

Emotion

Are Self-Conscious Emotions About the Self? Testing Competing Theories of Shame and Guilt Across Two Disparate Cultures

Mitchell Landers, Daniel Sznycer, and Patrick Durkee

Online First Publication, January 22, 2024. <https://dx.doi.org/10.1037/emo0001321>

CITATION

Landers, M., Sznycer, D., & Durkee, P. (2024, January 22). Are Self-Conscious Emotions About the Self? Testing Competing Theories of Shame and Guilt Across Two Disparate Cultures. *Emotion*. Advance online publication. <https://dx.doi.org/10.1037/emo0001321>

Are Self-Conscious Emotions About the Self? Testing Competing Theories of Shame and Guilt Across Two Disparate Cultures

Mitchell Landers¹, Daniel Sznycer², and Patrick Durkee³

¹Department of Psychology, University of Chicago

²Oklahoma Center for Evolutionary Analysis, Department of Psychology, Oklahoma State University

³Department of Psychology, California State University, Fresno

The emotions of guilt and shame play major roles in forgiveness, social exclusion, face-saving ploys, suicide, and honor killings. Understanding these emotions is thus of vital importance. The outputs of guilt and shame are already well understood: Guilt motivates amends; shame motivates evasion. However, the elicitors and functions of these emotions are disputed. According to attributional theory, guilt and shame are intrapersonal emotions elicited when negative outcomes are attributed to controllable/unstable (guilt) or uncontrollable/stable (shame) aspects of the self. By contrast, functionalist theory claims that guilt and shame are interpersonal emotions for minimizing the imposition of harm on valued others (guilt) and the cost of reputational damage on the self (shame). Although there is confirmatory evidence consistent with both theories, evidence ostensibly supporting one theory has been argued to actually support the other. To solve this problem of data interpretation, here we report contrastive critical tests of the two theories performed on online participant pools in the United States and India in 2021 ($N = 853$). Results in both countries support functionalist theory over attributional theory, suggesting that the intrapersonal effects reported in the emotion literature are tributary or incidental to the interpersonal functions of guilt and shame. Functionalist theory presents a promising framework for understanding the interpersonal and intrapersonal aspects of guilt, shame, and other self-conscious emotions.

Keywords: shame, guilt, attribution, intrapersonal processes, interpersonal processes

The importance of guilt and shame in human affairs is difficult to overstate. These self-conscious emotions play major roles in forgiveness, social exclusion, face-saving ploys, suicide, and honor killings (McCullough, 2008; Peristiany, 1970; Syme & Hagen, 2019; Tangney & Dearing, 2002). Guilt and shame sometimes are mobilized concurrently (Tangney, Wagner, & Gramzow, 1992). However, a

growing body of evidence shows stark differences between these two emotions (Smith et al., 2002; Tangney, Wagner, Fletcher, & Gramzow, 1992; Zhu, Feng, et al., 2019).

Guilt summons the good after the bad. When feeling guilt, people interrupt actions that would impose harm on others and try to make things right: People approach those whom they have harmed and confess their wrongdoing, accept responsibility, apologize, and make amends (Baumeister et al., 1994; de Hooge et al., 2007; Tangney, 1991). Shame, by contrast, summons the bad after the bad. Shame, like guilt, can motivate people to interrupt their selfish or reckless behavior (de Hooge et al., 2008). But shame's link to cooperative behavior is far weaker than guilt's (Declerck et al., 2014). Moreover, shame is associated with noncooperative and spiteful behavior, including evasion, concealment, externalization of blame, and aggression (Elison et al., 2014; Tangney, Wagner, Fletcher, & Gramzow, 1992; Zhu, Xu, et al., 2019).

What elicits guilt and shame? Why do these emotions mobilize the outputs that they do? Is shame a standard part of human nature or a pathology we are better off without? Although there is consensus about the outputs of guilt and shame, answers to these questions remain a matter of contention. Here, we aim to identify the elicitors and functions of guilt and shame by testing between central predictions derived from two major, competing theories: attributional theory and functionalist theory.

Attributional theory holds that guilt and shame are intrapersonal emotions elicited through self-reflective and self-evaluative processes (Tangney, 1990; Tracy & Robins, 2004; Weiner, 1985). On this account, both guilt and shame are elicited when individuals

Mitchell Landers  <https://orcid.org/0000-0003-4393-2871>

Mitchell Landers and Daniel Sznycer share first authorship. The authors would like to thank an anonymous reviewer for helpful comments on a previous version of this article. The Open Science Framework repository is available at https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3.

Mitchell Landers served as lead for writing—original draft and writing—review and editing, contributed equally to conceptualization, methodology, and validation, and served in a supporting role for data curation, formal analysis, and visualization. Daniel Sznycer served as lead for conceptualization and project administration and contributed equally to investigation, supervision, writing—original draft, and writing—review and editing. Patrick Durkee served as lead for data curation, formal analysis, methodology, and visualization and served in a supporting role for writing—review and editing.

Correspondence concerning this article should be addressed to Mitchell Landers, Department of Psychology, University of Chicago, 5848 S. University Avenue, Chicago, IL 60637-1554, United States, or Daniel Sznycer, Oklahoma Center for Evolutionary Analysis, Department of Psychology, Oklahoma State University, Psychology Building, Stillwater, OK 74078-3064, United States. Email: msblanders@uchicago.edu or daniel.sznycer@okstate.edu

attribute events incongruent with their identity goals to causes within themselves. Importantly, however, guilt is triggered when identity-incongruent outcomes are attributed to unstable or controllable aspects of the self, whereas shame is triggered when identity-incongruent outcomes are attributed to stable or uncontrollable aspects of the self. For example, if a good neighbor is the kind of person you aspire to be (your identity goal), you would feel guilt for damaging your neighbor's car if you attributed the damage to your lapse in attention (attribution: unstable and controllable), but you would instead feel shame if you attributed the damage to your brain-injury-induced tremors (attribution: stable and uncontrollable; Tracy & Robins, 2004, 2006). The affective, motivational, and behavioral characteristics of guilt and shame follow from the different patterns of attributions that elicit these emotions. Guilt feels bad, but the blame placed on the self is focalized and transient, and this leaves a guilty individual able to approach others and take reparative action. When feeling shame, by contrast, the individual blames not a specific action or inaction but their "global self" (Tangney, 1991). This belief in one's overarching defectiveness leads to depression, paranoid thoughts, and defensive maneuvers to deaden one's pain: hiding, blaming others, aggression, and more (Tangney, Wagner, Fletcher, & Gramzow, 1992). Attributional theory thus views shame as a maladaptive and pathogenic emotion (Tangney, Wagner, et al., 1996)—an inferior substitute for guilt. The attributional view that guilt and shame are fundamentally about the self is highly influential—so much so that in the scientific literature and beyond these emotions are generally referred to as "self-conscious emotions."

By contrast, functionalist theory argues that shame and guilt serve distinct interpersonal functions and thus are primarily about others (Baumeister et al., 1994; Gilbert, 1997; Leary et al., 2001; Sznycer, 2019; Tooby & Cosmides, 2008). Shame functions to minimize the spread of reputation-damaging information and the likelihood and costs of being devalued by others (Gilbert, 1997; Landers & Sznycer, 2022; Sznycer et al., 2016). And guilt functions to correct one's behavior toward valued others one has harmed. When harming valued others the individual indirectly harms herself, and guilt functions to correct this error going forward (Baumeister et al., 1994, 1995). On this account, shame is elicited more intensely when the individual, for example, fears others' devaluative reactions (e.g., aggression), while guilt is elicited more intensely when the individual imposes undue harm on valued others. For instance, if you unintentionally brought harm to your coworker, you would feel more shame (e.g., more intense motivations to evade) if that coworker were your tyrannical boss, whom you fear, than if that coworker were your likeable intern, whom you do not. This is because if your coworker found out what you had done and subsequently devalued you, you would incur more costs from your tyrannical boss than from your likeable intern. In contrast, if you unintentionally brought harm to your coworker, you would feel more guilt (e.g., more intense motivations to amend) if that coworker were your likeable intern than if that coworker were your tyrannical boss—because, indirectly, you would suffer higher costs for harming your intern, whose welfare you value more highly than your boss's. According to functionalist theory, guilt and shame sometimes activate concurrently because some actions or inactions simultaneously harm valued others (guilt) and carry the threat of being devalued by others (shame). But the functions, elicitors, computations, and outputs of these emotions differ. For example, someone

who feels both shame and guilt about an act of infidelity might refrain from it, whereas someone who feels only shame might continue to commit infidelity but conceal it. From a functionalist perspective, the characteristics of guilt and shame reflect the real adaptive problems these emotion systems evolved to solve, not different patterns of attributions generated by the individual. Guilt's outputs are clearly cooperative because guilt functions to curb the imposition of undue harm on valued others. Shame's outputs are sometimes cooperative because cooperation is sometimes a cost-effective way to enhance one's reputation (Barclay & Willer, 2007). But enhanced cooperation is not the sole means of minimizing the threat of being devalued, and avoiding, hiding, threatening, and aggressing, although socially undesirable, can counteract devaluative threats in some contexts (Leach & Cidam, 2015).

