010) found conceptually similar results using the CTHS, and results on the CTHS will also be reported below.

In sum, the findings of Study 1 were more complicated than our predictions. On the positive side, the control condition did yield a conceptual replication of Gill and Getty (2010)—i.e., more cynicism following external explanations—on one of our four measures of cynicism toward humanity. Also, collapsing across categorization tasks, female participants demonstrated the pattern found by Gill and Getty whereas male participants were uniformly cynical across explanations conditions. On the negative side,

there were unexpected effects of our categorization task—in particular, increased cynicism following internal explanations—when participants had previously categorized themselves as human. We can only offer speculations (see above) about why that might have happened.

Study 2: The Impact of Exculpatory Information

("He hates who he was when he did those deeds.")

In the external explanations condition, there is both a focal transgressor (Transgressor 1) and an additional transgressor (Transgressor 2) who is the external force "pushing" Transgressor 1 toward bad acts. It seems likely that the external explanations/cynicism link depends on participants viewing Transgressor 2 as inherently malevolent and blameworthy. Thus, altering that inference—getting participants to view Transgressor 2 as less blameworthy—should disrupt the external explanations/cynicism link. This would shift responses from the "complex blame game" alluded to above toward something that is "blame free." How might we change blame-laden perceptions of Transgressor 2?

One possibility would be to directly provide information about external explanations that have compelled Transgressor 2, or perhaps to indirectly activate such external explanations by having participants take the perspective of Transgressor 2 (e.g., Vescio, Sechrist, and Paolucci, 2003). The problem with this approach, we suspect, is that it merely adds another layer to the "blame game," and shifts blame from Transgressor 2 to Transgressor 3 and thus should still increase cynicism toward humanity (i.e., because *someone* is being viewed as malevolent).

On the other hand, several lines of research have linked forgiveness and attitude change towards transgressors with expressions of sincere apology (Darby & Schlenker, 1982; Davis & Gold, 2011). In fact, Darby and Schlenker (1982) demonstrated that repentant transgressors are perceived to be less blameworthy and less deserving of punishment. Indeed, across two experiments, Darby and Schlenker found that even when a transgressor was highly responsible for intentional acts that were motivated by bad intentions and those acts resulted in severe consequences (destroying property), if the transgressor expressed some remorse for their acts through apology, they were treated with compassion. Moreover, as the elaboration of the apology increased (i.e., when they were "really, really sorry"), so too did the compassionate response: The transgressor was judged more positively, blamed less, and forgiven for his misdeeds.

In light of these findings, we suspect that a more promising possibility for breaking the external explanations/cynicism link involves inducing participants to see Transgressor 2 as "acting based on a lack of understanding," having "overcome bad habits of the past," and currently apologetic and regretting his or her past transgressions (Darby & Schlenker, 1982; Davis & Gold, 2001; also see Pizarro, Uhlmann, & Salovey, 2003), thereby *shutting off blame* rather than merely *shifting it around*.

Method

Participants.

Fifty-eight (17 female) undergraduates from Lehigh University participated in order to fulfill the requirements of an introductory psychology course.

Procedure.

The procedures were essentially the same as Study 1, featuring the same "social memory" cover story. As in Study 1, some participants were in the *internal explanations* conditions (N = 15) and others were in the *external explanations* condition (N = 15). The major difference from Study 1 was the addition of two more conditions: An *external-explanations-plus-regret* condition (N = 14) and a *transgressor-3* condition (N = 14). To create the *external-explanations-plus-regret* condition (hereafter called the *regret* condition), three of the six external explanations vignettes were modified to indicate that "Transgressor 2" had "learned from his or her mistakes" and had come to regret previous harmful behaviors. To create the *transgressor-3* condition, the external explanations condition was modified to provide information suggesting that a third transgressor had transgressed against transgressor 2, and this is why transgressor 2 was mistreating the focal target in the vignette. See Appendices E and F for the modified explanations conditions.

After being randomly assigned to one of these four conditions and reading the appropriate vignettes, participants proceeded to the ostensible "filler tasks," the first of which was the Cynicism Towards Humanity Scale (CTHS; Gill & Getty, 2010; see Appendix C). However, in Study 2 this measure showed low reliability (M = 3.46, SD = .60, α = .434). Reducing the measure to a 2-item scale using the first two items—which were the only items significantly correlated with each other—increased its reliability (α = .64), but it remained marginal at best. This two-item scale will be used below (M = 3.72, SD = .84). Again, the Beliefs about Human Nature Scale (BaHN, Gill, 2010; see Appendix D) was included. Subscale scores were once again created for beliefs that humanity is inherently *aggressive* (M = 3.43, SD = .75 α = .78), *prosocial* (M = 3.56, SD

= .62, α = .69) and *selfish* (*M* = 4.07, *SD* = .54, α = .67). Unlike in Study 1, CTHS cynicism (based on two items here) did not significantly correlate with any of the BaHN dimensions (all *rs* < .2, *ns*).

After completing these and the other tasks mentioned above, participants returned to their original testing room to report their memories of the transgression vignettes. Finally, post-experimental probing was, again, conducted. For this purpose, after reporting their recollections, participants described what they believe the purpose of the study may "really" have been, and whether they took their participation seriously. All reported taking the study seriously.

Results

We predicted that we would replicate the Gill and Getty (2010) finding of increased cynicism in the external as compared to internal explanations condition. Furthermore, we predicted that the regret manipulation would shut off the tendency for external explanations to evoke cynicism, and thus that the regret condition would look like the internal explanations condition (i.e., relatively low cynicism). Finally, we predicted that the transgressor-3 condition would not break the external explanations/cynicism link, and thus the transgressor-3 condition should look like the external explanations condition (i.e., high cynicism).

To test these predictions, we began by analyzing our data in a 2 (sex: male, female) X 4 (explanations for transgressions: internal, external, regret, transgressor 3), mixed model multivariate analysis of variance (MANOVA) with our four measures of cynicism (i.e., rCTHS and the three subscales of the BaHN, respectively) entered as our within subjects factor. The test revealed a main effect of our cynicism measures (*Wilks' Λ*

= .494, $F(3, 72) = 16.404$, $p < .001$). We next conducted pairwise comparisons between each of the dependent variables to tease apart this main effect (see Figure 6 for the means). The comparisons revealed that belief in humanity's inherent selfishness was significantly higher than CTHS cynicism ($t(57) = 2.818$, $p = .007$, $d = .502$), belief in humanity's inherent aggressiveness ($t(57) = 6.379$, $p < .001$, $d = .983$), and belief in humanity's inherent prosociality ($t(57) = 4.731$, $p < .001$, $d = .87$). Furthermore, belief in humanity's inherent prosociality was statistically indistinguishable from belief in humanity's inherent aggressiveness and CTHS cyncism ($ts(57) < 2$, $ps > .2$). Finally, belief in inherent human aggressiveness was the lower that CTHS cynicism ($t(57) = 2.07$, $p = .043$, $d = .363$). No other significant between or within effects were found with this analysis.

Next, using the 2-item CTHS and BaHN scores as our dependent variables, we analyzed each dependent variable independently in a 2 (sex: male, female) X 4 (explanation condition: *internal explanation*, *external explanation*, *regret*, *transgressor 3*) factorial analysis of variance (ANOVA). In the analyses involving the BaHN subscales, there was no evidence of any effect of our explanation conditions, all Fs < 1, nor did explanations interact with participant sex, Fs < 1.

In our analysis of the CTHS, there were no main effects (ps > .15; see Figure 7 for means). There was, however, a significant sex by explanation condition interaction, $F(3, 50) = 3.41$, $p = .02$, partial-$\eta^2 = .17$. Because the cell sizes for female participants were never greater than $N = 6$ (and sometimes as small as $N = 3$), we will not dwell on the pattern of means from female participants. Suffice it to say that they did not align with predictions. Cell sizes for males were around 10 per cell. Contrary to our predictions,

follow-up comparisons found that males in the external explanations condition (M = 3.25, SD = .95) were *less* cynical than were males in all other conditions (Ms = 4.06, 4.00, 3.83, SDs = .88, .71, .60, ts(50) = 2.36, 2.20, 1.75, ps = .022, .034, .085, ds = .88, .89, .73) for the internal explanations, transgressor 3, and regret conditions, respectively). See Figure 8.

Discussion

The purpose of Study 2 was to examine the extent to which a tendency to direct blame and anger toward Transgressor 2—to fail to consider that perhaps Transgressor 2 is in some sense "innocent" —is a contributing factor to the external explanations/cynicism link. Specifically, we theorized that participants likely view transgressor 2—the negative, external influence on the focal transgressor of each vignette—as inherently malevolent and blameworthy, so they blame and become angry toward these secondary bad actors, which contributes to a dim view of humanity. Thus, we reasoned that if transgressor 2 could somehow be portrayed as "innocent" and sincerely apologetic this would break the external explanation/cynicism link. Therefore, Study 2 varied the presence of exculpatory information about transgressor 2, in particular whether there were external explanations contributing to the badness of transgressor 2 and whether transgressor 2 had become a "new person" with regret about his or her earlier bad acts.

Unfortunately, we were unable to find evidence supporting our hypotheses. In fact, the results of Study 2 make little sense at all. Indeed, with just one exception, even the items composing the CTHS did not correlate with each other! However, it might still

be possible to extract some meaning from these data if we take into account the heavily male skew of our sample.

The results related to female responding are unintelligible. There were simply too few females contributing data within each cell. On the other hand, the patterns coming from male participants are based on sufficient numbers to warrant discussion. Males reported greater cynicism in the internal explanations, regret and transgressor-3 conditions than in the external explanations condition, precisely the opposite of the pattern found by Gill and Getty (2010), and different from Study 1 in which males were unaffected by the explanations manipulation. Why would males respond with significantly less cynicism to external explanations than the other three conditions? The pattern of male responding might be due in part to pre-existing beliefs about humanity as well as a similar biased retrieval process in the face of overwhelming transgression, as described in Study 1. We will elaborate on these potential explanations.

Concrete examples of antisocial behavior (selfishly motivated people doing bad things)—as are depicted in the internal explanations condition—might reinforce cynical feelings among those who already possess negative pre-existing beliefs about humanity. On the other hand, because external explanations are more abstract and complex (e.g., Gilbert, Pelham, & Krull, 1988), males in the external explanations condition might be unable to "latch on" to a concrete example of wickedness, especially if their focus is on transgressor 1 and not transgressor 2. For this reason, their cynicism might be unchanged or even suppressed to some extent. This seems reasonable because in the external explanations condition, the main characters are essentially victims who have been "pushed" to depravity. They might even be perceived as "once innocent" victims. Just as

non-cynics might be unable to deny humanity's potential for evil in the face of evil acts, so too might the cynic be unable to deny the *humane* side of humanity in the face of "innocence corrupted." But why, then, is this not true for the transgressor 3 condition?

The increased cynicism among males in the transgressor 3 conditions might be due to the biased retrieval process suggested as an explanation for the original Gill & Getty (2010) findings. While males in Study 1 might have been unaffected by our manipulation, perhaps with the addition of transgressor 3, the sheer number of people behaving badly might have overwhelmed their capacity to see the "innocence corrupted" and left them seeing simply "lots of people doing bad things to lots of other people." This might have increased pre-existing high levels of cynicism to even higher levels ("Wow, people suck more than I originally thought!").

