

Social Psychology

Perpetrators' and Victims' Folk Explanations of Aggressive Behaviors and Desires for Apologies

Randy J. McCarthy^{1a}, Jared P. Wilson¹

¹ Department of Psychology, Northern Illinois University, DeKalb, IL, US

Keywords: folk explanations, victim-perpetrator asymmetry, apologies, aggression

<https://doi.org/10.1525/collabra.84918>

Collabra: Psychology

Vol. 9, Issue 1, 2023

After an aggressive interaction, perpetrators most want to offer apologies when they have unintentionally harmed another person and victims most want to receive an apology when another person intentionally harmed them. Perpetrators and victims also explain aggressive behaviors differently—perpetrators often explain their own aggressive behaviors by referring to beliefs they considered that led to their behaviors (i.e., “belief” explanations), whereas victims explain perpetrators’ behaviors by referring to background factors that do not mention the perpetrators’ mental deliberations (i.e., “causal history explanations”). Putting these ideas together, the current Registered Report had participants recall either a time they intentionally harmed another person or a time when they were intentionally harmed by another person. Participants then rated several characteristics of the recalled behavior, explained why the behavior occurred, and reported their desire for an apology. As predicted, we found that perpetrators who gave “belief” explanations wanted to give an apology much less than participants who gave “causal history explanations.” However, and inconsistent with our predictions, victims’ desire to receive an apology was similar regardless of how they explained the perpetrators’ behaviors. These findings underscore how perpetrators’ explanations can emphasize (or de-emphasize) the deliberateness of their harmful behaviors and how these explanations are related to their desire to make amends.

After an interpersonal transgression, victims often want to receive apologies and perpetrators often want to offer apologies (e.g., Lazare, 2005). However, victims’ and perpetrators’ desires for apologies come in response to different behaviors—victims most want to receive apologies when they have been intentionally harmed and perpetrators most want to offer apologies when they have unintentionally harmed another person (e.g., Leunissen et al., 2013). Relatedly, victims and perpetrators explain aggressive behaviors in self-serving ways (e.g., McCarthy et al., 2022). Within their explanations, victims try to frame the perpetrator’s aggressive behaviors as particularly harsh and unpredictable whereas perpetrators try to emphasize the reasonableness of their aggressive behaviors. Thus, just as past research has shown there are differences in victims’ and perpetrators’ desires for apologies from intentional and unintentional transgressions, there also may be differences in their desires for apologies from aggressive behaviors that are explained as being more or less deliberate.

The current Registered Report replicated and extended research comparing victims’ and perpetrators’ perceptions

of aggressive behaviors. We replicated past research showing that victims perceived an aggressive behavior as more harmful than perpetrators would (e.g., Elshout et al., 2017, 2020; McCarthy & Rivers, 2021), that victims and perpetrators explained the perpetrators’ aggressive behaviors differently (e.g., McCarthy et al., 2022), that victims were less likely than perpetrators to mention an apology when recounting an aggressive behavior (e.g., Leunissen et al., 2013), and that victims and perpetrators had different desires for apologies in response to an aggressive behavior (e.g., Leunissen et al., 2013). Extending this past research, we tested whether there is an association between the way victims and perpetrators explained aggressive behaviors and their desires for apologies.

Victims’ and Perpetrators’ Folk Explanations of Aggressive Behaviors

Victims and perpetrators are concerned with different social motives when perceiving and communicating about their aggressive behaviors (e.g., Baumeister, 1996; Fiske,

^a Correspondence concerning this article should be addressed to Randy J. McCarthy, Department of Psychology, Northern Illinois University, DeKalb, Illinois, 60115. Email: rmccarthy3@niu.edu

2018). As a few examples, they each want to avoid negative consequences such as punishment and reputational damage and may try to get benefits such as sympathy, social support, and restitution. In pursuit of these goals, victims often describe aggressive behaviors as unreasonable, disproportionately harmful, and unpredictable. Perpetrators also want to avoid blame and often describe their aggressive behaviors as sensible, justified, and, perhaps, reluctantly necessary (e.g., Baumeister et al., 1990; Stillwell et al., 2008; Stillwell & Baumeister, 1997). If instigated, perpetrators describe their behavior as a proportionate and reasonable response to the instigation (e.g., Elshout et al., 2017; Ent & Parton, 2020).

Differences between victims' and perpetrators' perceptions are not limited to their ratings of aggressive behaviors (e.g., *how severe was this behavior?*) or the content of their recalled memories of their aggressive behaviors (e.g., *did the victim mention an apology in their recollection?*). Differences between victims and perpetrators also can be seen in the subtleties of the language they use to communicate about their aggressive behaviors (e.g., Schütz & Baumeister, 1999). Recently, researchers have been examining victims' and perpetrators' "folk explanations" of why an aggressive behavior occurred (e.g., McCarthy et al., 2022). Folk explanations are simply the answers people give to the question, "Why did that person behave the way they did?" (e.g., Malle, 2004, 2011). Such explanations not only provide an account of what people believe caused a behavior, when expressed, they are social acts—they can be used to nudge, persuade, mislead, placate, etc. other people and to emphasize (or de-emphasize) information about what caused a behavior. Importantly for the current study, explanations can be used to emphasize the deliberateness of an aggressive behavior.

When explaining intentional behaviors such as aggression, people can use different "modes" of explanations (e.g., Malle, 2004, 2011). They can explain an aggressive behavior by referring to what the perpetrator was aware of and considered before behaving aggressively (i.e., *reason explanations*) or they can give an explanation that refers to other factors that caused the perpetrators' aggressive behavior (i.e., *causal history explanations*).

Reason Explanations

Reasons refer to "the contents of an agent's mental states in light of which and on the grounds of which the agent formed an intention to act" (pp. 492-493, Malle et al., 2007). Importantly then, reasons refer to the things the perpetrator considered and was aware of before deciding to behave aggressively (e.g., Malle, 2011; Malle et al., 2000). For example, when explaining why they yelled at their co-worker, the explainer (the perpetrator in this example) may say "I was upset about a joke he made earlier." Here, the perpetrator's explanation claims that they considered their co-worker's past behavior before deciding to act. Similarly, when explaining why a stranger shoved them and stole their backpack, the explainer (the victim in this example) may say, "he wanted my stuff." Here, the victim's explana-

tion claims that the perpetrator considered their desire to get "stuff" before deciding to steal the backpack.

There are strategic reasons why people may choose to explain a behavior with reason explanations. Following norms of communication (e.g., Grice, 1975), when a reason explanation is given, there is a mutual understanding that the given reason is relevant. Thus, when a reason explanation is given, people assume the perpetrator's thought process is relevant for understanding why an aggressive behavior occurred. Because reasons articulate the factors that a perpetrator considered and was aware of that led to their intention to act (e.g., Malle, 2011), a reason explanation implies the behavior was deliberate and the consequences were foreseeable. For example, perpetrators can try to portray their behaviors as sensible and justified by explaining their reasons (e.g., *I hit him because I was protecting myself*). And, upon hearing a perpetrator's reasons, listeners assume that these reasons are relevant (or at least what the perpetrator wants to be seen as relevant) to how the behavior ought to be understood. Likewise, victims, who often try to portray perpetrators' behaviors as capricious or unjustified, can choose to explain the aggressive behavior in ways that do not refer to the perpetrators' mental deliberations (these causal history explanations are described below).

Types of Reasons

There are many reasons that people can consider when deciding to act. Commonly, these are classified into three different types of reasons: *desires*, *beliefs*, and *valuings* (e.g., Malle, 2004, 2011).

Desires refer to the perpetrator's mental states that can be fulfilled. For example, "he yelled at his girlfriend *to make her cry*" refers to something the man was trying to achieve, or aspire towards, with his behavior. Namely, he *wanted* to make his girlfriend cry.

Beliefs refer to factors that a perpetrator believes to be true that were considered when deciding to behave aggressively. For example, "he swore at the other driver *because he was cut off in traffic*" refers to an event—i.e., getting cut off in traffic—the perpetrator believed was true.

Finally, there are reasons that refer to emotions, attitudes, and other sentiments. These reasons are called *valuings*. *Valuings* are distinct from *desires* because they are not aspirational (i.e., they are not something that can be fulfilled) and they are distinct from *beliefs* because they are not propositional (i.e., they are not true or false). For example, "I spread an untrue rumor about Charlie *because I don't like her*" is a *valuing* reason because the explanation claims the perpetrator considered that they do not like Charlie before they decided to act.

Although we expect perpetrators to provide more reasons than victims overall, we also expect nuances in the types of reasons provided by perpetrators and victims. People who want to portray a behavior as reasonable and logical, such as a perpetrator explaining their own aggressive behavior, could explain their behavior with *beliefs* (e.g., Malle et al., 2000). *Beliefs* are presented as *matters-of-fact*. Thus, explaining a behavior with a *belief* emphasizes factors that the perpetrator believes to be true and presents

(and implies) the perpetrator's rational deliberation leading up to their decision to act (see Malle et al., 2000, Experiment 5 for how people use reason explanations to convey rationality). Indeed, people generally explain their own behaviors with beliefs more than others do (e.g., Malle et al., 2007), and, more specific to aggression, perpetrators explain their own aggressive behaviors with beliefs more than victims do (e.g., McCarthy et al., 2022). People who want to de-emphasize a perpetrator's rationality can present a reason explanation that claims the perpetrator's behavior was due to a subjective preference (such as desire reasons or valuing reasons) or they can provide an explanation that does not refer to a perpetrator's mental deliberations (described below).