Attributional theory and functionalist theory have each prompted confirmatory tests leading to supporting evidence. For instance, in support of attributional theory, Niedenthal et al. (1994) found that when generating counterfactual thoughts to "undo" their feelings of shame and guilt, people undo shame by changing qualities of themselves but undo guilt by changing their actions. And in support of functionalist theory, Smith et al. (2002) found that shame, more than guilt, is sensitive to how much public scrutiny one's failures receive.

Despite this and other evidence consistent with attributional theory (e.g., Feiring et al., 1998; Giner-Sorolla & Espinosa, 2011; Tangney, 1991; Tangney, Miller, et al., 1996; Tangney, Wagner, Fletcher, & Gramzow, 1992; Tangney, Wagner, & Gramzow, 1992; Tracy & Robins, 2006) and functionalist theory (e.g., de Hooge et al., 2010; Dickerson et al., 2004; Durkee et al., 2019; Gilbert, 2000; Giner-Sorolla & Espinosa, 2011; Martens et al., 2012; Robertson et al., 2018; Sznycer & Cohen, 2021; Yu et al., 2014; Zhu, Feng, et al., 2019; Zhu, Xu, et al., 2019), no research has yet jointly tested competing predictions derived from each theory, thereby yielding data capable of weakening either theory (or both). This is unfortunate, because evidence ostensibly supporting one theory may in fact support the other. For example, emotion outputs generated by causal processes specified by theory A may give the illusion of supporting theory B if theory B capitalizes on concomitant associations observed in tests conceived to support theory B. Indeed, attributional researchers have argued that previous research on guilt and shame suffers from this problem. For instance, they have claimed that data apparently in support of functionalist theory in fact support attributional theory: "Taken together, [R. H.] Smith et al. [(2002)]'s findings [NB: the ones referred to above] are consistent with the notion that people focus on others' evaluations because they are feeling shame, *not vice versa*" (Tangney et al., 2007 p. 349, emphasis added). That is, according to these attributional researchers, associations between shame and a focus on others' evaluations run causally from shame to evaluative concerns rather than—as functionalist researchers claim to have demonstrated—from evaluative concerns to shame. In our view, while both the attributional and functionalist interpretations are plausible, they must be regarded as conjectural because the contrastive critical tests necessary to properly adjudicate between them have not been conducted. Thus, the relative merits of these two theories remain to be determined.

Analyzing guilt and shame with tests enabling strong inference (Platt, 1964) would improve our understanding of human nature. It also would improve interventions to manage these emotions and their behavioral correlates—anti-social behavior, drug-abuse, stigmatization,

suicidality, and others (Dearing et al., 2005; Goffnett et al., 2020; Ollivier et al., 2022)—because theories of these emotions influence what interventions aim to achieve.

To close this gap, we conducted preregistered critical tests in the United States and India, two cultures varying substantially along the individualism–collectivism spectrum (Muthukrishna et al., 2020)—a variable that is thought to cause the differential expression of guilt and shame (Benedict, 1946; Fessler, 2004). Using structural equation modeling, we tested the degree to which input variables hypothesized by each theory—attributions of controllability and stability, by attributional theory; liking and fear of a target individual, by functionalist theory—elicit distinctive outputs of the guilt and shame systems (motivations to amend and evade, respectively). These tests allow us to assess the validity of attributional theory and functionalist theory, in both absolute terms and contrastive terms, in India and the United States.

Method

Transparency and Openness

Procedure, stimuli, sample sizes, exclusion criteria, predictions, and analyses were preregistered before data collection began: <https://aspredicted.org/blind.php?x=t9ra3f>. In the preregistration protocol, we report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All stimuli, data, and code can be found in the Open Science Framework (OSF) repository: https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3.

Informed consent was obtained using a standard survey question. All studies were conducted in compliance with Institutional Review Board protocol.

Participants

As noted in our preregistration, rules-of-thumb for structural equation modeling suggest that 10 cases per indicator is a useful minimum to accurately estimate the model. Because we have 22 indicators, we aimed to collect at least 220 participants for each vignette in each country (440 participants per country). Anticipating having to exclude some participants (as per our preregistered criteria), we aimed to collect 500 participants in each country, for a total of 1,000 participants in both countries. Expecting that the associations between variables will correspond to at least medium-sized effects (i.e., r s between .3 and .5) or larger, this sample size provides around 90% power to achieve statistical significance ($p < .05$) with associations as small as $r = .16$ in each vignette and each country.

We recruited 500 participants from the United States and 508 from India in 2021 using Amazon Mechanical Turk. Following preregistered exclusion criteria, we removed from analyses 15 participants in the United States because they failed to pass an attention check. We removed from analyses 140 participants in India because they failed to pass an attention check, because they reported having only a “Basic” understanding of the English language, or both. Thereafter, the effective sample sizes were 368 (125 women) for India and 485 (261 women) for the United States. The mean age of participants was 30.69 ($SD = 6.83$) in India and 42.38 ($SD = 13.35$) in the United States.

The sociodemographic characteristics of the Indian sample are as follows: 45.9% of Indian respondents reported English as their first

language; 52.2% listed it as their second language; and 1.9% listed it as neither their first nor second language. Of the 198 Indian respondents who did not claim English as their first language, 54.6% cited Tamil as their first language, 25.0% cited Malayalam, and 20.4% cited other languages (two participants did not report any language). Additionally, 42.3% of Indian respondents reported living in a big town, 39.6% in a small- or medium-sized town, and 18.0% in a rural area or village (two participants did not respond). Finally, 68.3% of Indian participants hail from Tamil Nadu, 18.0% from Kerala, and 13.7% from other states (two participants did not respond).

Procedures and Materials

We created two vignettes in which participants were asked to imagine that they brought harm to a target individual unintentionally—an event that may elicit in participants guilt, shame, or both. We experimentally manipulated features of the vignettes that, according to each theory, differentially boost guilt or shame. According to attributional theory, a harm attributed to internal causes high in controllability and low in stability will differentially lead to guilt, whereas a harm attributed to internal causes low in controllability and high in stability will differentially lead to shame, irrespective of interpersonal factors. By contrast, functionalist theory predicts that a harm imposed unintentionally on a lower-status target whom one likes will differentially lead to guilt, whereas the same harm imposed on a disagreeable higher-status target whom one fears will differentially lead to shame, irrespective of intrapersonal factors.

For purposes of internal replication and contextual generalization, we created two different vignettes. In the “Botched Sale” vignette, participants inadvertently ruin the target’s big real estate sale. In the “Car Crash” vignette, participants accidentally bump into the target’s new car, damaging it. For each vignette, we used a 2×2 between-subjects, factorial design. First, we manipulated the description of the target to whom participants brought harm: This target was either the participant’s likeable intern or the participant’s mean boss (variables modulating guilt and shame, respectively, according to functionalist theory). Second, across target conditions, we manipulated factors that might lead participants to produce different patterns of controllability and stability attributions. More specifically, we manipulated the degree to which the harm participants brought to the target stemmed from participants’ actions or characteristics that appeared to be controllable and unstable versus uncontrollable and stable (variables modulating guilt and shame, respectively, according to attributional theory).

Thus, there were four conditions for each vignette. In the uncontrollable and stable conditions of the Botched Sale vignette, for example, participants’ Tourette syndrome causes them to shout profanities (an uncontrollable action attributable to a stable feature of themselves) and thus upset the target’s customer, thereby ruining the target’s big real estate sale. In the controllable and unstable conditions of the Botched Sale vignette, participants deliberately insult the target’s customer and thus ruin the target’s sale. In the likeable intern conditions, the target whose sale the participant ruins is both likeable and subordinate to the participant. In the mean boss conditions, the target whose sale the participant ruins is the participant’s tyrannical boss.

Participants in each country were randomly assigned to one of eight conditions (four conditions for each of two vignettes). Within each of the two core vignettes (Car Crash, Botched Sale),

participants were assigned to (a) the likeable intern/controllable and unstable condition, (b) the likeable intern/uncontrollable and stable condition, (c) the mean boss/controllable and unstable condition, or (d) the mean boss/uncontrollable and stable condition.

The vignettes were presented in two steps. Step 1 described background information (e.g., “You are a broker at a large real estate company”) and information about the target and their relationship to the participant. The likeable intern versus mean boss manipulation was introduced at Step 1. The Step 1 stimuli are the same across the Botched Sale vignettes and the Car Crash vignettes. Step 2 was presented subsequent to Step 1, on the next page. Step 2 described the harm the participant caused to the target: The participant either caused the target to lose a real estate sale (Botched Sale) or bumped into the target’s new car (Car Crash). The controllable and unstable versus uncontrollable and stable manipulation was introduced at Step 2. To increase the odds that participants actually read the vignettes, the “next” button allowing participants to proceed to the next screen was not displayed until Step 1 and Step 2 of the vignettes were displayed for 30 and 50 s, respectively.

After reading Step 2 of the vignette, participants rated the intensity of their shame and guilt, as well as their agreement with items hypothesized by each theory to mediate the elicitation of shame and guilt: controllability attributions, stability attributions, fear of the target, and liking of the target.

The lexical items “shame” and “guilt” lack precise counterparts across languages and cultures (Breugelmans & Poortinga, 2006; see Fiske, 2020). Moreover, the English words “shame” and “guilt” are frequently confused with each other (Tangney et al., 1998). By contrast, the motivational dimensions of evading versus amending appear to be central and structurally unique features of shame and guilt, respectively—and cross-culturally so (Breugelmans & Poortinga, 2006; T. R. Cohen et al., 2011; Giner-Sorolla & Espinosa, 2011; Tangney et al., 2007; Wicker et al., 1983). Therefore, following the frequently used method of measuring shame and guilt through their distinctive motivations (T. R. Cohen et al., 2011; Tangney et al., 2000; see Robins et al., 2007), we measured participants’ intensity of shame and guilt with items relevant to the motivations characteristic of these emotions.