What is even more difficult to explain, however, is why exculpatory information in the regret condition lead to exacerbated cynicism as compared to external explanations alone. A likely explanation could be that the marginal effect is simply an artifact of the statistical analysis, or a type I error. However, the explanation noted above might also account for this finding. It could be that adding exculpatory information only minimally affects males, perhaps because males have pre-existing negative beliefs about humanity. Those with such pre-existing beliefs might be unmoved by exculpatory information and perhaps even perturbed by expressions of regret because they doubt its sincerity.

Study 3: Concrete vs. Abstract Construal

Construal level theory (CLT) is founded on the idea that observers can represent any object, action or event at multiple levels of construal (Trope, in press). At a low level of construal, observers are more inclined to consider the subordinate details of an object

or event, or consider how the object is unique. On the other hand, when scrutinizing an object or event at a high level of construal, observers form abstract representations based solely on the relevant features of the object or event. For example, at a high level of construal, a gun can be abstractly construed as a "weapon" or a "means of self-defense." On the other hand, at a low level of construal, when people think in more concrete terms, they interpret the gun of interest in specific details (i.e., a Ruger Alaskan 44 caliber snub-nosed revolver).

Fujita, Trope, Liberman and Levin-Sagi (2006) found that self-generated lists of either superordinate or subordinate objects (like the gun described above) lead to differences in construal preferences for an unrelated situation description (Study 3a). Specifically, after generating one of the two category lists, participant were given a written description of an event and then asked to choose between two new conceptualizations of the same event. One described the event in more global, abstract terms (high-level construal) while the other emphasized specific details and used more concrete terms (low-level construal). They found that when participants generated a list of superordinate categories, they preferred the high-level construal conceptualization, whereas participants who generated a subordinate category list preferred the low-level construal conceptualization.

We suspect that a similar bias processing of the information in the vignettes might be an additional enabling condition for an external explanations/cynicism link. Specifically, we suspect that one of the potential mechanisms connecting external explanations to cynicism requires that people think abstractly (high-level construal) rather

than concretely (low-level construal; see Trope, in press). The reason is that when people think in abstract terms they are more likely to consider superordinate categories.

In our paradigm, we expect that when participants are forming judgments at a low level of construal, the "bad actors" in the external explanations will be seen as "a (limited) group of bad actors" and thus will not influence judgments of the abstract category "humanity." When participants are forming judgments at a more abstract level, however, the "bad actors" will be seen as "a group of bad human beings." Thus, this implies that we might be able to turn off the external to cynicism link by inducing people to think more concretely.

Method

Study 3 was a 2 (explanations for transgressions: internal, external) X 2 (construal: concrete, abstract) between-subjects design.

Participants.

Forty (19 females) Lehigh University undergraduates participated for credit in their introductory psychology course.

Procedure.

Once again, the procedures for Study 3 were similar to those used in Study 1. One change is that there was a slightly embellished social-memory cover story. Participants were informed that the purpose of the study is to examine "the relation between people's skill at making cognitive associations and their memory for social information." Prior to starting the social memory task, participants completed a "timed cognitive associations task" supposedly designed to measure participants' efficiency at making quick associations. In reality, the task was designed to prime high-level construal (abstract

thinking) or low-level construal (concrete thinking). In previous studies, this procedure was used to successfully demonstrate that priming high- versus low-level construal affects subsequent judgments in domains as diverse as self-control (Fujita, Trope, Liberman & Levin-Sagi, 2006, Studies 3a & 3b) and probability estimates (Wakslak & Trope, 2009). Therefore, participants were randomly assigned to receive one of these two priming manipulations to form an *abstract construal level* condition (N = 20) and a c*oncrete construal level* condition (N = 20).

The construal level manipulation involved presenting participants with 30 words for which they must complete category associations. In the low-level construal condition, for each word (Beer, King, Mountain, etc.), participants completed the following sentence: An example of _____is what? For example, an example of a Beer is what? Answer: *Ninkasi's Tricerahops*. In the high-level construal condition, for each word, participants completed the following sentence: _____ is an example of what? For example, Beer is an example of what? Answer: a tasty beverage. See Appendix G.

After completing the construal task, the remaining procedures were identical to Study 1 (although, of course, the categorization task was deleted). That is, participants were in the *internal explanations condition* (N = 20) or the *external explanations condition* (N = 20). After reading vignettes they were ushered to another room for the "filler tasks,"" which consisted of the same survey packet as in Study 2, which included the CTHS and the BaHN scales. Once again, the CTHS showed low reliability (M = 3.38, SD = .58, α = .434). Reducing the measure to a 3-item scale by removing the first item increased its reliability (α = .60), but the reliability remained marginal at best. The BaHN scale was scored as in the earlier studies, providing separate indexes of the extent to

which humanity is viewed as *aggressive* (M = 3.04, SD = .79 α = .81), *prosocial* (M = 3.58, SD = .69, α = .80) and *selfish* (M = 3.86, SD = .61, α = .83). In Study 3 as in Study 2, the shortened CTHS did not significantly correlate with any of the BaHN dimensions (all $rs < .2$, *ns*).

Finally, post-experimental probing was conducted. For this purpose, after reporting their recollections, participants described what they believe the purpose of the study may "really" have been, and whether they took their participation seriously. All participants indicated that they took their participation seriously; therefore, none were deleted.

Results

Our major prediction was a significant explanation by construal level interaction, with the effect of external explanations on heightened cynicism happening only in the abstract construal level condition. To test this prediction CTHS and BaHN scores acted as our dependent variables.

To test these predictions, we began by analyzing our data in a 2 (sex: male, female) X 2 (explanations for transgressions: internal, external) X 2 (construal: abstract, concrete), mix model multivariate analysis of variance (MANOVA) with our four measures of cynicism (i.e., rCTHS and the three subscales of the BaHN, respectively) entered as our within subjects factor. The test revealed a main effect of our cynicism measures (*Wilks' Λ* = .458, $F(3, 30) = 11.837$, $p < .001$). We next conducted pairwise comparisons between each of the dependent variables to tease apart this main effect (see Figure 9 for the means). The comparisons revealed that belief in humanity's inherent selfishness was significantly higher than CTHS cynicism ($t(39) = 3.91$, $p < .001$, d =

.847), belief in humanity's inherent aggressiveness ($t(39) = 5.536$, $p < .001$, $d = 1.153$), and belief in humanity's inherent prosociality ($t(39) = 3.069$, $p < .004$, $d = .439$). Furthermore, belief in humanity's inherent prosociality was greater belief in humanity's inherent aggressiveness (t (39) = 3.327, $p = .002$, $d = .711$), but statistically indistinguishable from CTHS cynicism ($t(39) = 1.536$, $p = .133$). Finally, belief in inherent human aggressiveness was the lower that CTHS cynicism ($t(39) = 1.883$, $p = .067$, $d = .385$).

Lastly, the MANOVA revealed a significant sex X within subjects cynicism interaction (*Wilks' Λ* = .558, $F(3, 30) = 7.909$, $p < .001$). Our follow up factorial analyses of variance will address this interaction. Thus, next the dependent variables were independently submitted to a 2 (sex: male, female) X 2 (explanation condition: internal explanation, external explanation) X 2 (construal condition: abstract, concrete) factorial analysis of variance (ANOVA). Analysis of the BaHN subscales revealed no significant effects, all Fs < 2.2, ps > .15, other than a main effect of participant sex suggesting that females ($M = 3.92$) view humanity as more intrinsically prosocial than do males ($M = 3.32$), $F(1, 27) = 7. 11$, $p = .013$.

Results involving the CTHS were more complex, although they did not support our primary prediction of an interaction between construal level and explanations, $F < 1$. The analysis did, however, reveal a significant main effect of sex, with males ($M = 3.54$, $SD = .57$) being more cynical than females ($M = 3.09$, SD = .67), $F(1, 32) = 4.65$, $p < .04$, partial-$\eta^2 = .13$. The main effect of sex was qualified by a significant sex by explanations condition interaction, $F(1, 32) = 7.43$, $p = .01$, partial-$\eta^2 = .19$. See Figure 10 for the means. Follow-up contrasts revealed that, as in Study 1, females showed the external

explanations/cynicism link uncovered by Getty and Gill (2010): They were more cynical in the external explanations condition than in the internal explanations condition, $t(36) = 2.68$, $p = .011$, $d = 1.25$. Males, on the other hand, were unaffected by the explanation manipulation, $t(36) = 1.4$, $p = .17$, just as in Study 1.

Discussion

In Study 3, we tested the idea that the mode in which people process information in the vignettes might be an additional enabling condition of the external explanations/cynicism link. Specifically, the mechanism connecting external explanations to cynicism might require people to think abstractly (high-level construal) rather than concretely (low-level construal; see Trope, in press) about the overall message of the vignettes. The reason is that when people think in abstract terms they are more likely to consider superordinate categories. We expected that when participants formed judgments at a low level of construal, the "bad actors" in the external explanations would be seen as individual bad actors and thus would not influence judgments of the abstract category "humanity." At a more abstract level, we predicted that they would be seen as "a group of bad human beings." Thus, we can turn off the external explanations/cynicism link by inducing people to think more concretely.

It seems that our construal manipulation failed to do anything at all. On the other hand, our failure to support our predictions could be evidence suggesting that levels of construal is not viable mechanism responsible for the external explanations/cynicism link. Nevertheless, Study 3 did uncover the same explanations by participant sex interaction that was uncovered in Study 1: Females seem to behave in accord with our predictions of an external explanations/cynicism link whereas males are cynical no matter

what. Because we found this same pattern across two studies, we will attempt to interpret it in the General Discussion.

General Discussion

We are interested in the psychology of social explanations and the ironic possibility that external explanations—which increase compassionate responding to those who commit bad acts (e.g., Gill & Andreychik, 2009; Weiner, 2006)—might foster cynicism toward humanity even as they increase compassion toward specific transgressors. As Gill and Getty (2010) suggested, external explanations might have this ironic side effect because they typically point to how the transgressing individual was harmed, hindered, or otherwise mistreated by other people. As one accumulates in her mind instances of "people driving other people toward bad deeds," one can come to view humanity in a negative light. In essence, external explanations can lead us to spare the individual transgressors the rod, but to damn humanity in the process.

We proposed several studies to examine how to break this external explanations/cynicism link. In Study 1, we tested the possibility that providing subtle reminders of positive instances of humanity (e.g., Gandhi, the self) might break this link. The rationale was that the external explanations/cynicism link depends on the operation of the availability heuristic (Tversky & Kahneman, 1973), with external explanations increasing cynicism because they create more available instances of "bad actors" in one's mind. Subtle reminders of positive instances of humanity should alter what is cognitively available, and thus nullify negative effects of external explanations.