Mental State Markers

Another notable feature of reason explanations is whether they contain a mental state verb such as "believed," "wanted," "knew," "needed," "tried," "felt," etc. Within the folk explanation framework, including a mental state verb into an explanation is called "marking" an explanation (e.g., Malle, 2004). For example, when explaining why "a man waved a knife," a victim could offer an unmarked belief such as "to intimidate me" or a marked belief such as "he *wanted* to intimidate me." Although the stated reason is the same, mental markers indicate that an actor possesses a thought (e.g., Rosenthal, 2002). Because desires and valuing are understood to be subjectively held by the person whose behavior is being explained, the possession of a thought that is indicated by mental markers are most meaningful when they are used with belief reasons (e.g., Malle et al., 2000).

Unmarked beliefs are stated as matters-of-fact and connote agreement with the stated beliefs (e.g., Malle, 2004; Malle et al., 2000). For example, an unmarked belief explanation such as "Lucas swore at Tim because Tim made an offensive gesture" presents the offensiveness of Tim's gesture as a given rather than Lucas's subjective interpretation of the gesture. Consequently, the unmarked belief explanation implies that the explainer agrees Tim's behavior was offensive. In comparison, a marked version of the explanation would be "Lucas swore at Tim because *he thought* Tim made an offensive gesture." Although the given reason is the same in both explanations, the implications are different: Adding a mental marker (i.e., "he thought...") singles out Lucas as the person who holds this belief and is non-committal about whether the explainer shares that belief. Thus, mental markers are a way to communicate the distinction between a statement about reality and a statement about a person's belief about reality—a distinction that adults typically understand (e.g., Wellman et al., 2001; however, see Birch & Bloom, 2007 about how adults' false-belief reasoning is not unbiased). Because mental state markers can imply agnosticism about whether the explainer agrees with the content of the belief, victims are more likely to provide marked belief explanations than perpetrators (e.g., McCarthy et al., 2022).

Causal History Explanations

Causal history explanations refer to factors that are believed to have caused an actor's reason for behaving intentionally, but do not explicitly refer to their reasons (e.g., Malle, 2011; Malle et al., 2007). Common causal history explanations for aggressive behaviors refer to personality traits (e.g., "*he is a jerk.*"), temporary mental states (e.g., "*I was drunk*"), and cultural factors (e.g., "*fighting is how disagreements are solved around here*"). For example, the explanation "he smashed his girlfriend's car window because he was drunk" declares that the perpetrator's drunkenness was a causal factor in their decision to smash the car window, but this explanation does not declare that the perpetrator considered the fact they were drunk and, in light of that consideration, decided to smash the car window. Importantly then, causal history explanations refer to factors that purportedly caused the behavior but do not presume that the actor considered, or was even aware of, these factors.

Providing causal history explanations can be used to imply that the relevant factors for understanding the behaviors are not the subjective mental deliberations that immediately preceded the aggressive behaviors but would be other background factors such as perpetrators' personalities or temporary mental states (e.g., Malle, 2004). If victims want to frame a perpetrator's aggressive behavior as capricious or unpredictable, they can choose to provide an explanation that emphasizes factors other than what the perpetrator considered before behaving aggressively. Providing a causal history explanation does not deny that a perpetrator had some mental deliberations before behaving aggressively, but these explanations shift the attention away from those deliberations and onto other factors.

Victims' and Perpetrators' Desires for Apologies

Sincere apologies convey many meanings: They can acknowledge that a transgression has taken place, express guilt and contrition, pledge that the transgression will not occur again, communicate that the apologizer values the relationship, etc. (e.g., Forster et al., 2021; Lazare, 2005). As with perceptions of the severity of an aggressive behavior, there are reliable differences in perpetrators' desires to offer apologies and victims' desires to receive apologies. When considering what evokes perpetrators' and victims' desires for an apology, a key feature is the extent to which the behavior was (perceived to be) deliberate and that the harmful consequences were (perceived to be) foreseen by the perpetrator (e.g., Leunissen et al., 2013).

A perpetrator who offers an apology for a deliberately harmful behavior acknowledges that they knowingly harmed the victim and, thus, "owe" the victim an apology (e.g., Lazare, 2005). When faced with managing the aftermath of an intentionally harmful behavior, perpetrators often resort to strategies such as denying the harmful consequences of the behavior (e.g., Dersley & Wootton, 2000), claiming the behavior was unintentional (e.g., Adams & Inesi, 2016, 2016), or providing justifications or excuses (e.g., Darley & Pittman, 2003; Malle et al., 2014), which often in-

volves explicating the reasons for a behavior so others can see the basis of the aggressive behavior. Perpetrators' aim to portray their aggressive behavior as a reasonable decision rather than an indefensible transgression that requires an apology. Somewhat ironically then, a perpetrator who deliberately harms another person likely has highly salient reasons for their behavior and, thus, view their behavior as justified and not requiring an apology (e.g., Leunissen et al., 2013; Tsang, 2002; see also McGraw, 1987).

In comparison, victims interpret apologies as an acknowledgment by the perpetrator that a transgression has taken place, as symbolic compensation for a transgression, and a pledge by the perpetrator that the transgression will not occur again (e.g., Lazare, 2005). Unsurprisingly then, victims' desire for apologies are strongest when perpetrators (are believed to have) behaved deliberately because these are the behaviors where the perpetrator knowingly and willfully behaved in a way that caused harm to the victim. Indeed, victims more strongly demand compensation and desire retribution in response to intentionally harmful behaviors (e.g., Darley & Pittman, 2003). In contrast, victims' desire for apologies are less strong when perpetrators behave unintentionally, such as when one person accidentally harms another person (e.g., Leunissen et al., 2013).

Hypotheses

Putting together the differences in how victims and perpetrators explain aggressive behaviors and how victims and perpetrators desire apologies in response to different behaviors, we formed a few novel hypotheses.

First, belief reasons convey (what the perpetrator believes is) the factual basis for why a perpetrator decided to act. Thus, perpetrators who provide belief reasons for their aggressive behaviors are likely to view their behavior as justified, which, in turn, should be associated with less of a desire to offer an apology. In contrast, belief explanations presume the perpetrator considered, and was aware of, the factors that led to their decision to behave aggressively. Thus, victims who provide belief reasons for a perpetrator's aggressive behavior would see the perpetrator as acting most deliberately, which should be associated with a strong desire to receive an apology (e.g., Leunissen et al., 2013).

Second, when causal history explanations are given, aggressive behaviors are explained by factors other than the perpetrators' thought processes leading up to their aggressive behaviors and there is no presumption that the perpetrators were aware of these factors before deciding to act. Thus, perpetrators who explain their aggressive behaviors with causal history explanations are choosing to offer explanations that de-emphasize their mental deliberations and, hence, should have a stronger desire to offer an apology. In comparison, victims who explain perpetrators' aggressive behaviors with causal history explanations also are choosing to explain the behavior with factors that do not include the perpetrator's thought process, nor do these explanations presume the perpetrator was aware of the provided causal factors, which may imply the behavior is seen as less deliberate. Consequently, victims who give causal

history explanations should have less desire to receive an apology (e.g., Leunissen et al., 2013).

To test these hypotheses, we asked participants to recall either a time another person intentionally harmed them (a "victim behavior") or a time they intentionally harmed another person (a "perpetrator behavior"). Participants then explained why the perpetrator behaved aggressively, rated their perceptions of the behavior, reported their desire for an apology, and whether an apology was actually offered. Thus, in addition to our novel hypotheses, our study also contains several built-in replications of previously published effects. The specific effects that will be used to evaluate our hypotheses are listed in [Table 1](#).

Hypothesis 1. Those who recalled a victim behavior will rate the behaviors as more harmful and less justified than those who recalled a perpetrator behavior (e.g., Elshout et al., 2017; McCarthy & Rivers, 2021). Further, we expect that participants who recalled a victim behavior will express a stronger desire for an apology than participants who recalled a perpetrator behavior (e.g., Leunissen et al., 2013).

Hypothesis 2. Participants will use more causal history explanations to explain their victim behaviors than their perpetrator behaviors. On the other hand, participants will use more reason explanations to explain their perpetrator behaviors than their victim behaviors. Within reason explanations, those who recalled a perpetrator behavior will use more belief and unmarked reasons than those who recalled a victim behavior (e.g., McCarthy et al., 2022).

Hypothesis 3. We also asked participants if the perpetrator offered an apology. As in Leunissen et al. (2013), we expected participants to report more apologies for those who recalled a perpetrator behavior than for those who recalled a victim behavior. Further, we expected that a stronger desire for an apology would predict whether an apology was reported. Finally, and also consistent with Leunissen et al. (2013), we expected participants' role (i.e., a victim or perpetrator) to interact with their desire for an apology such that a desire for an apology more strongly predicts perpetrators' reporting an apology than victims' reporting receiving an apology.

