

Blaming Blamers: Differential Obligation to Punish for Third-Parties Compared to Victims

by

Tessa Jean Gardner

Honors Thesis

Appalachian State University

Submitted to the Department of Psychology

in partial fulfillment of the requirements for the degree of

Bachelor of Arts

May, 2019

Approved by:

Andrew Monroe, Ph.D., Thesis Director

Alex Howard, Ph.D., Second Reader

Andrew Smith, Ph.D., Departmental Honors Director

Abstract

The present study examines the different ways we morally evaluate first-party (victims) and third-party (uninvolved observers) in judgement scenarios; when given the opportunity to punish or forgive a moral perpetrator. Participants were asked to read an interaction of individuals playing an economic game where one player cheats the other. Participants then read one of four accounts of either the player who was cheated punishing or forgiving the offender, or a third-party observer punishing or forgiving the offender. After reading the account the participants evaluated the judger in their scenario; we found that people are okay with victims who forgive or punish and third parties who punish, but participants saw third-parties who forgave as extremely immoral and sanctioned them for their decision. The results of this study support the theory that we hold third and first-parties to different moral norms; expecting third-parties to uphold what society believes is just, making punishment a rational protective choice for third-parties.

Permission is granted to Appalachian State University and the Department of Psychology to display and provide access to this thesis for appropriate academic and research purposes.

Blaming Blamers: Differential Obligation to Punish for Third-Parties Compared to Victims

In June of 2016 the ruling in the now infamous Brock Turner case, also popularly known as the Stanford swimmer case, was made public. Brock Turner, a student at Stanford University, was caught sexually assaulting an unconscious women one night by two other student who detained him. Despite the severity of the case, two eye-witness accounts, and a jury conviction, Aaron Persky, the judge presiding over the case, sentenced Turned to six months in prison, ignoring the statutory minimum sentence of two years. In the weeks and months that followed, Judge Persky's ruling was met with widespread protests (Gonzales & Domonoske, 2018). These protests ultimately culminated in a successful bid to recall Judge Persky, citing his failure to sufficiently punish Turner.

On its face, the reaction to Judge Persky's ruling is well explained by previous work arguing that people have an inherent "desire to punish" wrongdoers (Walsh, 2006). However, comparing people's strong negative reaction to Judge Persky to the overwhelmingly positive public reaction to the victims of the Charleston church shooting forgiving the shooter, Dylann Roof (Berman, 2015) suggests a more complex story and that the demand for punishment may not be as ubiquitous as previously theorized.

The present studies are motivated by these contrasting phenomena – condemnation for forgiving third-parties (e.g., judges, uninvolved observers) and praise for forgiving first-parties (i.e., victims). Specifically, my project examines how people morally evaluate first-parties versus third-parties who decide to punish or forgive moral wrongdoing, and we test whether differences in these judgments are explained by people applying different obligations to punish on victims vs. third-parties.

Social psychology has a long history examining humans' motivation to blame. Research shows that blame has been used to maintain a normal social and moral order in communities and cultures for centuries (Haidt, 2008). Psychologists have often viewed judgement and blame as an intuitive and automatic function, born from an innate want to uphold social values and expectations for behavior (Nadler & McDonnell, 2012). Further, people may be motivated to blame others who violate these moral codes of society in order to regulate and enforce prioritized shared values and norms (Haidt & Joseph, 2004). However, despite blame's usefulness as a tool for behavior regulation, the motivation to blame is often framed as a bias. For example, work by Ditto and colleagues (2009) suggests that people's motivation to blame is so robust that people may actually change their perceptions of events—inflating perceptions of intentionality, harm, or preventability—to validate the desire to blame wrongdoing. In our research we explore the idea that the motivation to blame – or at least to not let wrongdoers go completely unpunished – maybe a rational decision strategy for third parties in order to avoid condemnation from the community and build a positive social reputation. Recent research demonstrates that although it can be socially risky for a third-party to blame and punish others, punishing conveys a commitment to moral norms that increased people's willingness to trust punishers (Jordan, Hoffman, Bloom & Rand, 2016). Studies have outlined both the positive and negative social repercussions of being a punisher. When examining factors that explained the willingness of individuals to punish, Raihani & Bshary (2015a) found that when punishing someone who has violated a social norm the punishing third-party is rewarded with good social standing and the perception of a positive cooperative attitude. Meaning third-party judges can socially benefit from punishing wrongdoers. Further, although punishing wrongdoers has reputational benefits, when given the option people give greater reputational rewards to third-parties that help victims

instead of punishing perpetrators (Raihani & Bshary, 2015b). Costly helping, helping a victim despite possible personal expense, is viewed as a stronger signal of trustworthiness than punishment and gives the individuals better reputational benefits than punishment does (Jordan et al. 2016). Even when punishment is seen as justified, helping third-parties receive more rewards than punishing third-parties (Raihani & Bshary, 2015b).