We found evidence supporting this line of reasoning on one of four dependent variables. Specifically, although control participants reported increased belief in inherent

human aggressiveness following exposure to multiple external rather than multiple internal explanations (conceptually replicating Gill & Getty, 2010), external explanations showed no such negative effect when participants had been subtly reminded that they themselves are human. In fact, the subtle reminder of one's humanity reversed the Gill & Getty (2010) effect, and increased negative beliefs about inherent human aggressiveness in the *internal* explanations condition. Admittedly, we do not fully understand the cause of this reversal. Above, we speculated that it might occur because when presented with threatening information about humanity (e.g., in the external explanations condition), individuals whose human identity has been made salient become motivated to draw on the self (a powerfully positive anchor from which to judge humanity) rather than letting the negative information hold sway over the judgment. If they did not "defend humanity" in this manner, the negative information about humanity would reflect on the self because of the currently active human identity.

Interestingly, a reminder of Gandhi's humanity failed to disrupt the external explanations/cynicism link. We have speculated on several potential reasons for this. For example, preserving the valued self or a valued ingroup might be the only factors motivating people to defend humanity when threatened. However, the non-significant finding in this condition could simple be due to a cohort effect, in the sense that Gandhi might not represent a salient example of extreme human benevolence in the minds of Lehigh University undergraduates. However, we have speculated that other historical figures like Bob Marley and/or the Dalai Lama might represent more salient examples of human benevolence to this sample of participants. If so, then reminding these participants of either Bob Marley's humanity or the Dalai Lama's humanity might have the

hypothesized effect. Then again, extreme exemplars of human benevolence might have the reverse effect, leading participants to respond more negatively to humanity. If so, then our failure to find our hypothesized effect might have been due to a person schema confirming the available information about humanity. In other words, if Gandhi does represents a salient exemplar of extreme human benevolence, then compared to Gandhi, the rest of humanity is a bunch of bad actors, which confirms the information communicated by multiple external explanations for transgressions: humanity is overwhelmingly composed of bad actors.

In Study 2 we tested the hypothesis that a tendency to direct blame and anger toward Transgressor 2—to fail to consider that perhaps Transgressor 2 is in some sense "innocent" —is a contributing factor to the external explanations/cynicism link. Specifically, we theorized that participants likely view transgressor 2—the negative, external influence on the focal transgressor of each vignette—as inherently malevolent and blameworthy, and that this evaluation drives their increased cynicism toward humanity. Thus, we reasoned that regretful transgressor-2s would break the external explanation/cynicism link. We did not, however, find evidence confirming our hypothesis. We did find evidence suggesting that males—perhaps because of pre-existing negative beliefs about humanity—might be unmoved by exculpatory information. Future research might consider having participants report their beliefs about why people apologize. It could be that cynics do not see those who have "come to regret" some past grievance as sincere and are therefore unmoved.

In Study 3 we tested the idea that the mechanism connecting external explanations to cynicism might require people to think abstractly (high-level construal) rather than

concretely (low-level construal; see Trope, in press) about the overall message of the vignettes. Our reasoning was centered on the idea that when people think abstractly they are drawn to make inferences about superordinate categories. We reasoned that when participants form judgments at a low level of construal, the "bad actors" in the external explanations are seen as individual bad actors and thus will not influence judgments of the abstract category "humanity." At a more abstract level, we predicted that those bad actors would be seen as "a group of bad human beings." Thus, we expected to turn off the external explanations/cynicism link by inducing people to think more concretely. Unfortunately, our results did not support these predictions. The most likely reason: our construal manipulation failed to do anything. On the other hand, it could simply be that abstract construal is not an important mechanism enabling the external explanations/cynicism link.

Still, in both Studies 1 and 3 we replicated the findings of Gill and Getty (2010). In the control condition of Study 1, we demonstrated an external explanations/cynicism link. Furthermore, collapsing across our categorization task manipulation in Study 1, we replicated Gill and Getty among females but not males. Study 3 also replicated the Gill and Getty phenomenon but, once again, only among female participants. Among females, negative beliefs about humanity increase in response to external explanations. Males, on the other hand, appeared to have negative beliefs about humanity regardless of our manipulations. We believe that a likely explanation for these differences is rooted in pre-existing beliefs about humanity: Males and females enter our study with different assumptions about the worth of humanity (see Discussion following Study 1). Male beliefs were perhaps unaffected by our manipulations because they processed our

vignettes—regardless of the explanations—as further evidence for their negative beliefs or perhaps as "unsurprising." In contrast, perhaps females had more positive pre-existing beliefs about humanity. If that is the case, then females processed the transgressors in the internal explanations vignettes as a strong contrast to the rest of humanity, preserving their positive beliefs about humanity. On the other hand, females in the external explanations condition may have become cynical for two possible reasons: (1) when faced with multiple external explanations for transgressions, negative information about humanity becomes overwhelmingly available, leading to a cynical view of humanity and/or (2) the perception that one person having corrupted another person lead to moral outrage, which was expressed as increased cynicism.

While we did not find evidence for all of the mechanisms we examined, we have illuminated one: The valued self. People's human identity influences how they react to threatening information about humanity's value. This suggests the exciting possibility that people can relate to humanity as an ingroup member; a dynamic relationship overlooked in the literature. We have laid the foundation to consider this relationship in more detail. We have already speculated that there are a number of possible mechanisms by which human identity and pre-existing beliefs about humanity might interact when interpreting social explanations. We are encouraged to trudge on.

We have suggested that as humanity's value becomes more relevant to one's positive self-concept (i.e., when human identity has been activated), respondents should be motivated to use their pre-existing positive self-concept to combat any highly negative information about humanity's value (i.e., external explanations pointing to mass quantities of people harming other people). However, the question is, of course, whether

the mechanism responsible for anchoring on the valued self is the motivation to save humanity's face or one's own from the damning information? We suspect this anchoring occurs among individuals with positive pre-existing beliefs about humanity (i.e., non-cynics) in order to preserve humanity's (the ingroup's) positive image, while individuals with negative pre-existing beliefs about humanity (i.e., cynics) might do so only to preserve their own positive self-image.

When non-cynics are reminded of their humanity, they likely become strongly identified with their human ingroup, ready to defend and preserve the group's positive identity in the face of threatening information. They likely do so by using the most salient positive information available: The valued self. Cynics, on the other hand, might respond more like weak identifiers who begrudgingly acknowledge their own ties to their human ingroup with little motivation to preserve the group's positive image. Cynics, like non-cynics, likely call on salient positive information (the valued self) when faced with threatening information about humanity when their human identity is activated, but do so only to preserve their own positive self-image that would otherwise be polluted through their reluctant ties to the tainted group. Be that as it may, if given the change to distance themselves from the rest of humanity, cynics could be inclined to continue derogating humanity (respond cynically) as an attempt to differentiate themselves from their human brethren (*I am human, the best kind of human, a unique specimen—the rest are riffraff*). The mechanism responsible might ultimately be the urge to see the self as a unique, valued creature.

Indeed, the urge to see the self as a unique valued creature, differentiated from the rest of humanity is likely pronounced among cynics. It is possible that when their human

identity is activated, cynics might even experience dissonance when their cynical view of humanity clashes with their positive view of the self. To break the tension brought about by one's undeniable humanity (and hence essential connection to all those "despicable others"), they might want to differentiate themselves the best they can, while non-cynics would be less inclined to do so. Thus, if given the chance to distance the self from humanity, we would likely see cynics express greater cynicism in the face of threatening information about humanity as compared to non-cynics and cynics who are not given the opportunity to distance themselves from humanity.

One way to test these speculations would be to create a study that combines a measurement of dispositional cynicism and the opportunity to differentiate the self from humanity when confronted with damning information about humanity's value. The result would be a 2(dispositional cynicism: *cynics* versus *non-cynics*) X 2(differentiation: opportunity to *differentiate from humanity* versus *no opportunity to differentiate from humanity*) experimental design. Dependent variables would include the BaHN and/or CTHS as primary indices of cynicism, while actual differentiation scores (explained below) could be used as an indication of psychological distancing.

To conduct this study, we would first need to measure participant's dispositional cynicism in order to identify cynics from non-cynics in a pretest session. This could be facilitated by presenting participants with the BaHN and/or CTHS and then recruiting participants who scored one standard deviation or greater above the mean to represent cynics and participants who scored one standard deviation or lower below the mean to represent non-cynics. Then, in an ostensibly unrelated follow-up session, using the same social memory cover story noted above, we would present participants with threatening

information about humanity's value (i.e., the external explanations vignettes). Next, in order to activate their human identity, participants would complete only the self-as-human-being categorization portion of the IOS measure (Aron, Aron & Smollan, 1991) presented in Study 1 above. Then, we would manipulate whether participants are given the opportunity to differentiate themselves from humanity. To facilitate this, the BaHN or CTHS could be modified so as to reflect one's beliefs about the self (i.e., sBaHN and sCTHS). For example, items from the sBaHN might include: *Violence is obviously a central feature of MY nature,* and, *I naturally experience compassion when I see another person suffering.* Participants would complete both the standard and s-versions of the measures. Difference scores would be computed by comparing the standard BaHN and/or CTHS and its counterpart (the sBaHN and/or sCTHS) as indications of psychological distancing.

We have several predictions. First, we predict a main effect of dispositional cynicism on distancing, such that cynics would, if given the opportunity, distance themselves from humanity more than non-cynics. Our main prediction, however is a dispositional cynicism X differentiation interaction, indicating that cynics who are given the opportunity to distance themselves from humanity would express significantly more cynicism than cynics who are not give such an opportunity, or non-cynics in either differentiation condition. They would do so because their positive self-image would no longer be tied to their judgment of humanity. In fact, we would further expect non-differentiated cynics and non-cynics to report similar levels of cynicism because cynics' positive self-image would still be tied to their judgments of humanity.

If supported, these predictions would inform our understanding of how people's human identity and pre-existing beliefs about humanity interact to influences how they react to threatening information about humanity's value. When human identity has been activated, respondents should be motivated to use their pre-existing positive self-concept to combat any highly negative information about humanity's value. Non-cynics, who presumably have strong affiliation to humanity, do so motivated to save humanity's face. Cynics, who presumably have weak affiliation to humanity, do so motivation to save their own.

Limitations

There were a number of obvious limitations to these studies. First and foremost, our primary measure of interest, the CTHS, showed dramatically lower reliability in these studies than had been previously reported (Gill and Getty, 2010). Reducing the scale from a four to as little as a two-item scale seemed to improve its reliability but not to the level that has been reported previously. Further testing of this scale is obviously needed.

Further limitation came with the design of the study. We did not counter balance the presentation of the CTHS and BaHN cynicism measures. This lack of counter balancing may have interfered with our ability to extract clean data from each measure of cynicism. This was partly due to our hope to provide a clear test using the CTHS, which yielded promising results for Gill and Getty (2010). In order to best facilitate that, we presented the CTHS directly after the primary manipulation for each study; unfortunately, this may have directly interfered with supporting our predictions.

A related puzzle is the inconsistency of correlation of the cynicism measures within each study. There seems to have been a between-studies effect on these correlations.

Only in Study 1 did we see the expected relationship between the measures. This, again, may have been due to the lack of counterbalancing of measure presentation. Future research should be sure to include counterbalancing procedures.

An even more obvious limitation was gross gender imbalance in Study 2. Unfortunately, due to a male-dominated subject pool the skewed sample was unavoidable. Viewed in light of the explanations X sex interactions in Studies 1 and 3, the overwhelmingly male participants in Study 2 surely influences our data. Follow-up studies should ensure a more balanced sample of male and female participants.