Hypothesis 4. The use of belief reasons should emphasize the perpetrator's mental deliberations and subjective awareness of the factors that led to their aggressive behavior. This would have divergent effects for those who recalled perpetrator behaviors and those who recalled victim behaviors. For those who recalled a perpetrator behavior, the use of beliefs would be associated with less of a desire to offer an apology. For those who recalled a victim behavior, the use of beliefs would be associated with a greater desire for receiving an apology.

In contrast, a perpetrator's mental deliberations are de-emphasized when causal history explanations are given. This de-emphasis on the perpetrator's mental deliberations would have divergent effects for perpetrators and victims. For those who recalled a perpetrator behavior, the use of causal history explanations would be associated with a greater desire for an apology. For those who recalled a victim behavior, the use of causal history explanations would be associated with less of a desire for an apology.

Table 1. Summary of hypotheses and effects that will be used to test those hypotheses

Hypotheses and Outcome	Predicted effect to test our hypotheses	Was the hypothesis supported?
<i>Hypothesis #1 (replication)</i>		
Severity	Type of Memory (Perpetrator vs. Victim) × Rater (Self vs. Other) interaction	No
Intentionally Harmful	Main effect of Type of Memory (Victims > Perpetrators)	Yes
Actually Harmful	Main effect of Type of Memory (Victims > Perpetrators)	No
Justified	Main effect of Type of Memory (Perpetrators > Victims)	Yes
Desire for apology	Main effect of Type of Memory (Victims > Perpetrators)	No
<i>Hypothesis #2 (replication)</i>		
Reason Explanations	Main effect of Type of Memory (Perpetrators > Victims)	Yes
Belief Reasons	Main effect of Type of Memory (Perpetrators > Victims)	Yes
Desire Reasons	Main effect of Type of Memory (Victims > Perpetrators)	No
Marked Beliefs	Main effect of Type of Memory (Victims > Perpetrators)	Yes
Causal History Explanations	Main effect of Type of Memory (Victims > Perpetrators)	Yes
<i>Hypothesis #3 (replication)</i>		
An apology being issued	Main effect of Type of Memory (Perpetrators > Victims) Main effect of Desire for Apology Type of Memory (Perpetrator vs. Victim) × Desire for Apology interaction	Yes Yes Yes
<i>Hypothesis #4 (novel hypothesis)</i>		
Folk Explanations predict desires for apologies	Type of Memory (perpetrator vs. victim) × Type of Explanation (belief vs. causal history explanation) interaction	Partially supported

Note. Hypotheses #1-#3 are close replications of previously-published effects. The rightmost column describes whether the results of the current study supported the hypothesis or not.

Method

Sample Size Determination

Given there are no directly comparable studies on which to base our power analysis for the novel hypotheses (i.e., Hypothesis #4), we planned our proposed study to detect the smallest effect size of interest. The primary hypothesis tests were two simple effects, which would be used to interpret the expected Type of Memory (perpetrator vs. victim) × Type of Explanation (belief vs. causal history explanation) interaction. The first simple effect would examine desires for apologies for those who provided a belief explanation. The second simple effect would examine desires for apologies for those who provided a causal history explanation. Each of these simple effects would compare participants who recalled a perpetrator memory and those who provided a victim memory.

Among those who provided a belief explanation, and separately among those who provided a causal history explanation, we wanted to detect differences between perpetrators' and victims' desires for apologies as small as $d = 0.25$. Using a Type 1 error rate of 5% (i.e., $\alpha = .05$) and statistical power of 90% (i.e., $1 - \beta = .90$) to detect an effect of $d = 0.25$ requires 676 total participants. We anticipate about 20% of participants to be excluded from our primary analysis because they will provide both a belief explanation and a causal history explanation (described in greater detail be-

low), which means we would require (676×1.2) 812 participants. To account for attrition, we collected at least 1,000 "CloudResearch" approved participants.

Participants

A total of 1,178 participants visited the study link. Participants were excluded if they did not provide a valid memory ($n = 215$), a valid explanation ($n = 259$), or did not provide all the ratings of their memories ($n = 202$). We also included two attention check items (described below). Participants who did not pass these attention check items also were excluded ($n = 28$). These criteria resulted in the omission of 459 participants, which left a final sample of 719 participants.

The average age of participants in the final sample was 39.83 ($SD = 12.10$) years. Participants provided their gender and ethnicity in open-ended textboxes that were coded by the research team. In the final sample, 370 (51.5%) participants were men, 341 (47.4%) were women, 6 (0.9%) were non-binary, and 1 (0.1%) did not provide gender information. In the final sample, 538 (74.8%) participants were White, 62 (8.6%) were Asian-American, 54 (7.5%) were African-American/Black, 38 (5.3%) were Hispanic/Latinx, 21 (2.9%) were multi-ethnic, 6 (0.8%) were another ethnicity or did not report a usable response.

Procedures

Participants were recruited through [Amazon.com](https://www.amazon.com)'s Mechanical Turk website. After consenting to participate, participants wrote "I will answer open-ended questions" into a text box. This commitment exercise has been used in previous research to minimize attrition (e.g., Zhou & Fishbach, 2016).

Participants were then randomly assigned to either recall a time when they "intentionally harmed or offended another person" (i.e., a "perpetrator" memory) or a time when another person "intentionally harmed or offended you" (a "victim" memory).

Participants randomly assigned to recall a perpetrator memory were then asked to "Describe the harmful behavior in complete sentences." And were then asked "Did you know the person who you harmed prior to this incident? If so, what was your relationship (e.g., they were my significant other, we were co-workers, we were friends, etc.)", and "Why did you harm this other person?". Then participants rated how severe, excessive, and extreme they viewed their behavior on a scale with anchors 1 = not at all to 7 = completely (e.g., Elshout et al., 2017). Participants then rated their perpetrator behavior on three items: "I intended to harm or offend the other person", "I actually harmed or offended the other person", and "My behavior was justified" on a scale ranging from 1 = strongly disagree to 7 = strongly agree (e.g., McCarthy & Rivers, 2021). Participants then reported "To what extent did you want to offer an apology to this other person?" using a scale with anchors 1 = not at all to 7 = very much. Finally, participants were asked "Did you offer an apology to the other person?" (e.g., Leunissen et al., 2013).

Participants randomly assigned to recall a victim memory were then asked to "Describe the harmful behavior in complete sentences." And were then asked "Did you know the person who harmed you prior to this incident? If so, what was your relationship (e.g., they were my significant other, we were co-workers, we were friends, etc.)", and "Why did this other person harm you?". Then participants rated how severe, excessive, and extreme they viewed the behavior on a scale with anchors 1 = not at all to 9 = completely (e.g., Elshout et al., 2017). Participants then rated their perpetrator behavior on three items: "The other person intended to harm or offend me", "The other person actually harmed or offended me", and "The other person's behavior was justified" on a scale ranging from 1 = strongly disagree to 7 = strongly agree (e.g., McCarthy & Rivers, 2021). Participants then reported "To what extent did you want to receive an apology from this other person?" using a scale with anchors 1 = not at all to 7 = very much. Finally, they were asked "Did you receive an apology from the other person?" (e.g., Leunissen et al., 2013).

Next, all participants completed two items to screen for attentive responding: "I speak every language in the world" and "I was born before the year 1900." Participants then re-

ported their age, sex, gender (in an open-ended text box), and race/ethnicity (in an open-ended text box). Finally, participants were debriefed and thanked for their time.

Adherence to, and Deviations from, the IPA Version

Our methods and analyses followed our pre-registered plans with one exception: We proposed to have all participants rate how severe, excessive, and extreme they viewed the behavior on a scale with anchors 1 = not at all to 9 = completely. These ratings would have made our data directly comparable to the ratings from Elshout et al. (2017). Participants who recalled a victim memory rated their behaviors on a scale with anchors 1 = not at all to 9 = completely as planned. However, participants who recalled a perpetrator memory rated their recalled behavior on a scale with anchors 1 = not at all to 7 = completely. This error meant that those who recalled a victim memory could possibly give ratings from one to nine and those who recalled a perpetrator memory could only give ratings from one to seven. Thus, we were unable to analyze the severity of their recalled behaviors as we had proposed.¹

Results

Data Preparation

Several steps were taken to prepare our data for analysis.

First, a research assistant coded each memory as being valid or not. Missing memories; memories such as "none," "can't recall," etc.; non-sensical memories; and memories that were clearly written from the incorrect perspective were coded as invalid. Similarly, participants' explanations had to have contained an account of why the perpetrator behaved aggressively. This coding was used to exclude participants who did not provide usable data.