In addition to punishing having possible reputational benefits, recent work demonstrates that observers often make judgments of other's moral judgements. That is, people morally evaluate whether a person blames too harshly or too leniently as well as judging the reasons and motivations for one's moral decisions (Everette, Pizarro & Crockett, 2016). People tend to prefer empathetic people over individuals who make decisions in a purely logical (Everett et al., 2018). And, in a similar vein, people have implicit norms about how much blame is warranted by different types of wrongdoing, and that either over or under blaming are sanctioned in turn (Monroe & Malle, 2018).

This suggests that punishment – so long as it is justified – may not be a pervasive bias, but rather an adaptive heuristic for third-party judges not to let “bad” actors ‘off the hook’ because failing to blame in such cases may result in being blamed in turn. Since we know that blame has a history of being used to uphold the values and morals of a community, not blaming could be seen as a failure to protect the community by upholding its norms and values. To make sure its core values are respected a community could in turn put blame upon those who fail to blame, deterring other from making the same “mistake” and jeopardizing the norms of community. However, all this research has been focused on perceptions of and implications for third-parties, not first-parties (e.g. victims). Studies have shown that a victim's satisfaction with

a punishment can affect third-parties, but the social perceptions of a victim who blames or takes no action have not been evaluated (Gromet, Okimoto, Wenzel & Darley, 2012).

In our studies we wanted to explore the phenomenon of blaming blamers, and specifically, we wanted to test the proposition that people believe that third-parties are more obligated to punish wrongdoing compared to victims. Participants will be shown the interaction between individuals in an econ game in which one player gets cheated. They will then be shown the cheated player, the victim, choosing to forgive or punish or an observing third-party choosing to forgive or punish. After this participants will be asked to make judgements on the judger, the victims or third party, in five different areas: blame/praise, targets' moral character, participants' desire to punish desire to punish, anger, and participants' impression of whether the decider made the correct decision. We predict that there will be little difference between how victims who forgive or punish are rated. Concerning third-parties we believe those who punish will receive praise, be seen as highly moral, as having made the right decision, and will receive little to no anger or want to punish. We expect to see the opposite in third-parties who forgive; high rating of blame, anger, a strong desire to punish them, low rating of moral character, and little belief that they made the correct decision.

Method

Participants

We collected responses from 465 participants. The sample was almost evenly split between men ($n = 224$) and women ($n = 237$, with two participants identifying as non-binary). The majority of the sample identified as White ($n = 351$), followed by Asian Americans ($n = 37$), African Americans ($n = 33$), Native Americans ($n = 15$), Latinx individuals ($n = 14$), multi-ethnic individuals ($n = 13$), and one Middle Eastern participant. The sample was politically

moderate ($M = 3.80$, $SD = 1.79$) based on a 1 (very liberal) to 7 (very conservative) scale and moderately religious ($M = 2.67$, $SD = 1.49$) on a 1 (not at all religious) to 5 (very religious) scale.

Procedure

Participants were recruited online using Amazon Mechanical Turk (MTurk). They read and signed an informed consent sheet. Afterwards, they were informed that the experiment involved reading about and evaluating an interaction between two participants who played an economic game in a different experiment. On the next page, participants read about a cooperative economic game played by two players: Jon and Andy. Jon and Andy were described as being assigned to the role of Player 1 (Jon) or Player 2 (Andy); Player 1 was given \$10 and the option to send anywhere between \$1 and \$10 to Player 2, which would then be quadrupled, Player 2 would then have the option to send back as much or as little money as he wanted to Player 1. Once the interaction was over if Player 1 felt Player 2 had acted unfairly he was given the chance to either do nothing or to punish Player 2 by reducing his winnings to \$0. In the story Andy (Player 2) convinces Jon (Player 1) to trust him and to send all \$10, after the money is quadrupled, Andy decides to keep the resulting \$40, leaving Jon with \$0.

After reading one of the four reports, participants made judgements about the deciders (the punishing or forgiving first vs. third party in their report). Participants rated how much they blamed or praised the third/first party on a -5 (a lot of blame) to 5 (a lot of praise) scale. They judged the decider's moral character on a -5 (extremely bad moral character) and 5 (extremely good) scale. Participants also rated whether they thought the decider should be punished for their choice and whether the decider made the correct decision (1 absolutely no – 7 absolutely yes). Finally participants rated how angry they were with the decider on a 1 (not at all) to 7

(extremely) scale. After completing these measures, participants filled out a short demographics questionnaire and were debriefed.