We measured shame with five items (e.g., “I would feel like avoiding [Target] for a while,” “I would make up an excuse to avoid having to interact with [Target]”) coded from 1 (*not at all*) to 7 (*definitely*).

We measured guilt with four items (e.g., “I would go to [Target] and make things right with [Target],” “I would go to [Target] and apologize for it”) coded from 1 (*not at all*) to 7 (*definitely*).

We measured controllability attributions with three items (e.g., “I was in control of my own actions when I [caused harm to Target],” “I could have avoided [causing harm to Target] if I tried”) coded from 1 (*not at all*) to 7 (*definitely*).

We measured stability attributions with two items (e.g., “I [caused harm to Target] because I have a permanent, long-term problem,” “I [caused harm to Target] because of a chronic issue that I have”) coded from 1 (*not at all*) to 7 (*definitely*).

We measured fear of the target with four items (e.g., “How much do you fear [Target]?” coded from 1 (*not at all*) to 7 (*a lot*), “How bad would it be for you if you make [Target] angry?” coded from 1 (*not bad at all*) to 7 (*very bad*)).

We measured liking of the target with four items (e.g., “How much do you like [Target]?” coded from 1 (*not at all*) to 7 (*a lot*), “How close do you feel to [Target]?” coded from 1 (*not close at all*) to 7 (*very close*)).

Some items in the guilt, shame, controllability, and stability scales reference the harm caused by the participant, which varies across vignettes. Therefore, we tailored these items to the specific harm caused by the participant. However, the underlying semantic structure of each item remained the same across vignettes. For example, guilt item 1 is “I would go and confess to [Target] that I told [Target’s] customer to get out of the break room” (the harm done by the participant in the Botched Sale vignette) and “I would go and confess to [Target] that I damaged [Target’s] car” (the harm done by the participant in the Car Crash vignette).

Lastly, to control for possible effects of sex on judgments, participants were assigned same-sex targets with culturally common names (Robert and Ann in the United States; Amit and Divya in India).

Full vignettes and scales are available in the supplemental materials (see Items S1.1 and S1.2 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]). Henceforth, we use “S” to designate all items in the supplemental materials).

Analytic Procedures

Data wrangling was done in R. As per the preregistration, we conducted structural equation models (SEM) in Mplus where a shame latent variable and a guilt latent variable (each indicator defined by the shame and guilt scales, respectively, described above) were both simultaneously regressed on latent variables for liking [of target], fear [of target], controllability [attribution], and stability [attribution] (each indicator defined by its respective scale described above), and in turn, all of these latent variables were simultaneously regressed on two effect-coded variables representing each set of experimental conditions: (a) likeable intern versus mean boss and (b) uncontrollable and stable versus controllable and unstable. The latent variables for liking, fear, controllability, and stability were allowed to covary, as were guilt and shame. To formally compare the predictive power of variables, we specified model constraints that statistically tested the difference between absolute values of the specific paths or combinations of paths. We preregistered this comparison between magnitudes because we assumed (incorrectly, it turns out) that all variables hypothesized by the two theories to modulate shame and guilt would do so in their theoretically predicted directions. In some cases, following this preregistered plan of comparison masks true differences between predictions of the two theories. Therefore, to provide theoretically meaningful results, we account for the direction of the effects predicted by the theories where appropriate. However, to remain consistent with our preregistered plan, we note all such cases in our results and report the preregistered comparisons (not accounting for predicted direction) as well.

To examine trends across the two cultures and vignettes and in line with our preregistered plan, we ran six separate, focal models: one for India and one for the United States independently, with vignette treated as a stratification variable, and four additional models for each unique country and vignette combination (e.g., Car Crash in the United States). We also ran an overall model, in which we examined the associations of interest across the two cultures and vignettes, treating each as a stratification variable to account for nonindependence because of culture and vignette conditions. As shown in Table 1, the fit of all models ranged from acceptable to good.

Table 1
Model Fit Indices for SEM Models

Model	<i>n</i>	CFI	RMSEA
Across vignettes and countries	853	.949	.055
United States across vignettes	485	.945	.061
United States Botched Sale	242	.948	.068
United States Car Crash	243	.930	.082
India across vignettes	368	.933	.053
India Botched Sale	184	.896	.075
India Car Crash	184	.927	.068

Note. Each model estimated 93 parameters. SEM = structural equation model; CFI = comparative fit index; RMSEA = root-mean-square error of approximation.

Results

The standardized path estimates from the focal SEM analyses examining the associations across vignettes in each culture while accounting for the relevant dependencies in the data are shown in Figures 1 and 2. Overall path estimates (collapsing across cultures and vignettes) as well as culture-specific path estimates are reported within the text. For brevity, vignette-specific trends are described in the text and listed in Tables 2, 3, and 4. For full analyses of all vignette-specific trends, see Tables S3.2.1 and S3.5 of the supplemental materials, which are available on the OSF page (https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3). Descriptive statistics are shown in Table S2.3 in the supplemental materials, which are available on the OSF page (https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3).

Manipulation Checks

For the most part, the experimental manipulations worked as intended. Participants assigned to the mean boss conditions reported more fear ($\beta = .34$, 95% confidence interval, CI [.27, .40], $p < .001$) and less liking ($\beta = -.46$, 95% CI [-.51, -.41], $p < .001$) toward the harmed target compared with participants in the likable intern conditions. Additionally, participants assigned to the controllable and unstable conditions made attributions of greater controllability ($\beta = .47$, 95% CI [.42, .52], $p < .001$) and less stability ($\beta = -.42$, 95% CI [-.48, -.37], $p < .001$) than did participants in the uncontrollable and stable conditions. Moreover, the manipulations exerted unique effects on the specific mediators relevant to each theory: The target manipulations did not influence controllability or stability attributions, nor did the controllability or stability manipulations influence participants' liking or fear of the target (Figures 1 and 2; exact results in Item S3.2.1 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]). All the experimental manipulations worked as intended in the United States. In India, the experimental manipulations of controllability worked, but the experimental manipulations of stability did not: For neither vignette did participants in India reliably attribute greater stability in the uncontrollable and stable conditions than in the controllable and unstable conditions (exact results in Item S3.2.1 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]).

Tests of Functionalist Theory

Was greater fear of the target associated with more shame, as functionalist theory predicts? Yes. When participants harmed a target,

they reported more shame the more they feared the target, collapsing across countries and core vignettes ($\beta = .59$, 95% CI [.51, .68], $p < .001$). This pattern of results replicated within each country: within the United States ($\beta = .71$, 95% CI [.62, .79], $p < .001$) and within India ($\beta = .31$, 95% CI [.16, .46], $p < .001$). This pattern replicated for each vignette considered individually within the United States. And it was replicated in one of two vignettes in India (Item S3.2.1 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]).

Was greater liking of the target associated with more guilt, as functionalist theory predicts? Yes. When participants harmed a target, they reported more guilt the more they liked the target, collapsing across countries and core vignettes ($\beta = .69$, 95% CI [.62, .76], $p < .001$). Further, this pattern of results replicated within each country: within the United States ($\beta = .67$, 95% CI [.57, .77], $p < .001$) and within India ($\beta = .71$, 95% CI [.57, .84], $p < .001$). This pattern held for each vignette considered individually within both the United States and India (Item S3.2.1 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]).

Was shame better predicted by fearing the target than by liking the target, as functionalist theory predicts? Yes. Fearing the target they harmed led participants to report more shame than did liking the target, collapsing across countries and core vignettes ($\beta = .34$, 95% CI [.23, .46], $p < .001$). This pattern held across vignettes within the United States ($\beta = .55$, 95% CI [.40, .70], $p < .001$). It did not replicate within India ($\beta = .07$, 95% CI [-.06, .20], $p = .297$). This pattern held for each vignette considered individually within the United States but for neither vignette within India (Item S3.2.1 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]).