Second, two research assistants coded the explanations for the aggressive behaviors. When coding these open-ended responses, the coders completed training sessions where both authors reviewed the coding manual, jointly coded several cases, had each coder independently complete coding decisions with the authors, etc. Once the coders were confident in the coding process, two coders then independently coded all the explanations for whether they contained a reason explanation or causal history explanation. Inter-rater reliability was computed from their independently-completed coding decisions. Any discrepancies were then resolved with a discussion. After resolving the discrepancies for whether an explanation included a reason, the reason explanations were then independently coded for type of reason and whether belief reasons contained a mental marker. Inter-rater reliability was computed from their independently-completed coding deci-

¹ This error on our survey only affected participants' severity ratings for their recalled behaviors, this error did not affect our independent raters' ratings of the recalled behaviors.

Table 2. Victims' and perpetrators' ratings of their recalled aggressive behaviors

Rating	Perpetrator Memory		Victim Memory		<i>t</i>	<i>df</i>	<i>p</i>
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>			
Severe, Excessive, Extreme							
Self Ratings	0.52	0.28	0.76	0.24	-12.13	716.81	<.001
Independent Raters	3.47	1.60	5.30	2.12	-12.85	605.37	<.001
Intentionally harmful	5.26	1.59	6.16	1.13	-8.79	695.33	<.001
Actually harmful	5.38	1.32	5.79	1.45	-3.95	671.84	<.001
Justified	3.95	2.00	1.58	1.14	19.87	630.91	<.001
Desire for apology	3.48	2.34	3.72	2.53	-1.33	677.47	.18

Note. Participants rated their recalled behaviors on several traits. Due to an error in our survey, participants who recalled a perpetrator memory and participants who recalled a victim memory rated how "severe," "excessive," and "extreme" their behaviors were on scales with different numbers of anchors. These ratings were recoded by subtracting 1 from the ratings and dividing by the range of the scales (within their respective conditions). Thus, these recoded ratings ranged from 0 to 1. Two independent researchers also rated how "severe," "excessive," and "extreme" each behavior was on a scale ranging from 1 = *not at all* to 9 = *completely*. Ratings for "intentionally harmful," "actually harmful," and "justified" were made on scales ranging from 1 = *strongly disagree* to 7 = *strongly agree*. Ratings for "desire for an apology" were made on a scale ranging from 1 = *not at all* to 7 = *very much*.

sions (the inter-rater reliabilities are presented in Table 3). Any discrepancies were resolved with a discussion.

Third, two student research assistants who were unaware of the study's hypotheses rated how *severe*, *excessive*, and *extreme* they view each valid aggressive behavior. The means of these ratings were averaged across both raters (Coder A $\alpha = .97$, Coder B $\alpha = .95$, $r = .96$) to create a variable of how severe other people see the recalled behavior. These ratings were used as a "control" variable in several analyses (see Elshout et al., 2017; McCarthy et al., 2022).

We took several precautions to minimize bias during the coding process. First, neither author was a coder (although we both participated in the discussions to resolve discrepancies). Second, when coding the folk explanations and rating the severity of the behaviors, information about participants' responses, participants' demographic information, and the other coders' coding/rating decisions were hidden from the coders.

Hypothesis 1: Replication of Quantitative Ratings of Aggressive Behaviors

Participants' ratings of their recalled behaviors, the two independent raters' ratings of the severity of the behaviors, and comparison between ratings of the perpetrator memories and victim memories are shown in Table 2.

As described above, we erroneously asked those who recalled victim memories and those who recalled perpetrator memories to give their severity ratings on scales with different numbers of points, therefore, this analysis departs from what is described in our pre-data collection manuscript. For this analysis, we subtracted 1 from the mean ratings of severe, excessive, and extreme for each participant (thus, making the minimum value 0) and divided these recoded values by the range of the scale (i.e., for those who recalled victim memories, we divided their ratings by 8, and for those who recalled perpetrator memories, we divided their ratings by 6). Thus, the recoded ratings for participants in both conditions then ranged from 0 to 1. We then conducted a multi-level regression analysis with these recoded severity ratings nested within participants. We included memory type (0 = perpetrator memory vs. 1 = vic-

tim memory), rater (0 = rater vs. 1 = self-ratings), and the Memory Type \times Rater interaction. Intercepts were allowed to vary randomly. There were two main effects: Participants who recalled victim memories rated the behaviors as more severe than participants who recalled perpetrator memories, $b = 0.28$, $SE = 0.02$, $p < .001$, and participants rated the behaviors as more severe than the raters did, $b = 0.21$, $SE = 0.01$, $p < .001$. The interaction was not significant, $b = 0.01$, $SE = 0.01$, $p = .53$. We expected to observe an interaction similar to that reported in Elshout et al. (2017), thus, our hypothesis was not supported.

For the remaining ratings, we ran a series of linear regression analyses with the main effect of independent raters' (mean-centered) severity ratings, the type of memory a participant recalled (0 = perpetrator memory vs. 1 = victim memory), and the (mean-centered) Severity \times Type of Memory interaction as predictors. Thus, the "type of memory" main effect can be interpreted as the difference between those who recalled a perpetrator memory and those who recalled a victim memory for behaviors that were rated as being "average" in severity (i.e., the mean level of severity as rated by the independent raters).

When predicting whether the recalled aggressive behavior was intentionally harmful, we observed a main effect for severity ratings, $b = 0.25$, $SE = 0.04$, $p < .001$. As the behavior was more severe, participants rated the behaviors as more intentionally harmful. We also observed a main effect for the type of memory, $b = 0.51$, $SE = 0.11$, $p < .001$. As predicted, for a behavior that has an "average" level of severity, those who recalled victim memories rated their recalled behavior as 0.51 points more intentionally harmful than those who recalled perpetrator memories. There was not a Severity \times Type of Memory interaction, $b = -0.07$, $SE = 0.06$, $p = .24$.

When predicting whether the recalled aggressive behavior was actually harmful, we observed a main effect for severity ratings, $b = 0.20$, $SE = 0.04$, $p < .001$. As behaviors were rated as being more severe, participants rated their behaviors as more actually harmful. However, going against our predictions, there was not a main effect for the type of memory, $b = 0.07$, $SE = 0.11$, $p = .53$. There also was not a

Table 3. Victims' and perpetrators' folk explanations of their recalled aggressive behaviors

	Kappa (% agreement)	Victim memories		Perpetrator memories		χ^2	<i>p</i>
		Example	Percent	Example	Percent		
Reasons	85.3(96.9%)		80.0%		93.8%	29.94	<.001
Beliefs	95.9(98.0%)	"felt I wasn't being considerate enough of her"	44.3%	"The other person insulted me"	65.9%	28.75	<.001
Desires	98.3(99.2%)	"She wanted to make me look bad to make herself look better"	39.5%	"I wanted them to feel hurt"	34.1%	1.70	.19
Valuings	97.5(98.8%)	"She must have been pissed off that we canceled last minute"	47.2%	"I was irritated with her and her past behavior"	45.8%	1.53	.47
Causal history explanations	84.6(94.3%)	"He was drunk and angry"	36.7%	"I was angry and irritated from nicotine withdrawal"	17.0%	34.99	<.001

Note. Because participants sometimes give several explanations for a given behavior, these can add up to more than 100%. The analyses in this table are not controlling for the severity of the behaviors (as judged by the independent raters).

Severity \times Type of Memory interaction, $b = -0.03$, $SE = 0.05$, $p = .59$.

When predicting whether the recalled aggressive behavior was justified, we observed main effects for severity ratings, $b = -0.26$, $SE = 0.05$, $p < .001$, and type of memory, $b = -2.07$, $SE = 0.14$, $p < .001$. More severe behaviors were rated as less justified and, as predicted, behaviors judged from the victim perspective were rated as less justified. These main effects were qualified by a significant Severity \times Type of Memory interaction, $b = 0.17$, $SE = 0.07$, $p = .01$. Thus, the relationship between severity and justification was significantly different for those who recalled perpetrator memories and victim memories. As the behaviors were rated as more severe, those who recalled victim memories rated the behavior as less justified, $b = -0.09$, $SE = 0.03$, $p = .002$. However, as the behaviors were rated as more severe, those who recalled a perpetrator memory also rated their behavior as less justified, albeit to a greater extent than those who recalled a victim memory, $b = -0.26$, $SE = 0.06$, $p < .001$. An inspection of the distribution of the ratings suggests this interaction may be driven by a floor effect for those who recalled victim memories. In fact, for those who recalled victim memories, the modal rating of how justified the behavior was 1!

Finally, when predicting the desire for an apology, we observed a main effect for severity ratings, $b = 0.15$, $SE = 0.08$, $p = .05$. However, and going against our predictions, there was not a main effect for the type of memory, $b = -0.11$, $SE = 0.20$, $p = .59$. Finally, there was not a Severity \times Type of Memory interaction, $b = -0.15$, $SE = 0.08$, $p = .13$.

Hypothesis 2: Replication of Folk Explanations of Aggressive Behaviors

We tested whether victims and perpetrators gave different types of explanations for the perpetrators' behavior with a series of logistic regression analyses while accounting for differences in the severity of the behaviors. To test these hypotheses, we predicted whether a type of explanation was given (coded 0 = no and 1 = yes) with a model that included the main effect of independent raters' (mean-centered) ratings of the severity of the behaviors, the main effect of the type of behavior participants recalled (0 = perpetrator memory vs. 1 = victim memory), and the (mean-centered) Severity \times Type of Memory interaction as predictors.