Sadly, having a small sample size may have contributed to our inability to support our predictions in Study 2 and, surely Study 3. Again, given the low number in the subject pools and a lack of expendable assets to use in recruiting paid participants, the small sample sizes were unavoidable. Thus, follow-up studies should consider recruiting from a larger pool of participants.

Future Direction

Hitherto, the primary assumption of these studies came from Gill and Getty's (2010) findings suggesting that cynicism is a paradoxical side effect of subscribing to multiple external explanations for transgressions. Although we did find some further evidence for the Gill and Getty pattern, the sum of the present studies suggests that the relationship between social explanations and cynicism is a bit more dynamic and complex. We have found evidence pointing to the potential for internal explanations having the capacity to influence cynicism in addition to external explanations. Furthermore, it might also have been revealed that dispositional cynicism could influence how people interpret the meaning of social explanations whether internal or external.

Future research will be needed to better understand how the nature of that dispositional cynicism (non-cynics vs. cynics) determines how explanations are interpreted, leading to attitude maintenance or change.

How might future research tackle the problem of finding concrete evidence of these exciting possibilities? As previously suggested, we first need a pretest to measure baseline levels of cynicism. This test would help to determine whether differences in dispositional cynicism exist and, if they do, whether or not these differences relate to changes in the interpretation of social explanations.

Conclusion

We are interested in the psychology of cynicism and social explanations. The present set of studies has provided additional information suggesting that the link between social explanation and cynicism is potentially more complex than previously thought. We found that a subtle reminder of one's own humanity can break the link between external explanations and cynicism. It may be the case that being in touch with our human identity motivates us to simply lay down the metaphorical rod of blame in an attempt to *save the face* of our human ingroup or perhaps our own, but we cannot be certain at this time. To better understand these exciting possibilities, more research is needed. We have provided some ideas on how to illuminate these relationships in future research. Until them, it seems we can at least conclude that subtly reminding people of their humanity breaks the link between external explanations and cynicism; however, the mechanisms responsible for this break are still in question.

References

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin, 126(4),* 556-574.

Andreychik, M. R. (2006). *Social explanatory styles as a foundation of social orientation.* Unpublished Master's Thesis, Lehigh University

Andreychik, M. (2009). Social explanatory style as an aspect of ordinary psychology. *An Unpublished Dissertation for Lehigh University.*

Andreychik, M. R., & Gill, M. J. (2009). Ingroup identity moderates the impact of social explanations on intergroup attitudes: External Explanations are not inherently prosocial. *Personality and Social Pychology Bulletin, 35(12),* 1632- 1646.

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63(4),* 596-612.

Batson, C. D. (1991). *The altruism question: Toward a social-psychological answer.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., Klien, T. R., Highberger, L. (1997). Empathy and attitudes: can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology, 72,* 105-118.

Blair, R. J., Jones, L., Clark, F., & Smith, M. (1997). The psychopathic individual: a lack of responsiveness to distress cues? *Psychophysiology, 34(2),* 192–198.

Castano, E., & Kofta, M. (2009). Dehumanization: Humanity and its denial. *Group Processes & Intergroup Relations, 12(6),* 695-697.

Chandler, E. V., & Jones, C. S. (1979). Cynicism—An inevitability of police work? *Journal of Police Science and Administration, 7(1),* p. 65-68.

Darby, B. W., & Schlenker, B. R. (1982). Children's reactions to apologies. Journal of Personality and Social Psychology, 43, 742–753.

Davis, J. R., & Gold, G. J. (2011). An examination of emotional empathy, attributions of stability, and the link between perceived remorse and forgiveness. *Personality and Individual Differences, 50,* 392-397.

Davis, M. (1980). A multidimensional approach to individual differences in empathy. *Catalog of Selected Documents in Psychology, 10,* 85.

Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality & Social Psychology, 60(6),* 817-830.

Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin, 15,* 543–558.

Faulkner, J. L. (2007). *Examining the relationship between social explanatory style and social orientation.* Unpublished manuscript, Lehigh University.

Fujitia, K., Trope, Y., Liberman, N., & Levin-Sagi, M. (2006). Construal levels and self-control. *Journal of Personality and Social Psychology, 90(3),* 351-367.

Gill, M. J. (2010). The beliefs about humanity nature scale. *Unpublished measure, Lehigh University.*

Gill, M. J. & Andreychik, M. R. (2007). Explanation and intergroup emotion: Social

explanations as a foundation of prejudice-related compunction. *Group Processes and Intergroup Relations [Special Issue on Intergroup Emotion], 10,* 87-106.

Gill, M. J. & Andreychik, M. A. (2009). Getting emotional about explanations: Social explanations and social explanatory styles as bases of prosocial emotions and intergroup attitudes. *Social and Personality Psychology Compass, 3(6),* 1038-1054.

Gill, M. J., & Getty, P. D. (2010). Spare the rod, damn humanity: External explanations increase compassion for transgressors while reducing favorable impressions of humanity. Unpublished research, Lehigh University.

Gill, M. J., & Andreychik, M. R. (2011). Does compassion for the outgroup depend on seeing them as lacking control?: External explanations evoke compassion by increasing perceptions of suffering, independent of perceived controllability. *A manuscript in preparation for submission for publication.*

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70,* 491–512.

Heider, F. (1958). *The psychology of interpersonal relations.* New York, NY: Wiley.

Herr, P. M. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology, 52,* 1106-1115.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology, 2,* New York: Academic Press.

Jost, J. T., & Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: Consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology, 88(3),* 498-509.

Kelley, H. H. (1973). The process of causal attribution. *American Psychologist,* 107-128.

Likel, B., Hamilton, D. L., & Sherman, S. J. (2001). Elements of a lay theory of groups: Types of groups, relational styles and perception of group entitativity. *Personality and Social Psychology Review, 5(2),* 129-140.

Luke M. A., & Maio, G. R. (2009). Oh the humanity! Humanity-esteem and its social importance. *Journal of Research in Personality, 43,* p. 586-601.

Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology, 67(6),* 949-971.

Moskowitz, G. B. (2005). *Social cognition: Understanding self and others*. New York, NY: The Guilford Press.

Moskowitz, G. B. & Gill, M. J. (in press). Interpersonal perception: From snap judgments to the regulation of enduring relationships.

Pizarro, D., Uhlmann, E., & Salovey, P. (2003). Asymmetry in judgments of moral blame and praise: The role of perceived metadesires. *Psychological Science, 14(3),* 267-272.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology. 75,* 811-832.

Rosenberg, M. (1989). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press [Original work published in 1965].

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1–14). New Haven, CT: Yale University Press.

Tajfel, H., & Turner, J.C. (1986). The social identity theory of in- tergroup conflict. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7-24). Chicago: Nelson-Hall.

Tversky, A., & Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232.

Trope, Y. & Liberman, N. (in press). Construal level theory of psychological distance. *Psychological Review.*

Vescio, T. K., Sechrist, G. B., & Paolucci, M. P. (2003). Perspective taking and prejudice reduction: the meditational role of empathy arousal and situational attributions. *European Journal of Social Psychology, 33(4),* 455-472.

Wakslak C., & Trope, Y. (2009). The effects of construal level on subjective probability estimates. *Psychological Science, 20(1),* 52-58.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92,* 548-573.

Weiner, B. (2006). *Social motivation, justice, and the moral emotions: An attributional approach*. Mahwah, New Jersey: Lawrence Erlbaum, Inc.

Zagefka, H., Noor, M., Brown, R., Randsley De Moura, G., & Hopthrow, T. (2011). Donating to disaster victims: Responses to natural and humanly caused events. *European Journal of Social Psychology, 41,* 353-363.

Zucker, G. S. & Weiner, B. (1993). Conservatism and perceptions of poverty: An attributional analysis. *Journal of Applied Social Psychology, 23(12)*, 925-943.

Footnotes

[1]No significant effects were revealed using the CTHS (all *Fs* < 2, *ps* > .2) or the BaHN_prosocial subscale (all *Fs* < 2, *ps* > .2). Using the BaHN selfish subscale, however, a marginally significant main effect of sex was revealed, $F(1, 73) = 3.39$, $p = .07$. The pattern suggested that females ($M = 3.81$, $SD = .34$) are more likely than males ($M = 3.49$, $SD = .35$) to believe humanity naturally possesses selfish qualities. Because this finding had little bearing on our hypothesis and had no effect as a covariate, we did not discuss it further.