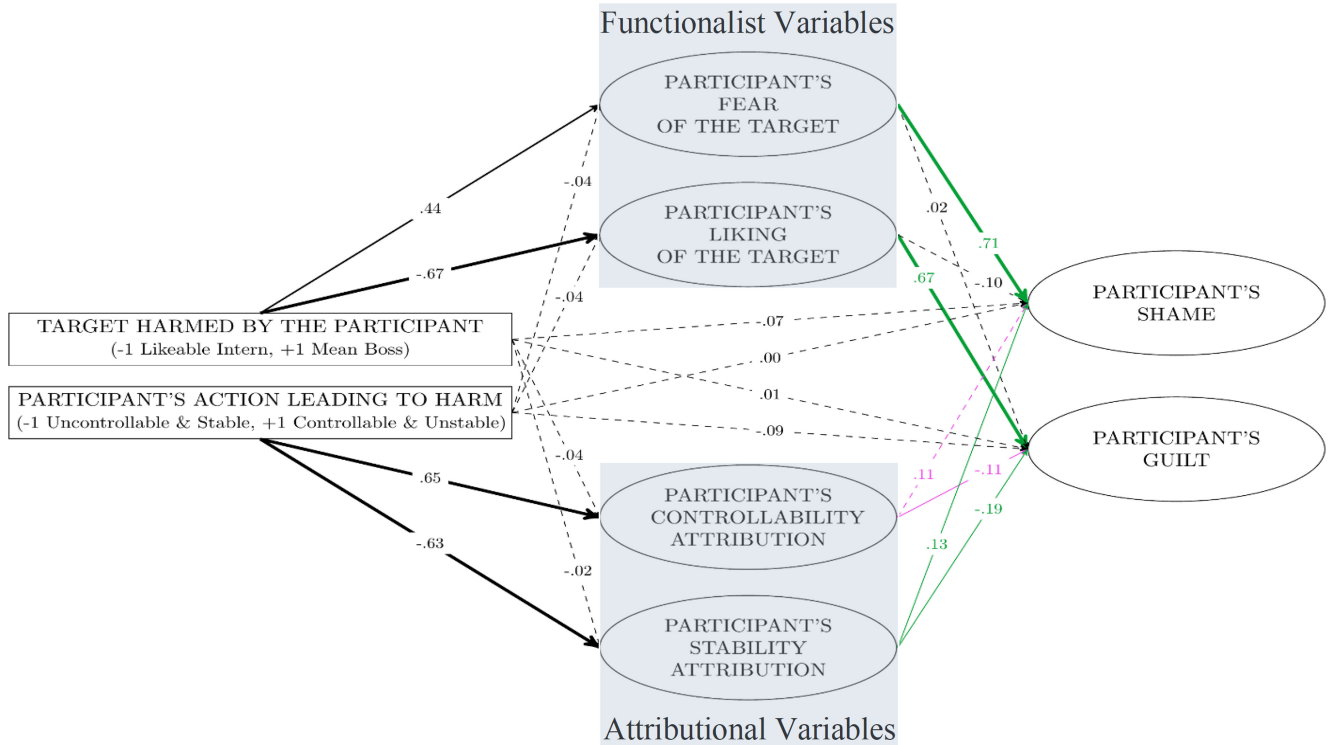
Was guilt better predicted by liking the target than by fearing the target, as functionalist theory predicts? Yes. Liking the target they harmed led participants to report more guilt than did fearing the target, collapsing across countries and core vignettes ($\beta = .67$, 95% CI [.55, .80], $p < .001$). This pattern of results held across vignettes within both the United States ($\beta = .71$, 95% CI [.53, .89], $p < .001$) and India ($\beta = .68$, 95% CI [.43, .93], $p < .001$). This pattern replicated for each vignette considered individually within each country (Item S3.2.1 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]).

Tests of Attributional Theory

Was attributing less controllability associated with more shame, as attributional theory predicts? No. In fact, when participants harmed a target, they reported *less* shame the less they attributed controllability to their actions, collapsing across countries and core vignettes ($\beta = .23$, 95% CI [.14, .32], $p < .001$). Greater controllability was marginally associated with more shame across vignettes within the United States ($\beta = .11$, 95% CI [.00, .22], $p = .057$). Further, an even stronger positive association between controllability and shame held across vignettes within India ($\beta = .40$, 95% CI [.22, .58], $p < .001$). This positive association was significant for each vignette considered individually within India but only marginally significant for one vignette within the United States (Item S3.2.1 in the supplemental materials, which are available on the OSF page

Figure 1

Depiction of the Focal SEM Analysis Examining the Predictors of Latent Shame and Guilt Reported by Participants in the United States



Note. Boxes represent observed variables (in this case, the effect-coded experimental conditions). Circles represent latent unobserved variables. To simplify the presentation, the indicators of the latent variables and their respective loadings are not depicted but are presented in the supplemental materials (Item S3.4 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]). Likewise, the covariances between the latent variables are not depicted but are reported in the supplemental materials (Item S3.3 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3])). The numbers on the paths depict the standardized model estimates. Dashed lines indicate that a path was not statistically significant ($p \geq .05$); solid lines indicate the path is statistically significant ($p < .05$); thickness of the lines is proportionate to effect size (thin: $|\beta| < .3$, thicker: $.3 \leq |\beta| \leq .5$, thickest: $|\beta| > .5$); and the color of the lines corresponds to the accuracy of predictions made either by functionalist theory or attributional theory as to the effect of fear and liking or stability and controllability, respectively, on shame and guilt (green [online]/dark gray [print] indicates an effect in the predicted direction by the corresponding theory and pink [online]/light gray [print] indicates an effect in the opposite direction predicted by the corresponding theory). SEM = structural equation model. See the online article for the color version of this figure.

[https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]).

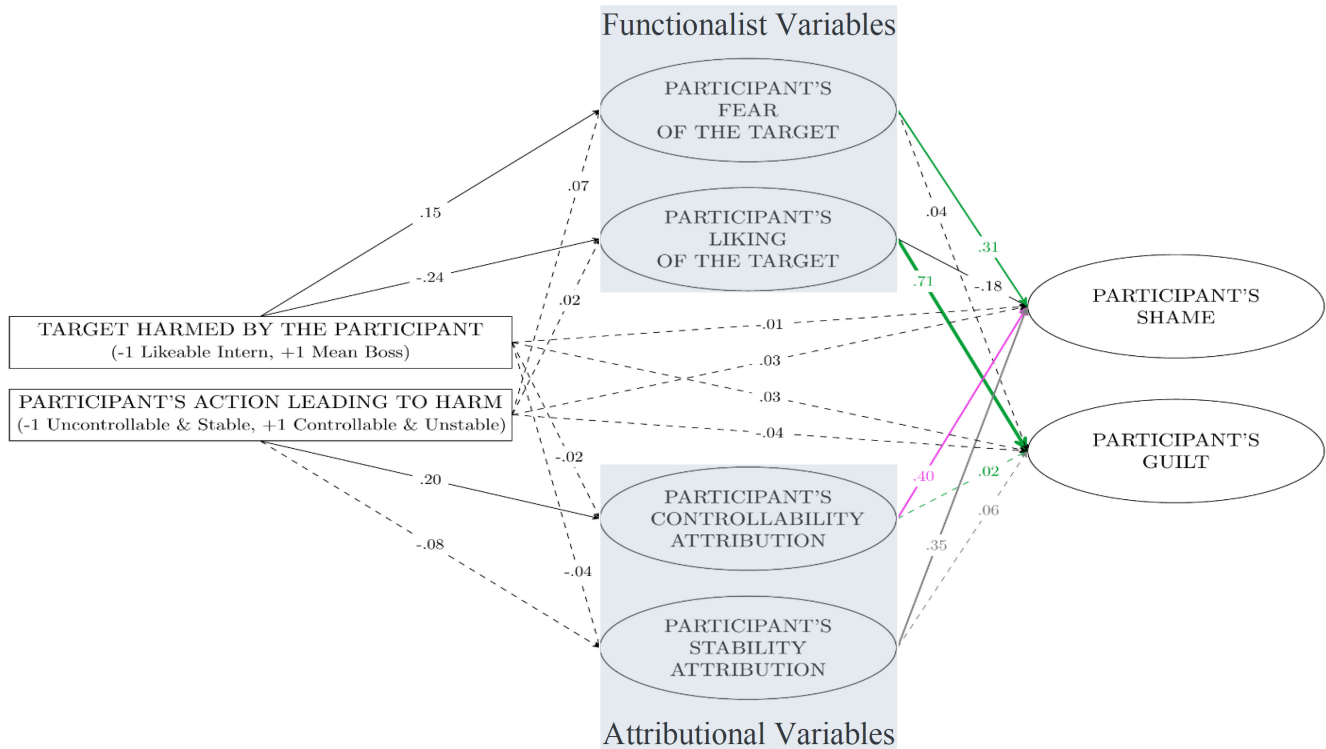
Was attributing greater stability associated with more shame, as attributional theory predicts? Yes. When participants harmed a target, they reported more shame the more they attributed stability to the cause of their actions, collapsing across countries and core vignettes ($\beta = .26$, 95% CI [.16, .35], $p < .001$). This pattern of results held within the United States ($\beta = .13$, 95% CI [.01, .25], $p = .029$) and within India ($\beta = .35$, 95% CI [.16, .55], $p < .001$). These effects replicated for both vignettes in India. In the United States, stability was positively associated with shame in one of two vignettes. The tests examining the effect of stability in the two core vignettes in India, including the test collapsing across vignettes in India, must be interpreted with caution, however, because in both vignettes in India the experimental manipulation of stability failed to exert its intended effect (Item S3.2.1 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3])).

Was attributing greater controllability associated with more guilt, as attributional theory predicts? No. When participants harmed a target, they did not report more guilt the more they attributed controllability to their actions, collapsing across countries and core vignettes ($\beta = -.04$, 95% CI [-.12, .04], $p = .320$). There was a weak negative association between controllability and guilt in the United States ($\beta = -.11$, 95% CI [-.22, -.01], $p = .040$). However, there was no relationship between controllability and guilt within India ($\beta = .02$, 95% CI [-.21, .25], $p = .875$). No relationship was found between controllability and guilt for either vignette considered individually in either country (Item S3.2.1 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3])).