As predicted, for a behavior that has an "average" level of severity, the odds that a reason explanation was given was less among those who recalled victim memories than among those who recalled perpetrator memories, $b = -1.22$, $z = -4.33$, $p < .001$, $OR = 0.29$. When predicting whether a reason explanation was given, there was no main effect of severity, $b = -0.03$, $z = -0.21$, $p = .83$, $OR = 0.97$, and there was no Severity \times Type of Memory interaction, $b = -0.05$, $z = -0.37$, $p = .71$, $OR = 0.95$. Also as predicted, for a behavior that has an "average" level of severity, the odds that a causal history explanation was given was greater among those who recalled victim memories than among those who recalled perpetrator memories, $b = 0.84$, $z = 4.42$, $p < .001$, $OR = 2.33$. When predicting whether a causal history explanation

was given, there was a main effect of severity, $b = 0.22$, $z = 2.63$, $p = .009$, $OR = 1.24$, such that more severe behaviors were more likely to be explained with causal history explanations. Finally, there was no Severity \times Type of Memory interaction, $b = -0.16$, $z = -1.66$, $p = .10$, $OR = 0.85$.

When looking at the types of reason explanations, for a behavior that has an "average" level of severity, as predicted, the odds that a belief explanation was given was less among those who recalled victim memory than among those who recalled a perpetrator memory, $b = -0.90$, $z = -4.91$, $p < .001$, $OR = 0.41$. Going against our predictions, those who recalled victim memories gave a similar amount of desire explanations, $b = 0.17$, $z = 0.91$, $p = .36$, $OR = 1.19$, and valuing explanations, $b = -0.18$, $z = -0.004$, $p = .98$, $OR = 1.00$, as those who recalled perpetrator memories did. As predicted, for a behavior that has an "average" level of severity, the odds that a mental markers was given for their belief explanations was greater among those who recalled a victim memory than among those who recalled a perpetrator memory, $b = 0.87$, $z = 3.41$, $p < .001$, $OR = 2.38$. For the types of reasons and the use of marked beliefs, there were no main effects of severity nor were there any Severity \times Type of Memory interactions.

Hypothesis 3: Replication of Desires for Apologies and Apology Behavior

Participants reported whether an apology was offered in response to their recalled aggressive behavior. In all, 37.0% of those who recalled a perpetrator memory reported offering an apology and 15.5% of those who recalled a victim memory reported receiving an apology. When not considering the severity of the behavior, the percentage of victims and perpetrators who reported an apology was significantly different, $X^2 (df = 1) = 40.92$, $p < .001$, $OR = 0.31$.

We used logistic regression to test whether participants' desire for an apology predicted whether an apology was actually given while accounting for differences in the severity of the behaviors. To test our hypotheses, we entered the type of memory (0 = perpetrator memory vs. 1 = victim memory), the (mean-centered) desire for an apology, and the Type of Memory \times Desire for Apology interaction as predictors of whether an apology was given (coded 0 = no apology given and 1 = an apology given). We also included the main effect of independent raters' (mean-centered) ratings of the severity of the behaviors as a control variable.

The severity of the behavior was unassociated with whether an apology was given, $b = -0.08$, $z = -1.36$, $p = .18$, $OR = 0.92$. As predicted, we observed a main effect of the type of memory, $b = -1.01$, $z = -3.73$, $p < .001$, $OR = 0.36$. For a behavior that was seen as average in severity and had an average desire for an apology, the odds that an apology was reported was greater for participants who recalled a perpetrator memory than for those who recalled a victim memory. We also observed a main effect for the desire for an apology, $b = 0.90$, $z = 11.08$, $p < .001$, $OR = 2.46$. As the desire for an apology increased, the odds of an apology being reported also increased. Finally, as predicted, we observed a Type of Memory \times Desire for Apology interaction, $b = -0.55$, $z = -5.05$, $p < .001$, $OR = 0.58$.

Table 4. Predicting desires for apology from victims' and perpetrators' folk explanations of their recalled aggressive behaviors

Coefficient	<i>b</i>	<i>SE</i>	<i>t</i>	<i>p</i>
(intercept)	3.01	0.18	16.83	<.001
Severity (mean-centered)	0.18	0.10	1.86	.06
Memory Type	1.04	0.32	3.27	.001
Explanation Type	2.08	0.42	4.96	<.001
Severity × Memory Type	-0.20	0.15	-1.34	.18
Memory Type × Explanation Type	-0.41	0.25	-1.64	.10
Severity × Memory Type × Explanation Type	-2.54	0.57	-4.48	<.001
Severity × Memory Type × Explanation Type	0.57	0.30	1.92	.06

Note. Regression analysis predicting participants' desires for apology from the severity of the behaviors (as judged by the independent raters), whether participants recalled a perpetrator memory or a victim memory, and whether participants explained the perpetrator's behavior with a belief reason or a causal history explanation.

To interpret the nature of this Type of Memory × Desire for Apology interaction, we computed the simple slopes separately for those who recalled a victim memory and for those who recalled a perpetrator memory. Among those who recalled a victim memory, for behaviors that were “average” in severity, there was a significant, although relatively weaker, relationship between the desire for an apology and whether an apology was received, $b = 0.36$, $z = 4.91$, $p < .001$, $OR = 1.43$. In comparison, among those who recalled a perpetrator memory, for behaviors that were “average” in severity, there was a significant and relatively stronger relationship between the desire for an apology and whether an apology was given, $b = 0.90$, $z = 11.00$, $p < .001$, $OR = 2.46$. Although both of these effects are significantly positive, the “desire for apology-apology” relationship for those who recalled a perpetrator memory was stronger than for those who recalled a victim memory.

Hypothesis 4: Association Between Folk Explanations and Desires for Apologies

To prepare the data to test our hypothesis #4, we needed groups of participants who provided either a belief reason explanation only or a causal history explanation only. Thus, we had to exclude participants who provided both a belief reason explanation and a causal history explanation. Participants who provided a belief reason and did not provide a causal history explanation, were our “belief reason” group (this group included participants who gave a belief explanation, did not give a causal history explanation, and they could have also given a desire and/or a valuing explanation). Participants who provided a causal history explanation and did not provide a belief explanation, were our “causal history explanation” group (this group included participants who gave a causal history explanation, did not give a belief explanation, and they could have also given a desire and/or a valuing explanation). These exclusions left a sample of 260 participants who recalled a perpetrator memory: 219 gave a belief reason (and not a causal history explanation) and 41 gave a causal history explanation (and not a belief reason). These exclusions also left us with a sample of 187 participants who recalled a victim memory: 93 gave a belief reason (and not a causal history explanation)

and 94 gave a causal history explanation (and not a belief reason).

With these groups created, we then predicted participants' desire for an apology from their type of memory (0 = perpetrator memory vs. 1 = victim memory), the type of explanation (0 = Belief vs. 1 = Causal History Explanation), and the Type of Memory × Type of Explanation interaction. We also included the main effect of independent raters' (mean-centered) ratings of the severity of the behaviors as a control variable.

There was a significant Type of Memory × Type of Explanation interaction, $b = -2.54$, $z = -4.48$, $p < .001$. As predicted, for participants who recalled a perpetrator memory, there was more of a desire for an apology for behaviors that were explained with causal history explanations ($M = 5.21$, $SD = 2.33$) than for behaviors explained with beliefs ($M = 2.86$, $SD = 2.17$), $b = 2.08$, $z = 5.30$, $p < .001$. However, and not predicted, for participants who recalled a victim memory, there was no difference in the desire for an apology for behaviors that were explained with causal history explanations ($M = 3.77$, $SD = 2.58$) and behaviors explained with beliefs ($M = 4.03$, $SD = 2.50$), $b = -0.45$, $z = -1.11$, $p = .27$ (see Table 4 and Figure 1).

Exploratory Analyses

To further test our hypotheses, we conducted the same analyses as in Hypothesis 4 except we predicted whether an apology was actually given rather than simply the desire for an apology. Because whether an apology was given is dichotomous, we used logistic regression.

There was a significant Type of Memory × Type of Explanation interaction, $b = -2.43$, $z = 3.05$, $p = .002$, $OR = 11.36$. For participants who recalled a perpetrator memory, the odds that an apology was reported was greater for behaviors that were explained with causal history explanations (61.9%) than for behaviors explained with beliefs (25.1%), $b = -1.52$, $z = -2.67$, $p = .008$, $OR = 0.22$. The opposite pattern emerged for participants who recalled a victim memory: The odds that an apology was reported was greater for behaviors explained with beliefs (21.5%) than for behaviors explained with causal history explanations (9.6%), $b = 1.34$, $z = 2.00$, $p = .05$, $OR = 3.84$.