Figure Captions

Figure 1

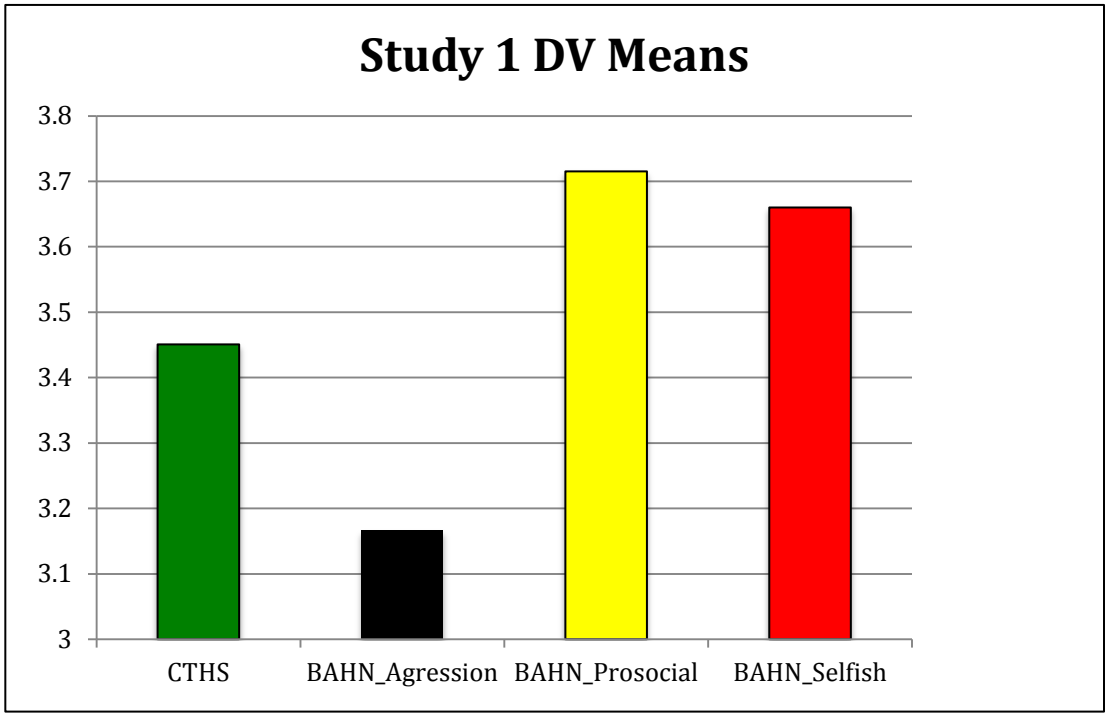


Figure 2

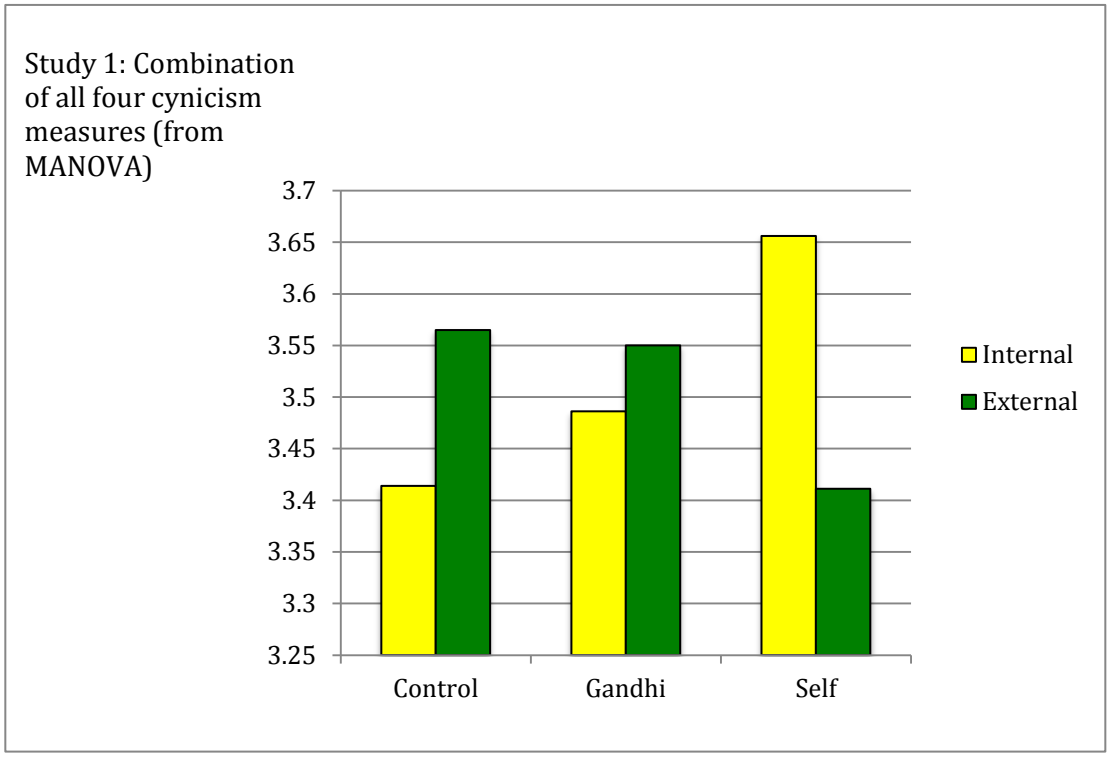


Figure 3

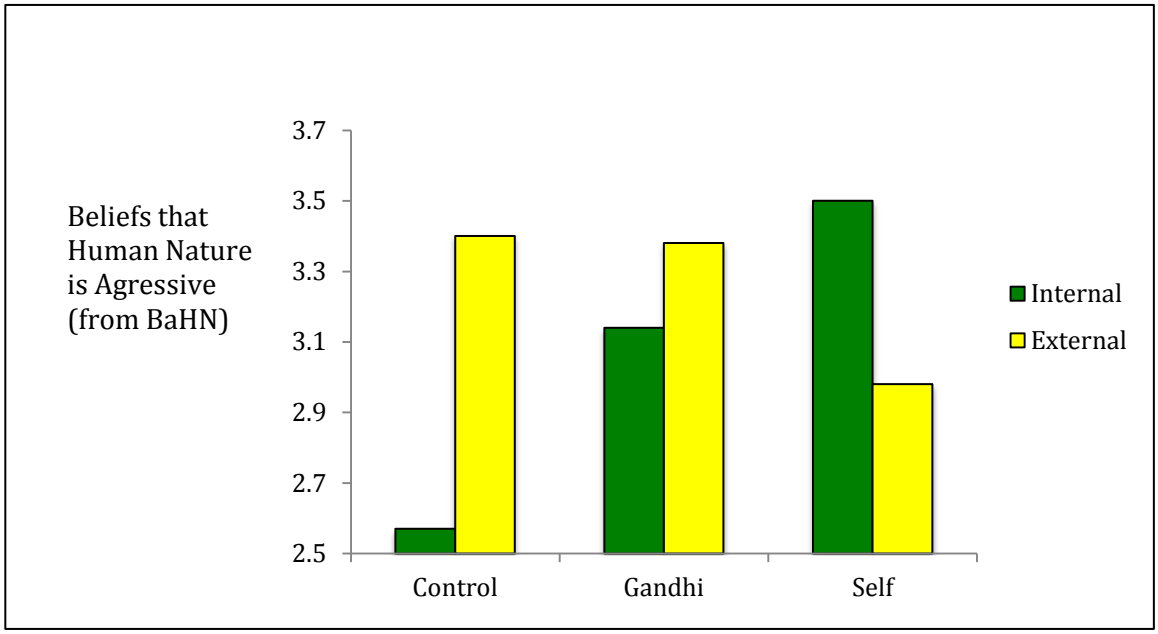


Figure 4

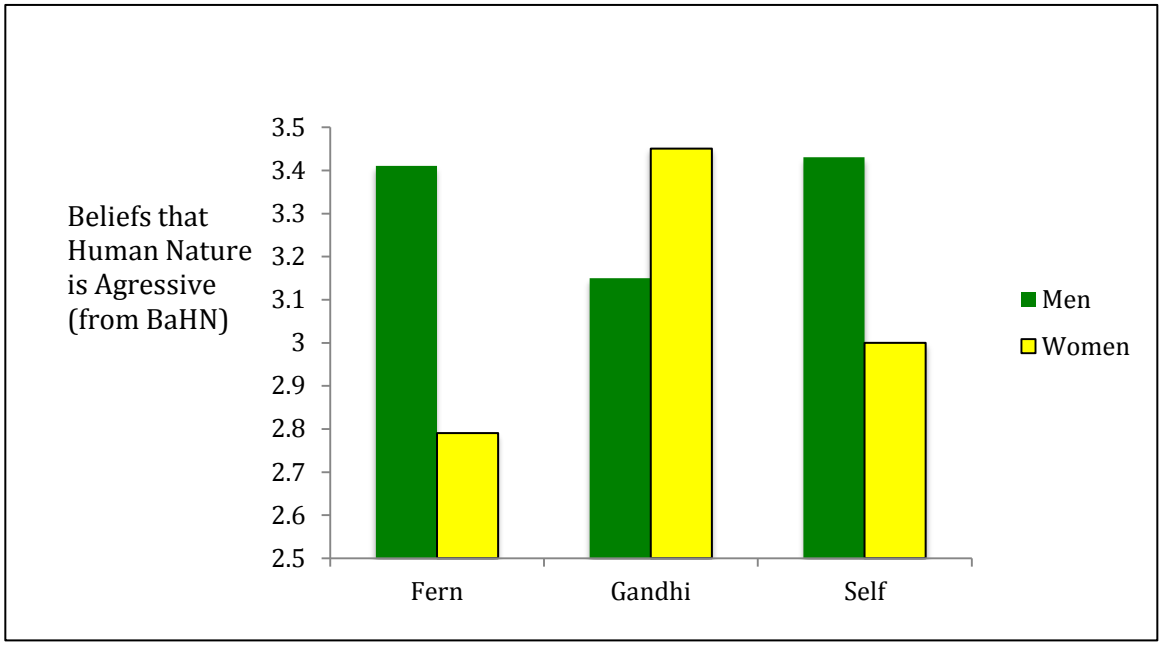


Figure 5

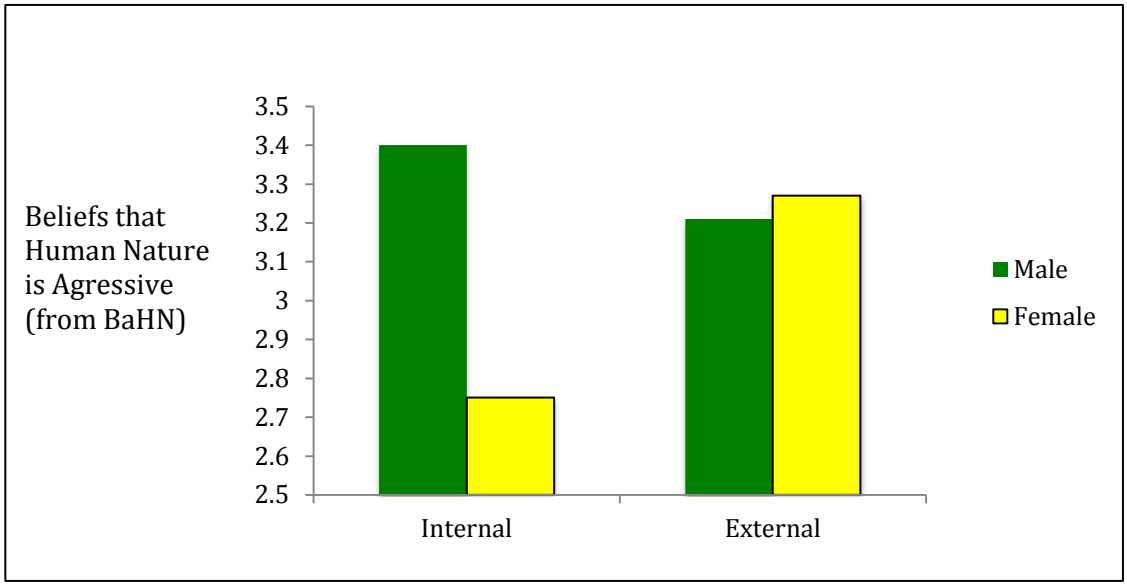


Figure 6

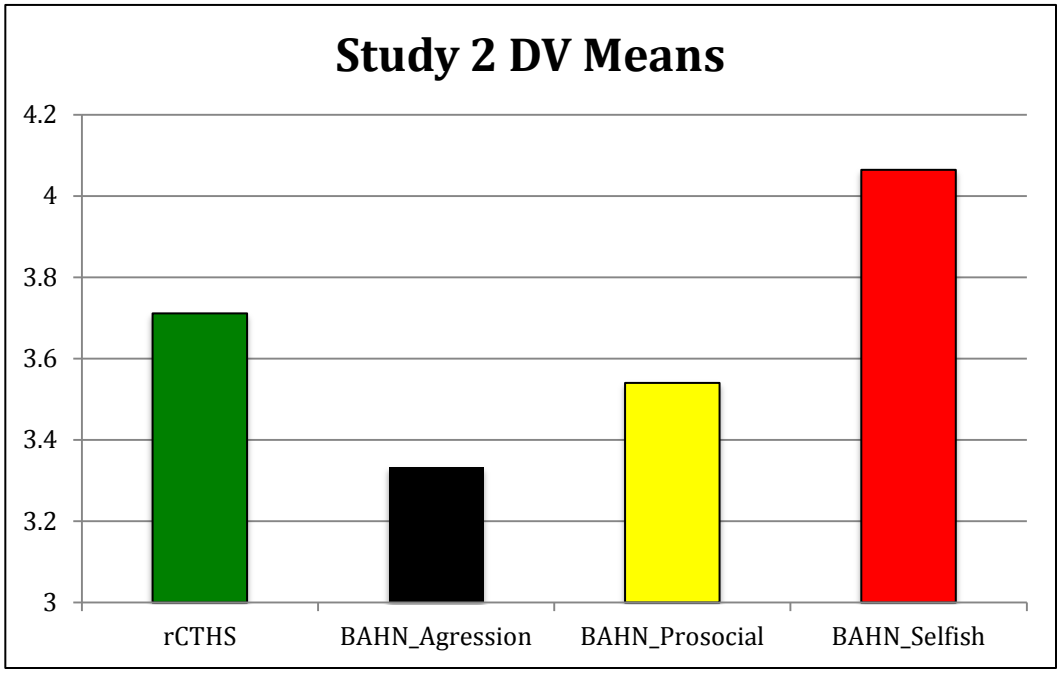


Figure 7

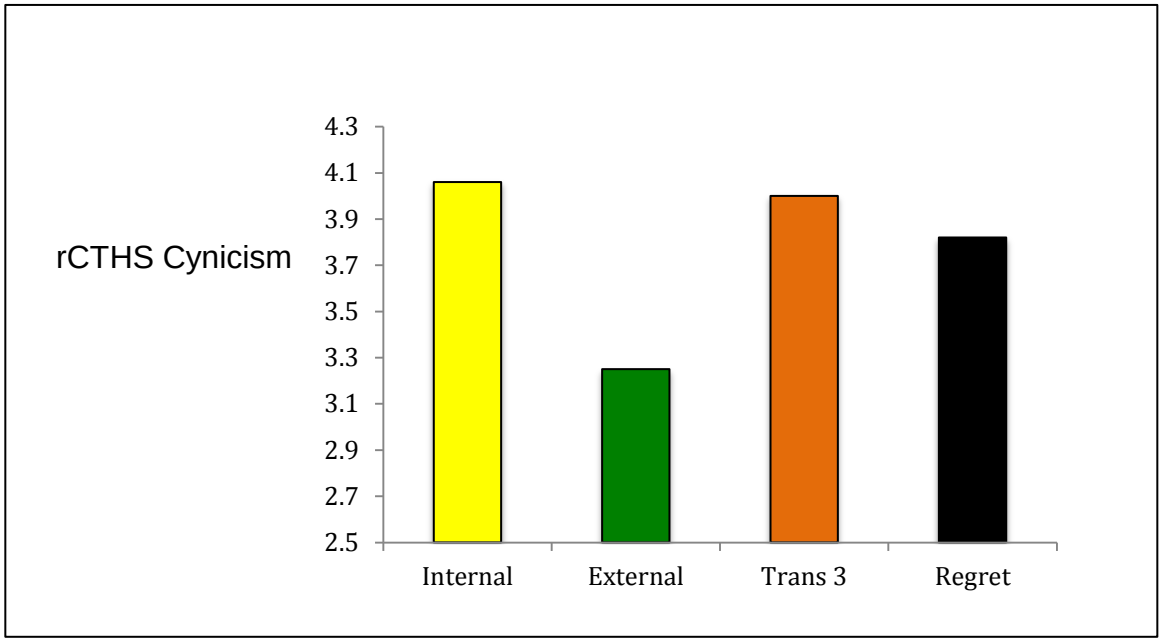


Figure 8

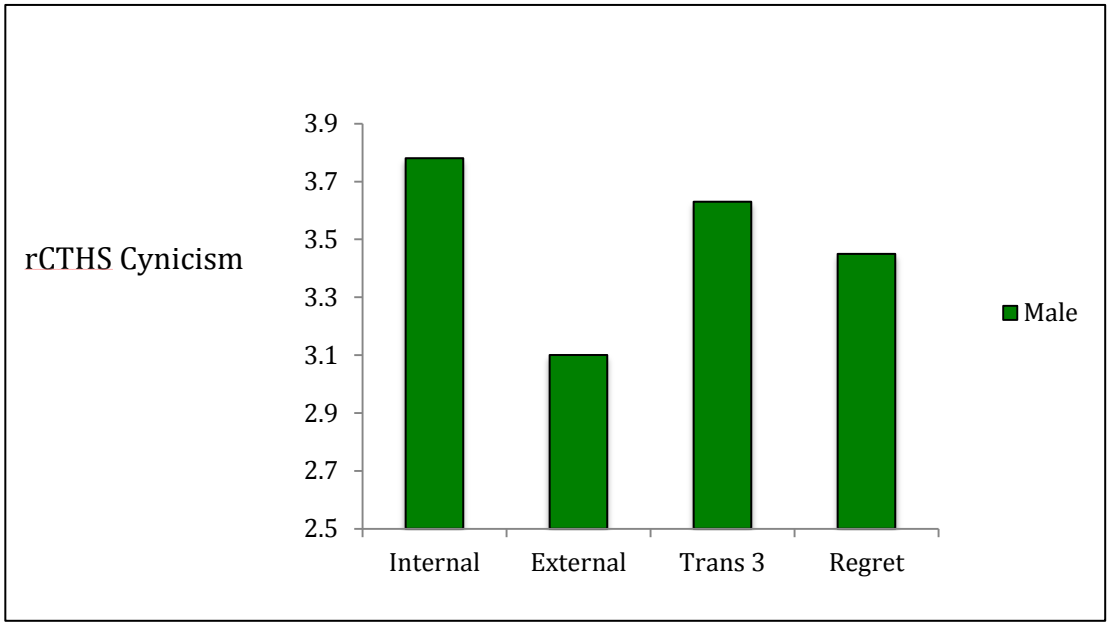


Figure 9

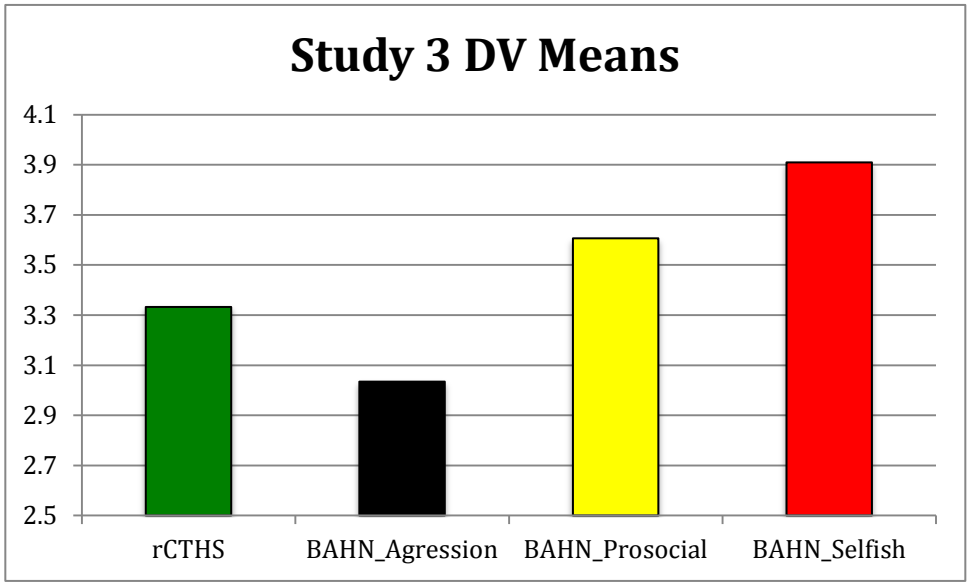
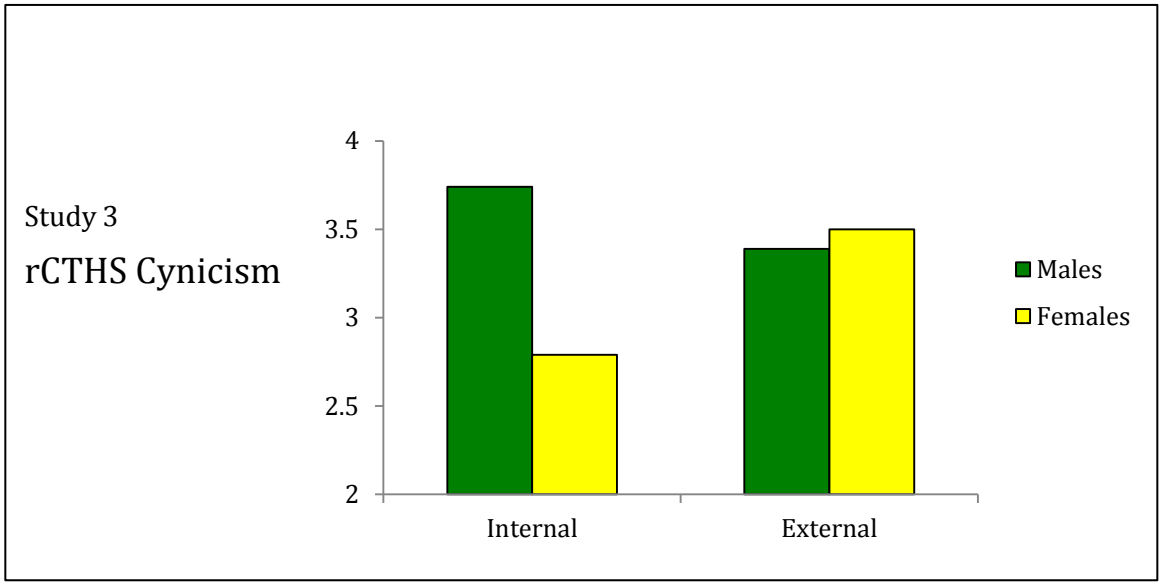


Figure 10



Appendix A: Vignettes With Explanations Manipulation

INTERNAL EXPLANATIONS CONDITION

(1) Frank has served in congress for years. Often, he has supported legislation that was inconsistent with his stated values. That is, there is a disconnect between what Frank *says* and what he *does*.
Frank loves the power and prestige of his office. Whatever he says and does as a political official is intended to fulfill his own desire for power and influence. Frank sees his constituents as pawns that must be manipulated during the campaign season in order to maintain his position of power.
(2) Megan is an accountant for a large company. Over a period of years, she "faked the books" and shifted company money into her personal bank account.
Megan wants more excitement and more material goods in her life (e.g., expensive clothing and vacations). These things motivate her to divert company money into her account.
(3) Bobby committed his first assault and robbery as a teenager. Since then, he has threatened people with a knife, punched and kicked them, and broken into their homes to steal their money and valuables.