Was attributing greater stability associated with less guilt, as attributional theory predicts? Yes, marginally. When participants harmed a target, they reported marginally less guilt the more they attributed stability to the cause of their actions, collapsing across countries and core vignettes ($\beta = -.08$, 95% CI [-.16, .00], $p = .063$). There was a negative association between stability and guilt in the United States

Figure 2

Depiction of the Focal SEM Analysis Examining the Predictors of Latent Shame and Guilt Reported by Participants in India



Note. Boxes represent observed variables (in this case, the effect-coded experimental conditions). Circles represent latent unobserved variables. To simplify the presentation, the indicators of the latent variables and their respective loadings are not depicted but are presented in the supplemental material (Item S3.4 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]). Likewise, the covariances between the latent variables are not depicted but are reported in the supplemental materials (Item S3.3 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]). The numbers on the paths depict the standardized model estimates. Dashed lines indicate that a path was *ns* ($p \geq .05$); solid lines indicate the path is statistically significant ($p < .05$); thickness of the lines is proportionate to effect size (thin: $|\beta| < .3$, thicker: $.3 \leq |\beta| \leq .5$, thickest: $|\beta| > .5$); and the color of the lines corresponds to the accuracy of predictions made either by functionalist theory or attributional theory as to the effect of fear and liking or stability or controllability, respectively, on shame and guilt (green [online]/dark gray [print] indicates an effect in the predicted direction by the corresponding theory; pink [online]/light gray [print] indicates an effect in the opposite direction predicted by the corresponding theory; and gray indicates an effect for which the manipulation failed). SEM = structural equation model. See the online article for the color version of this figure.

($\beta = -.19$, 95% CI $[-.31, -.08]$, $p = .001$). There was no relationship between stability and guilt within India ($\beta = .06$, 95% CI $[-.16, .29]$, $p = .583$). The overall relationship between stability and guilt found in the United States was found in neither of the two core vignettes in the United States and in neither of the two core vignettes in India (Item S3.2.1 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]). The tests examining the effect of stability in the two core vignettes in India, including the test collapsing across vignettes in India, must be interpreted with caution, however, because as noted above, in both vignettes in India the experimental manipulation of stability failed to exert its intended effect.

Was attributing greater controllability associated with more guilt than shame, as attributional theory predicts? No. Across countries and core vignettes, the relationship between controllability and shame slightly exceeded that between controllability and guilt ($\beta = -.10$, 95% CI $[-.19, -.01]$, $p = .028$). Note that both relationships were in the *opposite* direction of that predicted by attributional

theory. In the United States, greater controllability did not significantly lead to more guilt than shame ($\beta = .04$, 95% CI $[-.08, .16]$, $p = .529$). And in India, greater controllability led to more shame than guilt, though nonsignificantly ($\beta = -.16$, 95% CI $[-.41, .08]$, $p = .184$). For a full breakdown of the data for each individual vignette within each country, see Item S3.5 in the supplemental materials, which are available on the OSF page (https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3).

Was attributing greater stability associated with more shame than guilt, as attributional theory predicts? No. The relationship between stability and shame was not statistically different than that between stability and guilt, collapsing across countries and core vignettes ($\beta = -.08$, 95% CI $[-.17, .02]$, $p = .109$). This was true both in the United States ($\beta = .10$, 95% CI $[-.03, .22]$, $p = .118$) and in India ($\beta = -.10$, 95% CI $[-.32, .12]$, $p = .373$). The tests examining the effect of stability in the two core vignettes in India, including the test collapsing across vignettes in India, must be interpreted with caution, however, because in both vignettes in India the

Table 2*Summary of the Tests of Functionalist Theory in Each Country and for Each Vignette*

Functionalist theory	United States		India	
	Car crash	Botched sale	Car crash	Botched sale
1. Did greater fear of the target predict more shame?	.70 ^a	.74 ^a	.12 ^b	.47 ^a
2. Did greater liking of the target predict more guilt?	.52 ^a	.79 ^a	.77 ^a	.65 ^a
3. Was shame predicted more strongly by fear of the target than by liking of the target?	.56 ^a	.59 ^a	-.01 ^c	.10 ^b
4. Was guilt predicted more strongly by liking of the target than by fear of the target?	.54 ^a	.82 ^a	.70 ^a	.64 ^a

Note. Displayed on the four rightmost columns are observed path coefficients indicating the unique effects of the independent variable on the dependent variable or the difference between such coefficients. See the online article for the color version of this table.

^a Significant effect in the direction predicted by functionalist theory (13 of 16 effects). ^b Nonsignificant effect in the direction predicted by functionalist theory (two of 16 effects). ^c Nonsignificant effect in the opposite direction as predicted by functionalist theory (one of 16 effects).

experimental manipulation of stability failed to exert its intended effect. For vignette-specific analyses, see Item S3.5 in the supplemental materials, which are available on the OSF page (https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3).

Contrastive Tests of the Two Theories

Was shame better predicted by functionalist theory or attributional theory? By functionalist theory. Fear of the target was a better predictor of shame than both controllability attributions ($\beta = .31$, 95% CI [.20, .43], $p < .001$) and stability attributions ($\beta = .30$, 95% CI [.18, .43], $p < .001$), collapsing across countries and core vignettes. What is more, fear of the target was a better predictor of shame than both controllability attributions and stability attributions combined, ($\beta = .17$, 95% CI [.02, .31], $p = .023$). In the United States, fear of the target better predicted shame than controllability attributions ($\beta = .54$, 95% CI [.39, .70], $p < .001$) and stability attributions ($\beta = .54$, 95% CI [.38, .70], $p < .001$), as well as the latter two combined ($\beta = .46$, 95% CI [.26, .67], $p < .001$). In India, there was no reliable difference between how well fear of the target on the one hand and controllability attributions on the other predicted shame ($\beta = -.01$, 95% CI [-.14, .12], $p = .869$). However, this lack of difference in India is illusory—a consequence of assuming in the preregistration (incorrectly, it turns out) that the attributional variables would always yield effects in the direction predicted by attributional theory (we report these nominal effects here for the

purpose of matching the preregistration plan). As functionalist theory predicts, more fear did lead to more shame in India ($\beta = .31$, 95% CI [.16, .46], $p < .001$). But contra the prediction of attributional theory, more controllability led to more shame in India, not less (India: $\beta = .40$, 95% CI [.22, .58], $p < .001$). Therefore, the true and theoretically meaningful result in India is that fear of the target predicted shame better than controllability attributions did ($\beta = .35$, 95% CI [.18, .52], $p < .001$). There was no significant difference between how well fear of the target on the one hand and stability attributions on the other predicted shame in India ($\beta = .02$, 95% CI [-.13, .16], $p = .829$). Note, however, that the tests comparing the effects of fear on shame and of stability on shame within India must be interpreted with caution because the experimental manipulation of stability in India failed to exert its intended effect.

Within each vignette in the United States, fear of the target predicted shame better than either controllability or stability predicted shame, as well as the latter two combined. In each vignette in India, there was no difference between the predictive power of fear and controllability on shame. However, once again, this result is illusory: Contra the prediction of attributional theory, more controllability actually led to more shame in India. In reality, fear of the target predicted shame better than controllability attributions did for each vignette in India (Item S3.5 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]). In India, there was likewise no difference between the predictive power of fear

Table 3*Summary of the Tests of Attributional Theory in Each Country and for Each Vignette*

Attributional theory	United States		India	
	Car crash	Botched sale	Car crash	Botched sale
1. Did less controllability predict more shame?	.13 ^c	.03 ^c	.49 ^d	.37 ^d
2. Did greater stability predict more shame?	.21 ^a	-.01 ^c	.32 ^e	.36 ^e
3. Did greater controllability predict more guilt?	-.05 ^c	-.03 ^c	-.08 ^c	.07 ^b
4. Did greater stability predict less guilt?	-.12 ^b	-.07 ^b	.04 ^e	.17 ^e
5. Did greater controllability predict more guilt than shame?	-.04 ^c	.01 ^b	-.10 ^c	-.16 ^c
6. Did greater stability predict more shame than guilt?	-.03 ^c	.06 ^b	-.08 ^e	-.06 ^e

Note. Displayed on the four rightmost columns are observed path coefficients indicating the unique effects of the independent variable on the dependent variable or the difference between such coefficients. See the online article for the color version of this table.

^a Significant effect in the direction predicted by attributional theory (one of 18[24] effective[total] effects). ^b Nonsignificant effect in the direction predicted by attributional theory (five of 18[24] effective[total] effects). ^c Nonsignificant effect in opposite direction as predicted by attributional theory (10 of 18[24] effective[total] effects). ^d Significant effect in opposite direction as predicted by attributional theory (two of 18[24] effective[total] effects). ^e Test not clear because stability manipulation failed to exert intended effect (six of 24 total effects).

Table 4*Summary of the Tests Contrasting Functionalist Theory With Attributional Theory in Each Country and for Each Vignette*

Contrastive test	United States		India	
	Car crash	Botched sale	Car crash	Botched sale
1. Was shame predicted more strongly by fear than by stability?	.48 ^c	.63 ^c	-.06 ^b	.08 ^b
2. Was shame predicted more strongly by fear than by controllability?	.53 ^c	.62 ^c	.23 ^{a,c}	.48 ^{a,c}
3. Was guilt predicted more strongly by liking than by stability?	.43 ^c	.77 ^c	.76 ^b	.52 ^b
4. Was guilt predicted more strongly by liking than by controllability?	.50 ^c	.80 ^c	.73 ^c	.59 ^c

Note. Displayed on the four rightmost columns are observed path coefficients indicating the difference between unique effects of the independent variable on the dependent variable. See the online article for the color version of this table.