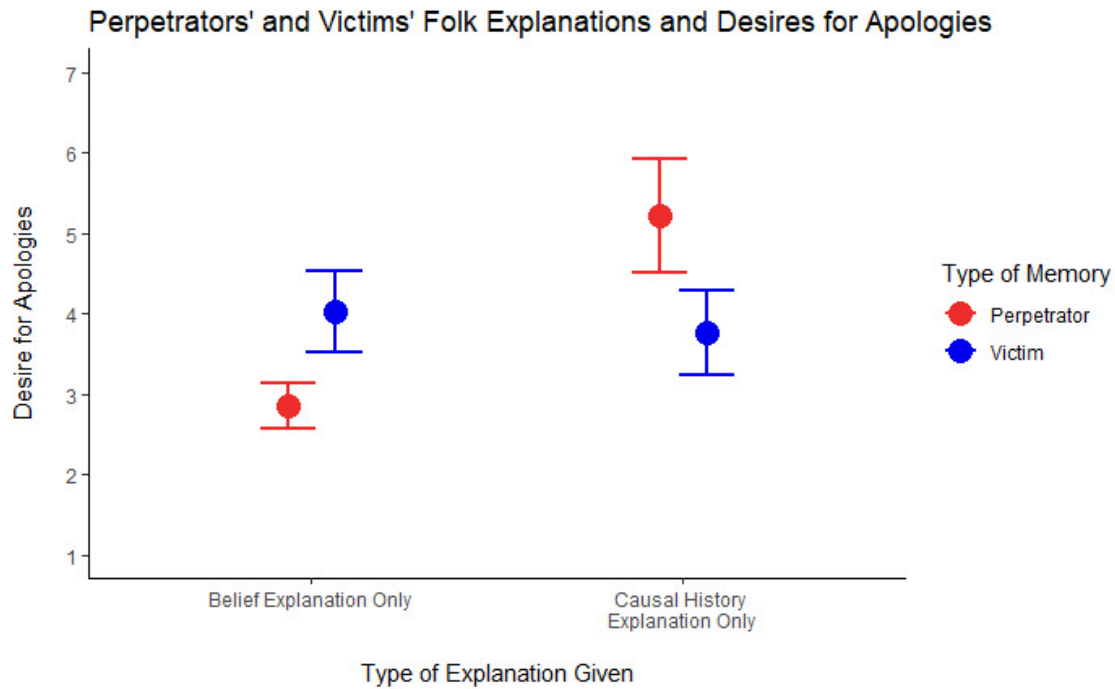


Figure 1. Desires for apology from victims' and perpetrators' folk explanations of their recalled aggressive behaviors

Note. Means of participants' desires for apology from the severity of the behaviors (as judged by the independent raters), whether participants recalled a perpetrator memory or a victim memory, and whether participants explained the perpetrator's behavior with a belief reason or a causal history explanation. Error bars represent 95% confidence intervals.

Discussion

In the current Registered Report, participants recalled either a time another person intentionally harmed them (i.e., a "victim" memory) or a time they intentionally harmed another person (i.e., a "perpetrator" memory). Participants then explained why the perpetrator—who was sometimes themselves and sometimes another person—behaved aggressively, rated the behavior on several traits, reported their desire for an apology, and reported whether an apology was actually given. These data allowed us to replicate several previously-published effects (Hypotheses 1-3) and to test novel hypotheses (Hypothesis 4) about how people's explanations for their aggressive behaviors are associated with their desires for apologies.

Hypothesis #1: Victims and Perpetrators Rate Aggressive Behaviors Differently

Our first hypotheses involved replicating previously-published effects on how people rate their recalled victim memories and perpetrator memories (e.g., Elshout et al., 2017; Leunissen et al., 2013; McCarthy et al., 2022; McCarthy & Rivers, 2021). In the current study, and consistent with previous research (e.g., McCarthy & Rivers, 2021), those who recalled victim memories viewed their recalled behaviors as more intentionally harmful and less justified than those who recalled a perpetrator memory did. However, and unlike past research, there was no difference in how actually harmful they viewed their recalled behaviors (e.g., McCarthy et al., 2022) and their desire for an apology

after an intentional transgression (e.g., Leunissen et al., 2013).

The current results are important for two reasons. First, given that these differences in how intentionally harmful and severe people rate their perpetrator memories and victim memories have been found across several samples by several research teams, now including findings from within a Registered Report, researchers should have high confidence in these effects. Second, participants' ratings to these face valid items demonstrates, in a straightforward manner, how influential a person's role is within an aggressive interaction: Perpetrators and victims simply perceive aggressive behaviors differently.

Hypothesis #2: Victims and Perpetrators Explain Aggressive Behaviors Differently

Our second hypotheses involved replicating previously-published effects on how people explain their own behaviors versus the behaviors of others for general behaviors (e.g., Malle et al., 2007) and specifically for aggressive behaviors (e.g., McCarthy et al., 2022). Specifically, after recalling an aggressive interaction, participants in the current study gave an explanation for why the perpetrator behaved aggressively. We coded the "mode" of explanation they gave. One mode (i.e., "reasons") refers to the mental deliberations that led to the perpetrators' behavior. Another mode (i.e., "causal history of reason explanations") refers to background factors that preceded the perpetrators' reasons for behaving aggressively such as their personality, demographic factors, etc.

As predicted, those who recalled perpetrator memories gave more reasons (especially unmarked beliefs) and fewer causal history explanations than those who recalled victim memories did. These differences in the way perpetrators and victims explain aggressive behaviors can be due to many cognitive factors. For instance, simple actor-observer differences can emerge because actors (e.g., perpetrators explaining their own behaviors) have privileged access to their reasons, whereas observers (e.g., victims explaining the perpetrators' behaviors) do not have such introspective access (e.g., Malle, 2004, 2011). Thus, actors, but not observers, have the option of using introspection when giving a reason explanation. When explaining their own aggressive behaviors, people can simply report the information they considered that led to their behavior. Of course, the differences in the way people explained their perpetrator memories and victim memories also are consistent with the idea that, to the extent that explanations can emphasize or de-emphasize the deliberateness of a behavior, people's explanations can be used to work towards their social goals. That is, although people generally give reason explanations for their own behaviors, sometimes they do not (we expand on this idea in our discussion of Hypothesis 4).

The current findings about perpetrators' and victims' folk explanations of aggressive behaviors are important because, not only do perpetrators and victims rate aggressive behaviors differently on scales that researchers provide them, but they *explain* aggressive behaviors differently in open-ended responses. Unlike rating scales that most participants could intuit what a more socially acceptable response would be, participants are unlikely to intuit what a more socially acceptable mode of folk explanations would be. These findings about folk explanations further emphasize that a person's role in an aggressive interaction is influential in how they perceive, think about, and *explain* their aggressive interactions.

Hypothesis #3: Desires for Apologies Predicts Actual Apologies Differently for Victims and Perpetrators

Our third hypothesis also replicated a previously-published effect showing that those who recalled a perpetrator memory more often reported that an apology was given than those who recalled a victim memory did. This difference in reported apologies is consistent with past research that explicitly asked participants to report whether an apology was given (e.g., Leunissen et al., 2013) and with past research when participants are simply asked to describe a past aggressive interaction and they are not explicitly asked to report whether an apology was given (e.g., Baumeister et al., 1990; Zechmeister & Romero, 2002). Although the current study found this difference in reported apologies to be replicable, unfortunately, the current study cannot speak to why people are more likely to report apologies when describing their perpetrator behaviors. This difference could be due to many factors such as an apology being a more memorable (and recallable) aspect of the aggressive interaction for perpetrators than for victims. Or this difference could reflect the idea that people will more often report

their apologies because they believe that information would paint them in a positive light (i.e., apologies would serve perpetrators' impression management goals more than victims'). It also is possible that perpetrators are simply more inclusive as to what they consider an "apology" than victims are when they are asked whether an apology was offered. For example, there is a distinction between apology acts, which can be perfunctory or insincere apologies or a routinized utterance (e.g., "sorry, I didn't see you there"), and sincere apologies (e.g., Smith, 2005). Perpetrators may be motivated to perform an apology act, such as insincerely uttering the phrase "I'm sorry," simply to get closure on an incident, to claim the moral status of a person who "owned up" to their mistakes, or to deter retaliation (e.g., Tavuchis, 1991). These "apologetic acts" would not have features of sincere apologies such as genuine contrition, acceptance of responsibility, or a commitment to change their future behavior. Nevertheless, when asked if an apology was offered, perpetrators may report their apology acts, which would include both insincere apologies and sincere apologies, and victims may only report (what they perceived to be) sincere apologies.

In addition to reporting whether an apology was given, participants in the current study also reported their "desire for an apology"—a desire to give an apology for those who described a perpetrator memory or a desire to receive an apology for those who described a victim memory. Although there was no overall difference in desires for an apology between those who described a perpetrator memory and those who described a victim memory, participants' desire for an apology more strongly predicted whether an actual apology was given for those who recalled perpetrator memories (this replicated the same interaction from Leunissen et al., 2013). This Role (Perpetrator vs. Victim) × Desire for Apology interaction most likely reflects the idea that people's thoughts predict their own behaviors better than others' behaviors. Apologies are, ultimately, something the perpetrator does: They say "I'm sorry," write an apologetic note, show remorse and contrition through their body language or tone of voice, etc. Thus, it is sensible that perpetrators' desires to give an apology strongly predict whether they actually gave an apology because this association reflects their desire predicting their behavior. In comparison, the extent that a victim wants to receive an apology is less predictive of whether the perpetrator, who is another person, actually gives an apology because this association reflects their (the victims') desire predicting another person's (the perpetrators') behavior.