Bobby likes to intimidate people. It makes him feel powerful and important. Bobby has no desire to help anyone else but himself, and he thinks of crime as a great way to get easy money.
(4) Pam is untruthful to lots of people. She has spent time socializing with girlfriends after work, but told her boyfriend she was working late. She has skipped work to spend time with her boyfriend, but told her supervisor she was sick.
Pam does what she wants when she wants to, even if that means being dishonest, especially with those that might try to inhibit her activities. She craves total independence. If she could, she would never answer to anyone but herself.
(5) John is very ungenerous. When friends ask for help, he does not respond. When people at work try to raise money for a sick co-worker, John offers nothing. He does not give presents for birthdays or holidays. He does not donate to charities or engage in volunteer activities in the community.
John simply has no feelings of care or concern regarding others. He considers it a waste of time or a waste of money to do anything to help others. John thinks that people should help themselves and if they are unable to do so, that's their problem and not his.
(6) At work, Laura does not follow principles of fairness in her decision making. She often gives herself the easiest tasks, and gives the hardest tasks to those she does not like. Her friends get better raises than her enemies even when their work performance is similar.
Laura considers others only as potential obstacles to her desires and success. She is intent on climbing the corporate latter even by means of sabotaging the opportunities of her coworkers and sees no reason to change. She has little regard for others; she really only cares about herself. Laura doesn't care if she hurts others to get what she wants.