^a See Item S3.5 in the supplemental materials, which are available on the OSF page (https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3), for results not accounting for direction predicted by attributional theory. ^b Test not clear because stability manipulation failed to exert intended effect (4 of 16 total effects). ^c Significant difference favoring functionalist theory (12 of 12[16] effective[total] effects).

and stability on shame (Item S3.5 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]). Note that tests between the effects of fear on shame and stability on shame (or between fear and both attributional variables combined) within India must be interpreted with caution because in both vignettes in India the experimental manipulation of stability failed to exert its intended effect.

Was guilt better predicted by functionalist theory or attributional theory? By functionalist theory. Liking the target was a better predictor of guilt than were both controllability attributions ($\beta = .65$, 95% CI [.56, .74], $p < .001$) and stability attributions ($\beta = .62$, 95% CI [.53, .71], $p < .001$), collapsing across countries and core vignettes. What is more, liking the target was a better predictor of guilt than both controllability attributions and stability attributions combined ($\beta = 0.58$, 95% CI [.48, 0.68], $p < .001$). In the United States, liking the target was a better predictor of guilt than both controllability attributions ($\beta = .62$, 95% CI [.48, .75], $p < .001$) and stability attributions ($\beta = .55$, 95% CI [.41, .69], $p < .001$), as well as the latter two combined, ($\beta = .43$, 95% CI [.26, .61], $p < .001$). Within India, too, liking the target better predicted guilt than both controllability attributions ($\beta = .71$, 95% CI [.46, .96], $p < .001$) and stability attributions ($\beta = .67$, 95% CI [.40, .95], $p < .001$), as well as the latter two combined, ($\beta = .66$, 95% CI [.41, .90], $p < .001$). Liking predicted guilt better than the two attributional variables combined for each individual vignette in the United States (Item S3.5 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]). And liking predicted guilt better than controllability for each individual vignette in India (Item S3.5 in the supplemental materials, which are available on the OSF page [https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3]). Note that tests comparing the effect of liking on guilt and stability on guilt (or between liking and both attributional variables combined) within India must be interpreted with caution because in both vignettes in India, the experimental manipulation of stability failed to exert its intended effect.

Exploratory Analyses

We conducted exploratory analyses to determine possible sex differences in our measures. In the United States, ratings of fear of the target and shame were significantly higher among females than males; ratings of attributions of controllability were significantly higher among males than females; there were no significant sex differences in ratings of liking

of the target, guilt, or attributions of stability. In India, ratings of attributions of controllability, attributions of stability, liking of the target, guilt, and shame were significantly higher among females than males; there were no significant sex differences in ratings of fear of the target. The only sex difference that emerged reliably across countries, therefore, was higher ratings of shame among females than males. Item S4 in the supplemental materials, which are available on the OSF page (https://osf.io/fw2cu/?view_only=f7b41aac89fb4015b0aa304b24cc26f3), includes descriptive and inferential statistics relevant to sex differences.

General Discussion

The results of these tests support functionalist theory. Greater fear of an individual one harmed unintentionally is associated with more intense shame motivations. This is expected if shame functions to minimize the cost of being devalued and if greater anticipated devaluative costs (here, from an intimidating person) elicit shame more intensely (A. S. Cohen et al., 2020; Durkee et al., 2019; Sznycer & Cohen, 2021; Sznycer et al., 2016, 2018). And greater liking of an individual one harmed unintentionally elicits more intense guilt motivations (see Sznycer & Lukaszewski, 2019). This is expected if guilt functions to curb the undue imposition of costs on valued others. Importantly, fear and liking exerted specific effects on shame and guilt, respectively. These effects were largely replicated both between cultures and within cultures, across different vignettes.

By contrast, the data credit attributional theory with far fewer supporting results. In the United States, high stability attributions exerted their predicted boosting effects on shame, while low stability attributions exerted their predicted boosting effects on guilt. Otherwise, the effects of stability and controllability attributions were either null or opposite to what attributional theory predicts. In the United States, high controllability attributions decreased guilt significantly and increased shame (though not significantly). In India, contra attributional theory's prediction, higher controllability attributions boosted shame. Importantly, across countries and core vignettes, attributions of controllability failed to elicit more guilt than shame, while attributions of stability failed to elicit more shame than guilt. Further, the attributional variables fared worse than the functionalist variables in the contrastive tests. Across countries and core vignettes, guilt was predicted more strongly by liking the target than by attributions of stability, controllability, or stability and controllability combined. Similarly, shame was predicted more strongly by fear of the target than by attributions of stability, controllability, or stability and controllability combined.

The observed strengths of functionalist theory tended to be echoed across both the United States and India. These cross-cultural regularities are striking, especially when considering that prominent cultural theories expect these disparate cultures to exhibit guilt and shame differentially because of their varying individualism–collectivism (Fessler, 2004; Muthukrishna et al., 2020). In addition, these findings suggest that other self-conscious emotions too—pride, for instance—may be more properly understood as interpersonal emotions (Durkee, 2022; Sznycer, 2019; Tracy et al., 2010).

If functionalist theory is correct and shame and guilt function to solve interpersonal challenges, this theory needs to explain the many reported effects that are consistent with attributional theory. We suggest that cognition about the self in shame and guilt is a concomitant of the interpersonal functions of these emotions. For example, minimizing the threat of being devalued (shame) is less within the actor's control than minimizing the excessive imposition of costs on valued others (guilt). This difference may explain why "controllability" has been associated with shame less than with guilt. Further, whereas guilt generally follows unintentional impositions of harm (McGraw, 1987), shame follows events both unintentional (e.g., being diseased) and intentional (e.g., stealing). This difference may explain why shame episodes feel more stably diagnostic of the individual's character ("stability") than guilt episodes do. Cognition about the self in shame and guilt may also be instrumental to interpersonal functions. For example, feelings of worthlessness or low self-esteem in shame may echo actual reductions in the social value that other people impute to the focal individual (Eisenberger et al., 2011; Leary et al., 1995; Sznycer et al., 2016; see also Cooley, 1902). That the intensity of shame in the individual tracks the intensity of devaluation in the audience—within and between industrial societies (Durkee et al., 2019; Leroux et al., 2023; Sznycer et al., 2016) and small-scale societies (Sznycer et al., 2018) and over millennia (Sznycer & Patrick, 2020)—is likely a design feature of the shame system, enabling the system to deploy anti-devaluation measures of the right magnitude and type. Additionally, the self-blaming that people engage in when ashamed might function as a signal of submission to deter attacks. Consistent with this, women with abusive partners blame themselves, but they blame their partners when they leave the relationship (Gilbert, 2016). Although some of these hypotheses remain to be evaluated, and further research is necessary to reassess and integrate the many findings consistent with attributional theory, functionalist theory represents a promising framework for explaining both interpersonal and intrapersonal aspects of shame and guilt (Landers et al., in press; Sznycer, 2019). Thus, functionalist theory can lead to "unification within an abstract and overarching theoretical framework that focuses on psychological processes common to the two classes of explanation and treats points of difference as tractable problems of parameter estimation" (Tetlock & Manstead, 1985, p. 74).

These findings have practical significance. Guilt and shame are implicated in the etiology of mental and physical health problems, including anxiety, depression, trauma, alcoholism, and suicide (Aakvaag et al., 2016; Kealy et al., 2021; Randles & Tracy, 2013). The attributional view that shame is a pathogenic emotion leads naturally to interventions to reduce the occurrence of shame (see, e.g., Goffnett et al., 2020)—a dominant therapeutic paradigm at present. By contrast, the functionalist perspective the data favor views shame as a well-designed adaptation that transmits bad news to a suite of countermeasures to defeat the devaluative threat (although it is of course possible that in some individuals the shame system is miscalibrated or impaired for a variety of reasons; see Cosmides & Tooby, 1999). Therefore, interventions to reduce shame

absent efforts to mitigate the causes of devaluation may be counterproductive, akin to interventions to reduce physical pain absent treatment of the underlying injury or disease.

Constraints on Generality

Our experimental manipulation of stability failed to exert its intended effect in India. We speculate this may be because of the relatively small effective sample size and low English proficiency in India. Recall that though the initial samples recruited from India and the United States were similar in size, preregistered exclusion criteria dictated that we remove from analyses over a quarter of the Indian sample because of failure to pass an attention check, failure to pass an English language check, or both. Even within the effective sample in India, nearly 22% of participants claimed only an "intermediate English" ability. The combined effects of a small effective sample and low English proficiency may have led to the failed manipulation of stability in India. Note that all the experimental manipulations, and not only the manipulation of stability, exerted smaller effects on the mediators (or manipulation checks) in India than in the United States. Note also that, though not significant, in each vignette the manipulation of stability exerted an effect in the expected direction in India. However, despite the failure of the stability manipulation in India, we were able to evaluate two key aspects of attributional theory: the effect of controllability on shame and guilt (cross-culturally, in the United States and India), and the effect of stability on shame and guilt (in the United States). Nevertheless, future research would profit from improved experimental manipulations of stability.