Hypothesis #4: Folk Explanations Predicts Desires for Apologies Differently for Victims and Perpetrators

Our fourth hypothesis tested the novel idea that the way people explain their aggressive interactions was related to their desires for an apology in response to that interaction. This hypothesis was rooted in the idea that people do not simply explain their perpetrator behaviors differently than their victim behaviors, but that people can flexibly and strategically explain their behaviors depending

on whether they desire for an apology for the behavior. Specifically, although people will most often explain their perpetrator behaviors with belief reasons, we hypothesized that they would give causal history explanations when they strongly want to give an apology and, although people will commonly give causal history explanations to explain their victim behaviors, we hypothesized they would give belief reasons when they strongly want to receive an apology. These hypotheses were partly supported: We found the predicted pattern of results for those who explained perpetrator memories but not for those who explained victim memories.

Consider participants who recalled and explained perpetrator memories. When explaining a time they intentionally harmed another person, perpetrators had less of a desire to apologize for behaviors they explained with beliefs and they had a greater desire to apologize for behaviors they explained with causal history explanations (notably, our exploratory analysis found the same pattern for whether people actually gave an apology). Providing a belief reason for behaviors they do not want to apologize for could be simply articulating their (in their mind, justifiable) thoughts or it could be an attempt to persuade others that their actions were defensible (and, hence, not requiring an apology; c.f., Malle et al., 2000). For example, when explaining why they punched another person, one participant explained, “[h]e was challenging me in front of my 2 friends and I decided to accept the challenge.” While this explanation also leaves no doubt that the behavior was intentional, it also portrays their behavior as a plausibly justified response that does not require an apology. The explanation lays out, in plain language, their thought process that led to their behavior. Further making reconciliation difficult, as put by Schumann (2018), “transgressors feel a boost to their sense of power and self-esteem when they refuse to apologize, suggesting that withholding an apology might be used as a strategy to alleviate self-image threat” (p. 76). Thus, people who explain their aggressive behaviors with belief reasons, and who subsequently withhold an apology, may be entrenched in their position that they do not want to offer an apology.

“Causal history of reasons” explanations, on the other hand, do not presume the perpetrator considered, or was even aware of, the factors that led to the aggressive behavior (e.g., Malle, 2004, 2011). Thus, people can use causal history explanations if they want to de-emphasize the deliberateness of their aggressive behavior, such as when they harmed another person and they now feel apologetic about their actions. For example, when explaining why they insulted their wife, one participant simply explained, “I was frustrated.” This explanation does not claim the behavior was the product of a logical reasoning process, rather, it implies that an emotion that overcame them like an external force that forced their behavior (e.g., Schütz & Baumeister, 1999). Because causal history explanations do not presume awareness, behaviors that are explained with causal history explanations are less self-threatening for perpetrators and, hence, are easier to apologize for.

To the extent that belief explanations portray the aggressive behaviors as deliberate, and that causal history ex-

planations de-emphasize the perpetrators' intentionality, these findings are in line with Leunissen et al. (2013)'s conclusion that “perpetrators prefer to apologize after unintentional than after intentional transgressions” (p. 322). Notably, Leunissen et al. (2013) compared perpetrators' desires to apologize after intentional and unintentional (i.e., purely accidental) transgressions. In comparison, participants in the current study only considered intentionally harmful behaviors. Thus, participants who gave causal history explanations in the current study were not denying the intentionality of their behavior, rather, their causal history explanations de-emphasized the deliberateness of their behavior.

Now consider participants who recalled and explained victim memories. Because reason explanations presume the perpetrator *chose* to behave aggressively, which should frame the perpetrator as responsible for their harmful behavior, we hypothesized that victims would more strongly want to receive an apology when they explained the perpetrator's behavior with a belief explanation relative to a causal history explanation. This hypothesis was based on Leunissen et al. (2013)'s conclusion that “[v]ictims have a stronger preference for an apology after intentional transgressions than after unintentional ones” (p. 322). However, this was not what we found: In the current study, victims wanted an apology a similar amount regardless of how they explained the perpetrators' behavior. However, as noted in the previous paragraph, participants only considered intentional behaviors in the current study. Thus, it seems that victims want an apology for any intentionally harmful behavior regardless of how they explained it (however, and notably, our exploratory analysis showed that participants who explained their victim memories reported receiving more apologies for behaviors they explained with belief reasons than for behaviors they explained with causal history explanations).

It is not simply the case that people will always give reason explanations for their own aggressive behaviors and that people will more often give causal history explanations for other people's behaviors. Rather, the current studies showed that people can be flexible and strategic about how they explain their aggressive behaviors.

Limitations and Future Directions

One limitation of using the autobiographical recall approach, like what was used in the current study, is that participants may recall different aggressive behaviors depending on whether they recall a behavior from the perpetrator perspective or the victim perspective. Although we rated the severity of the recalled behaviors, which allowed us to “statistically account” for the behaviors' (perceived) severity, it is possible the recalled behaviors differed in other ways. To address the issue of selective recall, future research could use existing dyads and have each member of the dyad consider the same aggressive interaction (e.g., Feeney & Hill, 2006; Mikula et al., 1998). The results of the current study would be greatly bolstered if perpetrators' and victims' explanations of *the same behavior* were associated with their desires for apologies in the same way as

we observed in the current study. Other than considering the same behavior, another possibility would be to create an aggressive interaction in a lab, which would allow participants to be randomly assigned a role in the interaction and would allow the researchers to have an objective record of the aggressive behavior. Notably, although researchers have created mild transgressions within a lab for these sorts of studies (e.g., Adams & Inesi, 2016), the severity of these behaviors are typically low, so the generalizability of those results to “real world” harm must be made cautiously (e.g., McCarthy & Elson, 2018; Simons et al., 2017; however, see Anderson & Bushman, 1997; Mook, 1983). Nevertheless, studies with strong internal validity are important for making strong claims about of how a person’s role (i.e., either being a “perpetrator” or “victim”) in an aggressive interaction may causally affect their explanation of the behavior and how those explanations are related to their desires for apologies.

Another future direction would be to apply the current findings to post-conflict resolution. Apologies are least likely to be offered when perpetrators explain their behaviors with belief explanations—they feel as if their behavior was justified and that they simply do not owe an apology. One idea would be to prompt people to consider causal history explanations for their aggressive behaviors and see if they would then have a stronger desire to offer an apology to the person they harmed. The idea would be that generating a causal history explanation would make their transgression seem less personally threatening to the self, which would then make an apology easier to give.

Besides a desire to give an apology, other ideas could test whether different explanations for an aggressive behavior would make apologies more or less effective. For example, victims believe perpetrators have less of a desire for forgiveness for intentional transgressions (e.g., Adams & Inesi, 2016), victims are more likely to reject an apology when the perpetrator is viewed as responsible for the transgression (e.g., Bennett & Earwaker, 1994), and victims are less likely to actually grant forgiveness in response to an apology for an intentional transgression (e.g., Struthers et al., 2008). Consequently, to the extent that victims’ folk explanations are related to the explainer’s perception of the deliberateness of the behavior, we could predict that vic-

tims who explain the perpetrator’s behaviors with reasons, which portrays the behavior as more deliberate, may be less forgiving of the perpetrator *even when an apology has been offered*.

Conclusion

It seems trite to say that “the words we use matter,” but that is exactly what the current study showed. In situations where a perpetrator explains their behavior with belief explanations, apologies (and possibly other forms of conflict resolution) seem unlikely to happen—they believe their behavior is justified and they have little desire to apologize. Victims, on the other hand, want to receive an apology when they are harmed regardless of how they explain the behavior.

That perpetrators and victims explain their aggressive interactions differently is no surprise. But the current study underscores how the way people naturally explain their aggressive interactions can make it easy for perpetrators and victims to talk past one another when deciding how much they want an apology after an aggressive interaction.

Competing Interests

We have no competing interests.

Author Contributions

Randy J. McCarthy: Conceptualization, writing, data collection and analysis, and project administration.

Jared P. Wilson: Conceptualization, writing, data collection and analysis, and project administration.

Data Accessibility Statement

Stimuli, data, and code for the current study are accessible at the project’s Research Box page: <https://research-box.org/516>.