EXTERNAL EXPLANATION CONDITION

(1) Frank has served in congress for years. Often, he has supported legislation that was inconsistent with his stated values. That is, there is a disconnect between what Frank *says* and what he *does*.
Indeed, leaders within Frank's political party are ruthless. Over the years, they have strong-armed Frank into voting in line with their desires—even when these contradict Frank's desires—by threatening to smear him and make him unelectable in the future. This frightens Frank because the party leaders are powerful and could likely destroy his career.
(2) Megan is an accountant for a large company. Over a period of years, she "faked the books" and shifted company money into her personal bank account.

Megan typically works 60 hours a week for a meager salary. Despite this, her boss is ungrateful and insulting. In fact, her boss is always criticizing her and taunting her about being "stuck" at her job due to constraints on her ability to move her family. Her boss gives her terrible raises because of her being "stuck." This motivates Megan to divert company money into her account.

(3) Bobby committed his first assault and robbery as a teenager. Since then, he has threatened people with a knife, punched and kicked them, and broken into their homes to steal their money and valuables.

Bobby grew up in a violent home. His father yelled at him and "cut him down to size" on a regular basis. On several occasions, Bobby's father hit Bobby in a fit of rage. Bobby's mother is afraid of her husband and therefore has done little over the years to protect Bobby or get him to a safer, healthier environment.

(4) Pam is untruthful to lots of people. She has spent time socializing with girlfriends after work, but told her boyfriend she was working late. She has skipped work to spend time with her boyfriend, but told her supervisor she was sick.

Pam had an overbearing, unkind mother. She made Pam feel ashamed and terrible any time Pam wanted to do something "for fun." Pam grew up feeling that she "didn't deserve" to have fun, to spend time relaxing, and so on. So, she developed an unconscious habit of fibbing about her need for fun and relaxation, for fear that admitting to such needs makes her a selfish person.

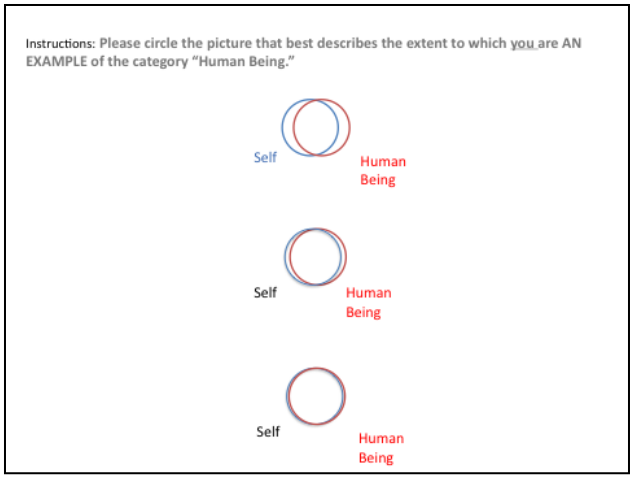
(5) John is very ungenerous. When friends ask for help, he does not respond. When people at work try to raise money for a sick co-worker, John offers nothing. He does not give presents for birthdays or holidays. He does not donate to charities or engage in volunteer activities in the community.

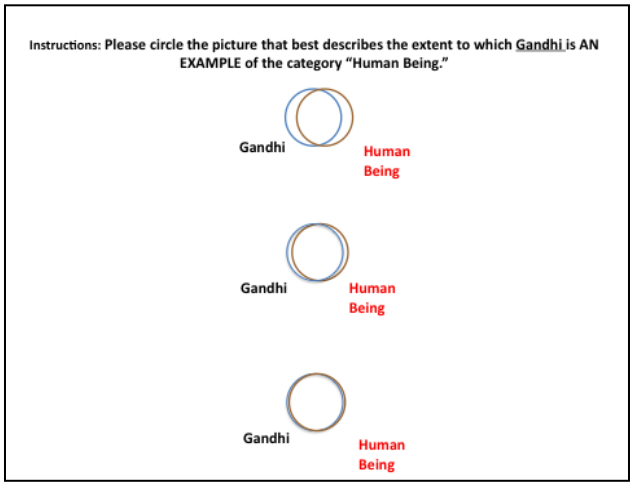
Growing up, John suffered a lot from social exclusion and disconnection. He was very unpopular, had few friends, and was beaten up several times by children who came to his school from rough neighborhoods. These experiences hurt him a lot and gave him a negative image of other people, inhibiting any tendency to be generous toward others.

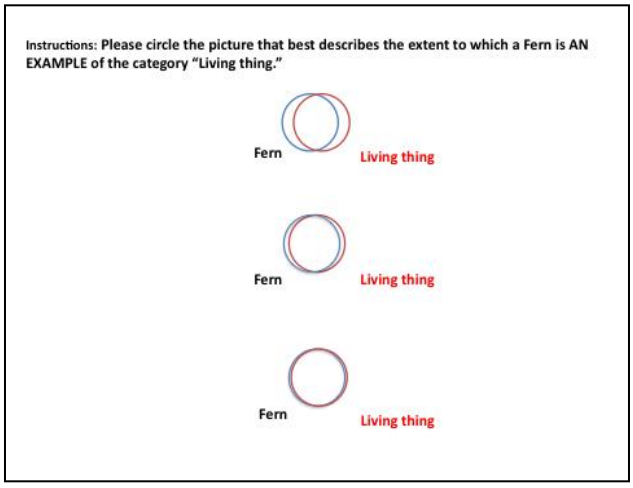
(6) At work, Laura does not follow principles of fairness in her decision making. She often gives herself the easiest tasks, and gives the hardest tasks to those she does not like. Her friends get better raises than her enemies even when their work performance is similar.

Laura has a history of working for deceitful and harassing corporate executives. Laura is regularly the subject of her boss's smarmy comments and abuse. In addition to inappropriate behavior, Laura's boss continues to overlooks her for promotion even though she displays superior performance and competence. She is not appreciated for the long hours she works and is rarely given positive feedback or praise for her contribution to the company's success.

Appendix B: Modified IOS Scales







Appendix C: Cynicism Towards Humanity Scale

Please read each statement carefully and CIRCLE the number that best characterizes your feelings.

1. The way people act can sometimes make me feel cynical and pessimistic about humanity.

1----------2----------3----------4----------5

NO YES

2. Sometimes I cannot help but think: "People are jerks."

1----------2----------3----------4----------5

NO YES

3. Many aspects of human behavior are *NOT* admirable at all.

1----------2----------3----------4----------5

NO YES

4. I think that, for the most part, people are good and admirable.

1----------2----------3----------4----------5

NO YES

Appendix D: Beliefs about Human Nature Scale

Below are some items describing ideas about Human Nature, or the traits that are "built in" to the human species. Please indicate the extent to which you personally agree with each item. Please use the following scale:

1----------2----------3----------4----------5

Strongly DISAGREE **Strongly AGREE**

____ (1) Violence is obviously a central feature of human nature.
____ (2) Human beings are naturally aggressive.
____ (3) Human beings are obviously *not* peaceful by nature.
____ (4) Given human nature, there will always be wars.
____ (5) Human beings naturally experience compassion when they see another person suffering.
____ (6) The tendency to care for others is built in to human nature.
____ (7) When people help a person in need, the help is motivated by true feelings of compassion or caring.
____ (8) It is human nature to have a sense of concern about the well-being of others.
____ (9) It is human nature to be primarily concerned with "looking out for #1" (i.e., oneself).
____ (10) It is natural for a person to put him- or herself first, and others second.
____ (11) When things get tough in a community, people will naturally focus on their own well-being and show less concern for others.
____ (12) As a species, one of our fundamental motives is to take care of ourselves first and foremost, to be somewhat "selfish."

Appendix E: External Explanations Vignettes with Regret

EXTERNAL-EXPLANATION-PLUS-REGRET CONDITION

(1) Frank has served in congress for years. Often, he has supported legislation that was inconsistent with his stated values. That is, there is a disconnect between what Frank *says* and what he *does*.

Indeed, leaders within Frank's political party are ruthless. Over the years, they have strong-armed Frank into voting in line with their desires—even when these contradict Frank's desires—by threatening to smear him and make him unelectable in the future. This frightens Frank because the party leaders are powerful and could likely destroy his career.

(2) Megan is an accountant for a large company. Over a period of years, she "faked the books" and shifted company money into her personal bank account.

Megan typically works 60 hours a week for a meager salary. Despite this, her boss is ungrateful and insulting. In fact, her boss is always criticizing her and taunting her about being "stuck" at her job due to constraints on her ability to move her family. Her boss gives her terrible raises because of her being "stuck." This motivates Megan to divert company money into her account.

Lately, Megan's boss has been reflecting on his behavior. He has decided he has been really awful and unfair. He has apologized to Megan and is now sincerely committed to improving his relationship with her.

(3) Bobby committed his first assault and robbery as a teenager. Since then, he has threatened people with a knife, punched and kicked them, and broken into their homes to steal their money and valuables.

Bobby grew up in a violent home. His father yelled at him and "cut him down to size" on a regular basis. On several occasions, Bobby's father hit Bobby in a fit of rage. Bobby's mother is afraid of her husband and therefore has done little over the years to protect Bobby or get him to a safer, healthier environment.

As Bobby's dad has matured, it has dawned on him that his "parenting style" is ruining his son. Because of this, he has enrolled in parenting classes and is perpetually improving his "anger management" techniques. He is sincerely regretful about his "former self."

(4) Pam is untruthful to lots of people. She has spent time socializing with girlfriends after work, but told her boyfriend she was working late. She has skipped work to spend time with her boyfriend, but told her supervisor she was sick.

Pam had an overbearing, unkind mother. She made Pam feel ashamed and terrible any time Pam wanted to do something "for fun." Pam grew up feeling that she "didn't deserve" to have fun, to spend time relaxing, and so on. So, she developed an unconscious habit of fibbing about her need for fun and relaxation, for fear that admitting to such needs makes her a selfish person.

(5) John is very ungenerous. When friends ask for help, he does not respond. When people at work try to raise money for a sick co-worker, John offers nothing. He does not give presents for birthdays or holidays. He does not donate to charities or engage in volunteer activities in the community.

Growing up, John suffered a lot from social exclusion and disconnection. He was very unpopular, had few friends, and was beaten up several times by children who came to his school from rough

neighborhoods. These experiences hurt him a lot and gave him a negative image of other people, inhibiting any tendency to be generous toward others.

At a recent 10-year class reunion, John was approached by at least six of the former bullies. Each bully had grown up a lot and described how guilty he or she felt about his or her cruelty to John. They all hoped that John was doing well and that he would forgive them.

(6) At work, Laura does not follow principles of fairness in her decision-making. She often gives herself the easiest tasks, and gives the hardest tasks to those she does not like. Her friends get better raises than her enemies even when their work performance is similar.