Whether the results observed here generalize to other eliciting situations, to age classes other than the ones evaluated here, and to participants from other cultures remains to be determined. Likewise, future research might explore whether the results obtained with our vignette-based design generalize to more direct, behavioral measures of shame and guilt. Regarding this point, we note that measuring state shame and guilt via self-report has a long history in psychology (e.g., Otterbacher & Munz, 1973; Turner, 1998), and results obtained from carefully constructed vignette-based studies can indeed mirror those obtained from behavioral measures (e.g., Hainmueller et al., 2015). Finally, we note that our experimental manipulations of the target's identity were yoked in crossover fashion: Less guilt and more shame were elicited by the mean boss, compared to the likeable intern. But if, as functionalist theory argues, guilt and shame are different systems, then their activation need not be so yoked. Further research is needed to address these questions.

Concluding Remarks

The present findings cast doubt on the attributional account of shame and guilt. They suggest instead that shame and guilt function to solve interpersonal challenges: to minimize the direct costs of being devalued by others and the indirect costs of imposing undue harm on valued others, respectively. The present critical tests are the first to our knowledge to pit competing predictions of attributional and functionalist theories against each other. When evaluating competing theories, performing such critical tests is preferable: Critical tests decrease the odds of false positives to which simple confirmatory tests are more susceptible. As we suggest above, functionalist theory does not dispute the narrow attributional claim that guilt and shame recruit intrapersonal processes. Rather, the functionalist argument is that the intrapersonal aspects of guilt and shame are proximate, tributary, or incidental to the interpersonal adaptive functions of these emotions. In short, the self-conscious emotions may not be about the self after all.

References

- Aakvaag, H. F., Thoresen, S., Wentzel-Larsen, T., Dyb, G., Røysamb, E., & Olff, M. (2016). Broken and guilty since it happened: A population study of trauma-related shame and guilt after violence and sexual abuse. *Journal of Affective Disorders*, 204, 16–23. <https://doi.org/10.1016/j.jad.2016.06.004>
- Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences*, 274(1610), 749–753. <https://doi.org/10.1098/rspb.2006.0209>
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin*, 115(2), 243–267. <https://doi.org/10.1037/0033-2909.115.2.243>
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1995). Personal narratives about guilt: Role in action control and interpersonal relationships. *Basic and Applied Social Psychology*, 17(1–2), 173–198. https://doi.org/10.1207/s15324834basp1701&2_10
- Benedict, R. (1946). *The chrysanthemum and the sword*. Houghton Mifflin.
- Breugelmans, S. M., & Poortinga, Y. H. (2006). Emotion without a word: Shame and guilt among Rarāmuri Indians and rural Javanese. *Journal of Personality and Social Psychology*, 91(6), 1111–1122. <https://doi.org/10.1037/0022-3514.91.6.1111>
- Cohen, A. S., Chun, R., & Sznycer, D. (2020). Do pride and shame track the evaluative psychology of audiences? Preregistered replications of Sznycer et al. (2016, 2017). *Royal Society Open Science*, 7(5), Article 191922. <https://doi.org/10.1098/rsos.191922>
- Cohen, T. R., Wolf, S. T., Panter, A. T., & Insko, C. A. (2011). Introducing the GASP scale: A new measure of guilt and shame proneness. *Journal of Personality and Social Psychology*, 100(5), 947–966. <https://doi.org/10.1037/a0022641>
- Cooley, C. H. (1902). *Human nature and the social order*. Scribner's.
- Cosmides, L., & Tooby, J. (1999). Toward an evolutionary taxonomy of treatable conditions. *Journal of Abnormal Psychology*, 108(3), 453–464. <https://doi.org/10.1037/0021-843X.108.3.453>
- Dearing, R. L., Stuewig, J., & Tangney, J. P. (2005). On the importance of distinguishing shame from guilt: Relations to problematic alcohol and drug use. *Addictive Behaviors*, 30(7), 1392–1404. <https://doi.org/10.1016/j.addbeh.2005.02.002>
- Declerck, C. H., Boone, C., & Kiyonari, T. (2014). No place to hide: When shame causes proselvs to cooperate. *The Journal of Social Psychology*, 154(1), 74–88. <https://doi.org/10.1080/00224545.2013.855158>
- de Hooge, I. E., Breugelmans, S. M., & Zeelenberg, M. (2008). Not so ugly after all: When shame acts as a commitment device. *Journal of Personality and Social Psychology*, 95(4), 933–943. <https://doi.org/10.1037/a0011991>
- de Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2007). Moral sentiments and cooperation: Differential influences of shame and guilt. *Cognition & Emotion*, 21(5), 1025–1042. <https://doi.org/10.1080/02699930600980874>
- de Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2010). Restore and protect motivations following shame. *Cognition & Emotion*, 24(1), 111–127. <https://doi.org/10.1080/02699930802584466>
- Dickerson, S. S., Gruenewald, T. L., & Kemeny, M. E. (2004). When the social self is threatened: Shame, physiology, and health. *Journal of Personality*, 72(6), 1191–1216. <https://doi.org/10.1111/j.1467-6494.2004.00295.x>
- Durkee, P. K. (2022). *Probing the fundamental features of human status psychology* [Doctoral dissertation]. University of Texas. <https://repositories.lib.utexas.edu/handle/2152/115120>
- Durkee, P. K., Lukaszewski, A. W., & Buss, D. M. (2019). Pride and shame: Key components of a culturally universal status management system. *Evolution and Human Behavior*, 40(5), 470–478. <https://doi.org/10.1016/j.evolhumbehav.2019.06.004>
- Eisenberger, N. I., Inagaki, T. K., Muscatell, K. A., Byrne Haltom, K. E., & Leary, M. R. (2011). The neural sociometer: Brain mechanisms underlying state self-esteem. *Journal of Cognitive Neuroscience*, 23(11), 3448–3455. https://doi.org/10.1162/jocn_a_00027
- Elison, J., Garofalo, C., & Velotti, P. (2014). Shame and aggression: Theoretical considerations. *Aggression and Violent Behavior*, 19(4), 447–453. <https://doi.org/10.1016/j.avb.2014.05.002>
- Feiring, C., Taska, L., & Lewis, M. (1998). The role of shame and attributional style in children's and adolescents' adaptation to sexual abuse. *Child Maltreatment*, 3(2), 129–142. <https://doi.org/10.1177/1077559598003002007>
- Fessler, D. (2004). Shame in two cultures: Implications for evolutionary approaches. *Journal of Cognition and Culture*, 4(2), 207–262. <https://doi.org/10.1163/1568537041725097>
- Fiske, A. P. (2020). The lexical fallacy in emotion research: Mistaking vernacular words for psychological entities. *Psychological Review*, 127(1), 95–113. <https://doi.org/10.1037/rev0000174>
- Gilbert, P. (1997). The evolution of social attractiveness and its role in shame, humiliation, guilt and therapy. *British Journal of Medical Psychology*, 70(2), 113–147. <https://doi.org/10.1111/j.2044-8341.1997.tb01893.x>
- Gilbert, P. (2000). The relationship of shame, social anxiety and depression: The role of the evaluation of social rank. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, 7(3), 174–189. [https://doi.org/10.1002/1099-0879\(200007\)7:3<174::CO;2-U](https://doi.org/10.1002/1099-0879(200007)7:3<174::CO;2-U)
- Gilbert, P. (2016). *Depression: The evolution of powerlessness*. Routledge.
- Giner-Sorolla, R., & Espinosa, P. (2011). Social cuing of guilt by anger and of shame by disgust. *Psychological Science*, 22(1), 49–53. <https://doi.org/10.1177/0956797610392925>
- Goffnett, J., Liechty, J. M., & Kidder, E. (2020). Interventions to reduce shame: A systematic review. *Journal of Behavioral and Cognitive Therapy*, 30(2), 141–160. <https://doi.org/10.1016/j.jbct.2020.03.001>
- Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, 112(8), 2395–2400. <https://doi.org/10.1073/pnas.1416587112>
- Kealy, D., Treeby, M. S., & Rice, S. M. (2021). Shame, guilt, and suicidal thoughts: The interaction matters. *British Journal of Clinical Psychology*, 60(3), 414–423. <https://doi.org/10.1111/bjc.12291>
- Landers, M., & Sznycer, D. (2022). The evolution of shame and its display. *Evolutionary Human Sciences*, 4, Article e45. <https://doi.org/10.1017/ehs.2022.43>
- Landers, M., Sznycer, D., & Al-Shawaf, L. (in press). Shame. In L. Al-Shawaf & T. K. Shackelford (Eds.), *The Oxford handbook of evolution and the emotions*. Oxford University Press. <https://psyarxiv.com/nh2yj/>
- Leach, C. W., & Cidam, A. (2015). When is shame linked to constructive approach orientation? A meta-analysis. *Journal of Personality and Social Psychology*, 109(6), 983–1002. <https://doi.org/10.1037/pspa0000037>
- Leary, M. R., Koch, E. J., & Hechenbleikner, N. R. (2001). Emotional responses to interpersonal rejection. In M. R. Leary (Ed.), *Interpersonal rejection* (pp. 145–166). Oxford University Press.
- Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an interpersonal monitor: The sociometer hypothesis. *Journal of Personality and Social Psychology*, 68(3), 518–530. <https://doi.org/10.1037/0022-3514.68.3.518>
- Leroux, A., Héту, S., & Sznycer, D. (2023). The shame system operates with high precision. *Evolutionary Psychology*, 21(3), Article 14747049231203394. <https://doi.org/10.1177/14747049231203394>
- Martens, J. P., Tracy, J. L., & Shariff, A. F. (2012). Status signals: Adaptive benefits of displaying and observing the nonverbal expressions of pride and shame. *Cognition & Emotion*, 26(3), 390–406. <https://doi.org/10.1080/02699931.2011.645281>
- McCullough, M. (2008). *Beyond revenge: The evolution of the forgiveness instinct*. John Wiley & Sons.
- McGraw, K. M. (1987). Guilt following transgression: An attribution of responsibility approach. *Journal of Personality and Social Psychology*, 53(2), 247–256. <https://doi.org/10.1037/0022-3514.53.2.247>
- Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD) psychology: Measuring and mapping