Submitted: August 18, 2021 PDT, Accepted: July 19, 2023 PDT



References

- Adams, G. S., & Inesi, M. E. (2016). Impediments to forgiveness: Victim and transgressor attributions of intent and guilt. *Journal of Personality and Social Psychology, 111*(6), 866–881. <https://doi.org/10.1037/pspi0000070>
- Anderson, C. A., & Bushman, B. J. (1997). External validity of "trivial" experiments: The case of laboratory aggression. *Review of General Psychology, 1*(1), 19–41. <https://doi.org/10.1037/1089-2680.1.1.19>
- Baumeister, R. F. (1996). *Evil: Inside human cruelty and violence*. W H Freeman/Times Books/ Henry Holt & Co.
- Baumeister, R. F., Stillwell, A., & Wotman, S. R. (1990). Victim and perpetrator accounts of interpersonal conflict: Autobiographical narratives about anger. *Journal of Personality and Social Psychology, 59*(5), 994–1005. <https://doi.org/10.1037/0022-3514.59.5.994>
- Bennett, M., & Earwaker, D. (1994). Victims' Responses to Apologies: The Effects of Offender Responsibility and Offense Severity. *The Journal of Social Psychology, 134*(4), 457–464. <https://doi.org/10.1080/00224545.1994.9712196>
- Birch, S. A. J., & Bloom, P. (2007). The Curse of Knowledge in Reasoning About False Beliefs. *Psychological Science, 18*(5), 382–386. <https://doi.org/10.1111/j.1467-9280.2007.01909.x>
- Darley, J. M., & Pittman, T. S. (2003). The Psychology of Compensatory and Retributive Justice. *Personality and Social Psychology Review, 7*(4), 324–336. https://doi.org/10.1207/s15327957pspr0704_05
- Dersley, I., & Wootton, A. (2000). Complaint Sequences Within Antagonistic Argument. *Research on Language & Social Interaction, 33*(4), 375–406. https://doi.org/10.1207/s15327973rlsi3304_02
- Elshout, M., Nelissen, R. M. A., Beest, I. van, Elshout, S., & Dijk, W. W. van. (2020). Real-life revenge may not effectively deter norm violations. *The Journal of Social Psychology, 160*(3), 390–399. <https://doi.org/10.1080/00224545.2019.1681351>
- Elshout, M., Nelissen, R. M. A., & van Beest, I. (2017). Your act is worse than mine: Perception bias in revenge situations. *Aggressive Behavior, 43*(6), 553–557. <https://doi.org/10.1002/ab.21713>
- Ent, M. R., & Parton, D. M. (2020). "I Avenge; Others Aggress": A Victim–Perpetrator Asymmetry in Judging Whether a Transgression Was Motivated by Revenge. *Psychological Reports, 123*(4), 1355–1364. <https://doi.org/10.1177/0033294119841846>
- Feeney, J. A., & Hill, A. (2006). Victim-perpetrator differences in reports of hurtful events. *Journal of Social and Personal Relationships, 23*(4), 587–608. <https://doi.org/10.1177/0265407506065985>
- Fiske, S. T. (2018). *Social Beings: Core Motives in Social Psychology*. John Wiley & Sons.
- Forster, D. E., Billingsley, J., Burnette, J. L., Lieberman, D., Ohtsubo, Y., & McCullough, M. E. (2021). Experimental evidence that apologies promote forgiveness by communicating relationship value. *Scientific Reports, 11*(1), 13107. <https://doi.org/10.1038/s41598-021-92373-y>
- Grice, H. P. (1975). Logic and Conversation. *Speech Acts, 41*–58. https://doi.org/10.1163/9789004368811_003
- Lazare, A. (2005). *On Apology*. Oxford University Press.
- Leunissen, J. M., De Cremer, D., Reinders Folmer, C. P., & van Dijke, M. (2013). The apology mismatch: Asymmetries between victim's need for apologies and perpetrator's willingness to apologize. *Journal of Experimental Social Psychology, 49*(3), 315–324. <https://doi.org/10.1016/j.jesp.2012.12.005>
- Malle, B. F. (2004). *How the Mind Explains Behavior: Folk Explanation, Meaning and Social Interaction*. The MIT Press. <https://doi.org/10.7551/mitpress/3586.001.0001>
- Malle, B. F. (2011). Time to Give Up the Dogmas of Attribution. In *Advances in Experimental Social Psychology* (Vol. 44, pp. 297–352). Elsevier. <https://doi.org/10.1016/b978-0-12-385522-0.00006-8>
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A Theory of Blame. *Psychological Inquiry, 25*(2), 147–186. <https://doi.org/10.1080/1047840x.2014.877340>
- Malle, B. F., Knobe, J. M., & Nelson, S. E. (2007). Actor-observer asymmetries in explanations of behavior: New answers to an old question. *Journal of Personality and Social Psychology, 93*(4), 491–514. <https://doi.org/10.1037/0022-3514.93.4.491>
- Malle, B. F., Knobe, J., O'Laughlin, M. J., Pearce, G. E., & Nelson, S. E. (2000). Conceptual structure and social functions of behavior explanations: Beyond person–situation attributions. *Journal of Personality and Social Psychology, 79*(3), 309–326. <https://doi.org/10.1037/0022-3514.79.3.309>
- McCarthy, R. J., & Elson, M. (2018). A Conceptual Review of Lab-Based Aggression Paradigms. *Collabra: Psychology, 4*(1), 4. <https://doi.org/10.1525/collabra.104>
- McCarthy, R. J., Jensen, A. P., Wilson, J. P., & Rivers, A. K. (2022). Perpetrators', Victims', and Witnesses' Folk Explanations of Aggressive Behaviors. *Psychological Reports, 00332941221123781, 00332941221123781*. <https://doi.org/10.1177/00332941221123781>
- McCarthy, R. J., & Rivers, A. K. (2021). Does Trait Hostility Moderate the Asymmetry in Victim–Perpetrator Perceptions of Aggressive Behaviors? *Psychological Reports, 125*(2), 703–722. <https://doi.org/10.1177/0033294120984123>
- 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>

- Mikula, G., Athenstaedt, U., Heschgl, S., & Heimgartner, A. (1998). Does it only depend on the point of view? Perspective-related differences in justice evaluations of negative incidents in personal relationships. *European Journal of Social Psychology*, 28(6), 931–962.
- Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 38(4), 379–387. <https://doi.org/10.1037/0003-066x.38.4.379>
- Rosenthal, D. (2002). Consciousness and the Mind. *Iyyun: The Jerusalem Philosophical Quarterly / עיון: רבעון פילוסופי*, 251–227. <https://doi.org/10.2307/23352846>
- Schumann, K. (2018). The Psychology of Offering an Apology: Understanding the Barriers to Apologizing and How to Overcome Them. *Current Directions in Psychological Science*, 27(2), 74–78. <https://doi.org/10.1177/0963721417741709>
- Schütz, A., & Baumeister, R. F. (1999). The Language of Defense: Linguistic Patterns in Narratives of Transgressions. *Journal of Language and Social Psychology*, 18(3), 269–286. <https://doi.org/10.1177/0261927x99018003003>
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on Generality (COG): A Proposed Addition to All Empirical Papers. *Perspectives on Psychological Science*, 12(6), 1123–1128. <https://doi.org/10.1177/1745691617708630>
- Smith, N. (2005). The Categorical Apology. *Journal of Social Philosophy*, 36(4), 473–496. <https://doi.org/10.1111/j.1467-9833.2005.00289.x>
- Stillwell, A. M., & Baumeister, R. F. (1997). The Construction of Victim and Perpetrator Memories: Accuracy and Distortion in Role-Based Accounts. *Personality and Social Psychology Bulletin*, 23(11), 1157–1172. <https://doi.org/10.1177/01461672972311004>
- Stillwell, A. M., Baumeister, R. F., & Priore, R. E. D. (2008). We're All Victims Here: Toward a Psychology of Revenge. *Basic and Applied Social Psychology*, 30(3), 253–263. <https://doi.org/10.1080/01973530802375094>
- Struthers, C. W., Eaton, J., Santelli, A. G., Uchiyama, M., & Shirvani, N. (2008). The effects of attributions of intent and apology on forgiveness: When saying sorry may not help the story. *Journal of Experimental Social Psychology*, 44(4), 983–992. <https://doi.org/10.1016/j.jesp.2008.02.006>
- Tavuchis, N. (1991). *Mea culpa: A sociology of apology and reconciliation*. Stanford University Press.
- Tsang, J.-A. (2002). Moral Rationalization and the Integration of Situational Factors and Psychological Processes in Immoral Behavior. *Review of General Psychology*, 6(1), 25–50. <https://doi.org/10.1037/1089-2680.6.1.25>
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child Development*, 72(3), 655–684. <https://doi.org/10.1111/1467-8624.00304>
- Zechmeister, J. S., & Romero, C. (2002). Victim and offender accounts of interpersonal conflict: Autobiographical narratives of forgiveness and unforgiveness. *Journal of Personality and Social Psychology*, 82(4), 675–686. <https://doi.org/10.1037/0022-3514.82.4.675>
- Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, 111(4), 493–504. <https://doi.org/10.1037/pspa0000056>

Supplementary Materials

Peer Review History

Download: https://collabra.scholasticahq.com/article/84918-perpetrators-and-victims-folk-explanations-of-aggressive-behaviors-and-desires-for-apologies/attachment/174688.docx?auth_token=Y4ZhBQb65XEej5vQHqa