Laura has a history of working for deceitful and harassing corporate executives. Laura is regularly the subject of her boss's smarmy comments and abuse. In addition to inappropriate behavior, Laura's boss continues to overlooks her for promotion even though she displays superior performance and competence. She is not appreciated for the long hours she works and is rarely given positive feedback or praise for her contribution to the company's success.

Appendix F: External Explanations Vignettes with Transgressors 3

EXTERNAL-EXPLANATION-PLUS-TRANSGRESSOR 3 CONDITION

(1) Frank has served in congress for years. Often, he has supported legislation that was inconsistent with his stated values. That is, there is a disconnect between what Frank *says* and what he *does*.

Indeed, leaders within Frank's political party are ruthless. Over the years, they have strong-armed Frank into voting in line with their desires—even when these contradict Frank's desires—by threatening to smear him and make him unelectable in the future. This frightens Frank because the party leaders are powerful and could likely destroy his career.

At the same time, the leaders who oppress Frank are themselves exploited. The leaders of Frank's party are practically enslaved by the special interests of corporations who contribute millions of dollars to their campaign funds. If they fail to "produce results" that please these corporate donors, the party leaders will soon lack any power whatsoever because their bank accounts will run dry.

(2) Megan is an accountant for a large company. Over a period of years, she "faked the books" and shifted company money into her personal bank account.

Megan typically works 60 hours a week for a meager salary. Despite this, her boss, Bertram, is ungrateful and insulting. In fact, Bertram is always criticizing her and taunting her about being "stuck" at her job due to constraints on her ability to move her family. So Bertram gives Megan terrible raises and over works her, as she sees it, because of her being "stuck." This motivates Megan to divert company money into her account.

Although Bertram behaves like a tyrant, he is also a victim. Bertram's boss, Clive, the president of their company, routinely bullies Bertram into cutting expenses anyway he can. Bertram's behavior towards Megan is largely a product of Bertram's fear of what Clive will do to him if he does not cut costs by keeping Megan's salary very low.

(3) Bobby committed his first assault and robbery as a teenager. Since then, he has threatened people with a knife, punched and kicked them, and broken into their homes to steal their money and valuables.

Bobby grew up in a violent home. His father yelled at him and "cut him down to size" on a regular basis. On several occasions, Bobby's father hit Bobby in a fit of rage. Bobby's mother is afraid of her husband and therefore has done little over the years to protect Bobby or get him to a safer, healthier environment.

Bobby and Bobby's father are part of a long line of troubled individuals. The abuse Bobby has suffered at the hands of his father is simply a perpetuation of the same abuse Bobby's father suffered at the hands of Bobby's grandfather. Bobby's father learned from his father that physical violence is an appropriate way to "teach a lesson" to children.

(4) Pam is untruthful to lots of people. She has spent time socializing with girlfriends after work, but told her boyfriend she was working late. She has skipped work to spend time with her boyfriend, but told her supervisor she was sick.

Pam had an overbearing, unkind mother, Margaret. Margaret made Pam feel ashamed and terrible any time Pam wanted to do something "for fun." Pam grew up feeling that she "didn't deserve" to have fun, to spend time relaxing, and so on. So, she developed an unconscious habit of fibbing about her need for fun and relaxation, for fear that admitting to such needs makes her a selfish person.

Margaret's behavior, in turn, was caused by her experiences growing up. Her family was desperately poor and turned to fundamentalist religion to give them hope. Unfortunately, their religion required coldness and harshness toward one's children, supposedly to "get the sinful nature out of them." Because of this treatment, Margaret suffered a lot. Now, Margaret is inflicting the same coldness and criticism on her own daughter that made her suffer in the past.

(5) John is very ungenerous. When friends ask for help, he does not respond. When people at work try to raise money for a sick co-worker, John offers nothing. He does not give presents for birthdays or holidays. He does not donate to charities or engage in volunteer activities in the community.

Growing up, John suffered a lot from social exclusion and disconnection. He was very unpopular, had few friends, and was beaten up several times by children who came to his school from rough neighborhoods, inhibiting any tendency to be generous toward others.

Of course, the kids that ostracized and bullied John had their own problems. They came from homes that provided little guidance in the norms of positive social relations. Thus, these kids learned to rely on violence to gain "status" and a reputation for being "tough." Without such a reputation, they feared that they would be targets of bullying.

(6) At work, Laura does not follow principles of fairness in her decision-making. She often gives herself the easiest tasks, and gives the hardest tasks to those she does not like. Her friends get better raises than her enemies even when their work performance is similar.

Laura has a history of working for deceitful and harassing corporate executives. Laura is regularly the subject of her boss Richard's smarmy comments and abuse. In addition to inappropriate behavior, Richard continues to overlook Laura for promotion even though she displays superior performance and competence. Laura is not appreciated for the long hours she works and is rarely given positive feedback or praise for her contribution to the company's success.

Naturally, Richard is not the sole cause of Laura's difficulties. Based on anxiety about his own ability to succeed, Richard has been sucked into the "good ol' boy mentality" of his corporate leadership. Within this mentality, one must favor friends (e.g., those with whom one "hits the golf course") and devalue those who are not part of the "network." Richard conforms to the "good ol' boy" mentality primarily because he fears the negative consequences of not doing so: Receiving the same type of negative treatment he inflicts on Laura.

Appendix G: Construal Priming Manipulation

Low-level Construal items

We will now switch gears, and ask you to complete a different type of task having to do with thinking about categories and examples. In each question below, you will be provided with a category and will be asked to provide an example of something that belongs to it.
For instance: "An example of a SKYSCRAPER is *the Empire State Building"*

1. An example of an ACTOR is … _____
2. An example of a BEER is … _____
3. An example of a BOOK is … _____
4. An example of a CANDY is … _____
5. An example of a COIN is … _____
6. An example of a COLLEGE is … _____
7. An example of a COMPUTER is … _____
8. An example of a DANCE is … _____
9. An example of a GAME is … _____
10. An example of a KING is … _____
11. An example of LUNCH is … _____
12. An example of MAIL is … _____
13. An example of MATH is … _____
14. An example of a MOUNTAIN is … _____
15. An example of a MOVIE is … _____
16. An example of a NEWSPAPER is … _____
17. An example of a PAINTING is … _____
18. An example of PASTA is … _____
19. An example of a PHONE is … _____
20. An example of a PROFESSOR is … _____
21. An example of a RESTAURANT is … _____
22. An example of a RIVER is … _____
23. An example of a SENATOR is … _____
24. An example of a SHOE is … _____
25. An example of a SINGER is … _____
26. An example of a SOAP OPERA is … _____
27. An example of a SODA is … _____
28. An example of a SPORT is … _____
29. An example of a TREE is … _____
30. An example of a WHALE is … _____

High-level Construal

We will now switch gears, and ask you to complete a different type of task having to do with thinking about categories and examples. In each question below, you will be provided with an example and will be asked to identify a category that it that belongs to.
For instance: "A SKYSCRAPER is an example of *a building"*

1. An ACTOR is an example of ... _____
2. A BEER is an example of ... _____
3. A BOOK is an example of ... _____
4. A CANDY is an example of ... _____
5. A COIN is an example of ... _____
6. A COLLEGE is an example of ... _____
7. A COMPUTER is an example of ... _____
8. A DANCE is an example of ... _____
9. A GAME is an example of ... _____
10. A KING is an example of ... _____

11. LUNCH is an example of … _____
12. MAIL is an example of … _____
13. MATH is an example of … _____
14. A MOUNTAIN is an example of … _____
15. A MOVIE is an example of … _____
16. A NEWSPAPER is an example of … _____
17. A PAINTING is an example of … _____
18. PASTA is an example of … _____
19. A PHONE is an example of … _____
20. A PROFESSOR is an example of … _____
21. A RESTAURANT is an example of … _____
22. A RIVER is an example of … _____
23. A SENATOR is an example of … _____
24. A SHOE is an example of … _____
25. A SINGER is an example of … _____
26. A SOAP OPERA is an example of … _____
27. A SODA is an example of … _____
28. A SPORT is an example of … _____
29. A TREE is an example of … _____
30. A WHALE is an example of … _____

Phillip D. Getty
Department of Psychology
17 Memorial Drive East
Bethlehem, PA 18015-3068
Ph#: 610-597-3579
Email: pdg209@lehigh.edu

EDUCATION

2011—M.S., Psychology, Lehigh University (Expected).
Thesis: Dropping the metaphorical rod of blame to save humanity: Testing multiple means of breaking the link between external explanations for transgressions and increased cynicism toward humanity

2009—B.A., Psychology, University of Oregon, Summa Cum Laude, with department honors.
Honors Thesis: "Autistic-like" personality traits in neurotypical adults: Examining possible differences in the ease, speed, and hierarchy of social inference.

1999—Dive Master, Lane Community College.

PUBLICATIONS & PRESENTATIONS

Getty, P. D. (2009). "Autistic-like" personality traits in neurotypical adults: Examining possible differences in the ease, speed, and hierarchy of social inference. *McNair Scholars Research Journal, University of Oregon, (1)*, p. 61-117.

Getty, P. D. (November, 2008). "Autistic-like" personality traits in neurotypical adults: Examining possible differences in the ease, speed, and hierarchy of social inference. Poster presented at the 17th annual Ronald E. McNair Scholars Research Conference, Delavan, WI.

PUBLICATIONS (NON-ACADEMIC)

Getty, Phillip D. (2007), Web site publication. *Capitellowines.com.*

Getty, Phillip D. (May 24, 2007). LCC goes global: President Mary Spilde plans internationally. *Eugene Weekly, p. 13.*

Getty, Phillip D. (April 26, 2007). Chicken tacos and drunken nachos. *Eugene Weekly (Chow addition)*, p. 7.

MANUSCRIPTS IN PROGRESS

Gill, M. J. & Getty, P. D. (*in prep.*). All this compassion is making me cynical!: External

explanations increase cynicism toward humanity even as they elicit compassion for the individual transgressor.

RELEVANT WORK EXPERIENCE:

2009-Present—Graduate Research Assistant for Dr. Mike Gill

2010-Present—News Editor for Wiley-Blackwell Publishing at socialpsychologicaleye.com

Present—Graduate Teaching Assistant (Statistics) with Dr. Hupbach

Spring 2011—Graduate Teaching Assistant (Social Psychology) with Dr. Gordon Moskowitz

Fall 2010—Graduate Teaching Assistant (Developmental Psychology) with Dr. Susan Barrett

2008-2009—Research Assistant for Dr. Gerard Saucier

2007-2009—Research Assistant for Dr. Sara Hodges

2007-2009—Research Assistant for Dr. Bertram F. Malle

HONORS AND AWARDS

2011—Lehigh University, College of Arts and Sciences, Graduate Summer Research Fellowship

2009—Lehigh University, McNair Scholars Graduate Fellowship

2008—Ronald E. McNair Scholarship (McNair Scholars Program). Thesis: "Autistic-like" personality traits in neurotypical adults: Examining possible differences in the ease, speed, and hierarchy of social inference.

2008—Non-Traditional Student Award, University of Oregon.

2007—Honorable Mention, League for Innovation's Annual Essay Contest for Community Colleges. Essay: The Legacy of Constantine.