- scales of cultural and psychological distance. *Psychological Science*, 31(6), 678–701. <https://doi.org/10.1177/0956797620916782>
- Niedenthal, P. M., Tangney, J. P., & Gavanski, I. (1994). “If only I weren’t” versus “If only I hadn’t”: Distinguishing shame and guilt in counterfactual thinking. *Journal of Personality and Social Psychology*, 67(4), 585–595. <https://doi.org/10.1037/0022-3514.67.4.585>
- Ollivier, F., Soubelet, A., Duhem, S., & Thümler, S. (2022). Shame and guilt in the suicidality related to traumatic events: A systematic literature review. *Frontiers in Psychiatry*, 13, Article 951632. <https://doi.org/10.3389/fpsyt.2022.951632>
- Otterbacher, J. R., & Munz, D. C. (1973). State-trait measure of experiential guilt. *Journal of Consulting and Clinical Psychology*, 40(1), 115–121. <https://doi.org/10.1037/h0033979>
- Peristiany, J. G. (1970). *Honour and shame: The values of Mediterranean society*. University of Chicago Press.
- Platt, J. R. (1964). Strong Inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science*, 146(3642), 347–353. <https://doi.org/10.1126/science.146.3642.347>
- Randles, D., & Tracy, J. L. (2013). Nonverbal displays of shame predict relapse and declining health in recovering alcoholics. *Clinical Psychological Science*, 1(2), 149–155. <https://doi.org/10.1177/2167702612470645>
- Robertson, T. E., Sznycer, D., Delton, A. W., Tooby, J., & Cosmides, L. (2018). The true trigger of shame: Social devaluation is sufficient, wrongdoing is unnecessary. *Evolution and Human Behavior*, 39(5), 566–573. <https://doi.org/10.1016/j.evolhumbehav.2018.05.010>
- Robins, R. W., Nofle, E. E., & Tracy, J. L. (2007). Assessing self-conscious emotions: A review of self-report and nonverbal measures. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 443–467). Guilford Press.
- Smith, R. H., Webster, J. M., Parrott, W. G., & Eyre, H. L. (2002). The role of public exposure in moral and nonmoral shame and guilt. *Journal of Personality and Social Psychology*, 83(1), 138–159. <https://doi.org/10.1037/0022-3514.83.1.138>
- Syme, K. L., & Hagen, E. H. (2019). When saying “sorry” isn’t enough: Is some suicidal behavior a costly signal of apology? *Human Nature*, 30(1), 117–141. <https://doi.org/10.1007/s12110-018-9333-3>
- Sznycer, D. (2019). Forms and functions of the self-conscious emotions. *Trends in Cognitive Sciences*, 23(2), 143–157. <https://doi.org/10.1016/j.tics.2018.11.007>
- Sznycer, D., & Cohen, A. S. (2021). Are emotions natural kinds after all? Rethinking the issue of response coherence. *Evolutionary Psychology*, 19(2), Article 147470492110160. <https://doi.org/10.1177/14747049211016009>
- Sznycer, D., & Lukaszewski, A. W. (2019). The emotion–valuation constellation: Multiple emotions are governed by a common grammar of social valuation. *Evolution and Human Behavior*, 40(4), 395–404. <https://doi.org/10.1016/j.evolhumbehav.2019.05.002>
- Sznycer, D., & Patrick, C. (2020). The origins of criminal law. *Nature Human Behaviour*, 4(5), 506–516. <https://doi.org/10.1038/s41562-020-0827-8>
- Sznycer, D., Tooby, J., Cosmides, L., Porat, R., Shalvi, S., & Halperin, E. (2016). Shame closely tracks the threat of devaluation by others, even across cultures. *Proceedings of the National Academy of Sciences*, 113(10), 2625–2630. <https://doi.org/10.1073/pnas.1514699113>
- Sznycer, D., Xygalatas, D., Agey, E., Alami, S., An, X.-F., Ananyeva, K. I., Atkinson, Q. D., Broitman, B. R., Conte, T. J., Flores, C., Fukushima, S., Hitokoto, H., Kharitonov, A. N., Onyishi, C. N., Onyishi, I. E., Romero, P. P., Schrock, J. M., Snodgrass, J. J., Sugiyama, L. S., ... Tooby, J. (2018). Cross-cultural invariances in the architecture of shame. *Proceedings of the National Academy of Sciences*, 115(39), 9702–9707. <https://doi.org/10.1073/pnas.1805016115>
- Tangney, J. P. (1990). Assessing individual differences in proneness to shame and guilt: Development of the Self-Conscious Affect and Attribution Inventory. *Journal of Personality and Social Psychology*, 59(1), 102–111. <https://doi.org/10.1037/0022-3514.59.1.102>
- Tangney, J. P. (1991). Moral affect: The good, the bad, and the ugly. *Journal of Personality and Social Psychology*, 61(4), 598–607. <https://doi.org/10.1037/0022-3514.61.4.598>
- Tangney, J. P., & Dearing, R. L. (2002). *Shame and guilt*. The Guilford Press.
- Tangney, J. P., Dearing, R. L., Wagner, P. E., & Gramzow, R. H. (2000). *The Test of Self-Conscious Affect–3 (TOSCA-3)*. George Mason University.
- Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, 70(6), 1256–1269. <https://doi.org/10.1037/0022-3514.70.6.1256>
- Tangney, J. P., Niedenthal, P. M., Covert, M. V., & Barlow, D. H. (1998). Are shame and guilt related to distinct self-discrepancies? A test of Higgins’s (1987) hypotheses. *Journal of Personality and Social Psychology*, 75(1), 256–268. <https://doi.org/10.1037/0022-3514.75.1.256>
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, 58(1), 345–372. <https://doi.org/10.1146/annurev.psych.56.091103.070145>
- Tangney, J. P., Wagner, P., Fletcher, C., & Gramzow, R. (1992). Shamed into anger? The relation of shame and guilt to anger and self-reported aggression. *Journal of Personality and Social Psychology*, 62(4), 669–675. <https://doi.org/10.1037/0022-3514.62.4.669>
- Tangney, J. P., Wagner, P., & Gramzow, R. (1992). Proneness to shame, proneness to guilt, and psychopathology. *Journal of Abnormal Psychology*, 101(3), 469–478. <https://doi.org/10.1037/0021-843X.101.3.469>
- Tangney, J. P., Wagner, P. E., Hill-Barlow, D., Marschall, D. E., & Gramzow, R. (1996). Relation of shame and guilt to constructive versus destructive responses to anger across the lifespan. *Journal of Personality and Social Psychology*, 70(4), 797–809. <https://doi.org/10.1037/0022-3514.70.4.797>
- Tetlock, P. E., & Manstead, A. S. (1985). Impression management versus intrapsychic explanations in social psychology: A useful dichotomy? *Psychological Review*, 92(1), 59–77. <https://doi.org/10.1037/0033-295X.92.1.59>
- Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 114–137). Guilford.
- Tracy, J. L., & Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry*, 15(2), 103–125. https://doi.org/10.1207/s15327965pli1502_01
- Tracy, J. L., & Robins, R. W. (2006). Appraisal antecedents of shame and guilt: Support for a theoretical model. *Personality and Social Psychology Bulletin*, 32(10), 1339–1351. <https://doi.org/10.1177/0146167206290212>
- Tracy, J. L., Shariff, A. F., & Cheng, J. T. (2010). A naturalist’s view of pride. *Emotion Review*, 2(2), 163–177. <https://doi.org/10.1177/1754073909354627>
- Turner, J. E. (1998). *An investigation of shame reactions, motivation, and achievement in a difficult college course*. The University of Texas at Austin.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4), 548–573. <https://doi.org/10.1037/0033-295X.92.4.548>
- Wicker, F. W., Payne, G. C., & Morgan, R. D. (1983). Participant descriptions of guilt and shame. *Motivation and Emotion*, 7(1), 25–39. <https://doi.org/10.1007/BF00992963>
- Yu, H., Hu, J., Hu, L., & Zhou, X. (2014). The voice of conscience: Neural bases of interpersonal guilt and compensation. *Social Cognitive and Affective Neuroscience*, 9(8), 1150–1158. <https://doi.org/10.1093/scan/nst090>
- Zhu, R., Feng, C., Zhang, S., Mai, X., & Liu, C. (2019). Differentiating guilt and shame in an interpersonal context with univariate activation and multivariate pattern analyses. *NeuroImage*, 186, 476–486. <https://doi.org/10.1016/j.neuroimage.2018.11.012>
- Zhu, R., Xu, Z., Tang, H., Liu, J., Wang, H., An, Y., Mai, X., & Liu, C. (2019). The effect of shame on anger at others: Awareness of the emotion-causing events matters. *Cognition and Emotion*, 33(4), 696–708. <https://doi.org/10.1080/02699931.2018.1489782>

Received July 7, 2023

Revision received September 28, 2023

Accepted October 9, 2023 ■