Results

In order to test the prediction that third parties are held to higher standards than first parties and always expected to punish, we conducted a 2 (Agent: victim vs. third party) x 2 (Decision: forgive vs. punish) between subjects ANOVA on participants' ascriptions of blame/praise, targets' moral character, participants' desire to punish, anger, and participants' impression of whether the decider made the correct decision. For blame/praise judgements, the ANOVA revealed the predicted interaction, $F(1, 347) = 136.0, p < .001$ (See Figure 1). Whereas participants praised victims slightly more when they forgave compared to when they punished wrongdoers ($p = .041$), 3rd party judges were severely blamed when they forgave the wrongdoer, but praised when they punished ($p < .001$). Similar interaction effects emerged for judgments of a target's moral character, $F(1, 346) = 180.2, p < .001$. People saw victims as having strong moral character regardless of if they punished or forgave the perpetrator (though character is perceived as better for forgiving victims relative to punishing ones, $p < .001$); 3rd parties, however, were seen as having bad moral character if they forgave the offender, but positive moral character if they punished the wrongdoer, $p < .001$ (See Figure 2). For desire to punish, the ANOVA revealed the predicted interaction, $F(1, 346) = 84.5, p < .001$. Regardless of the decision the choice the victim made, to punish or forgive, people did not want to punish them for their decision ($p = .224$); in the case of third-parties those who punished were seen as undeserving of punishment but overall people wanted third-parties who forgave to be punished, $p < .001$ (See Figure 3). For anger at the decider, the ANOVA revealed the predicted interaction,

$F(1, 346) = 45.4, p < .001$. Overall, participants reported very little anger toward victims, though they were slightly angrier when victims forgave than when they punished ($p = .016$). Similarly, participants were not angry toward third-parties who punished, but anger significantly increased toward third-parties who choose to forgive the offender, $p < .001$ (See Figure 4). Finally, evaluations of whether the target made the correct decisions showed the predicted interaction, $F(1, 345) = 34.3, p < .001$. Although people generally thought victims made a correct decision, they viewed punishing as more correct than forgiving ($p < .001$). This difference was substantially exacerbated for judgements of third-parties. Punishing third-parties were viewed as deciding correctly; whereas forgiving third-parties were judged as deciding incorrectly ($p < .001$) (See Figure 5).

General Discussion

The present study tested if people evaluate forgiving versus punishing first-parties (victims) differently from third-party judges and whether differences in these judgements are explained by people applying different obligations to punish on first versus third-parties. The study demonstrated that victims were allowed broad latitude regarding their decision to punish or not; regardless of their choice victims were viewed positively and their moral character was not smeared. By contrast, third-parties are held to a different, and more stringent standard. Third-parties were perceived as being obligated to punish wrongdoers, and when they did so, they were appraised positively and similarly to victims. However, forgiving third-party judges were blamed, viewed as having bad moral character, and sanctioned at significantly higher levels compared to victims. Interestingly, for both victims and third-parties, participants viewed punishing wrongdoers as making the correct decision; however, forgiving victims appeared to pay little for making the ‘wrong choice.’ They were still appraised as having good moral

character and not being blameworthy. Conversely, forgiving third-parties were heavily penalized for making the identical wrong choice.

Together, these findings suggest that past demonstrations that third-party judges are motivated to blame (Ditto, Pizarro & Tannenbaum, 2009) may be a rational reaction to perceived social expectations rather than a punitive bias. Our data suggest that social perceivers expect third-parties to punish wrongdoing when possible, and that having bad moral character, and possibly even punished themselves. This aligns with similar research arguing that blaming, while potentially costly (Monroe & Malle, 2019) also confers benefits (if blame judgements are appropriate), as punishing third-parties are often praised for their decisions and viewed as trustworthy (Jordan et al., 2016). This may also explain why we see third-parties being praised more than victims for punishing. Over-praising could be a way for society to condition third-parties to continue making the decisions they like. Whereas future work future work is needed to test whether judges are aware of these expectations, the present data indicate that the expectation on the part of social perceivers is robust.

Despite our effort to make this realistic for the participants these are still hypothetical situations; there is a chance that in the real-world people would act differently. Similarly, in the lab setting there are no real stakes; participants do not risk any consequences when they blame the third-party and their anger does not accumulate into any real-life sanctions. In the future, we would like to study how people feel about third-parties who help victims, compared to third-parties who punish or forgive. We would also like to create media or a lab setting in which participants can observe the interaction between the victim, wrongdoer, and third party in order to remove the level of detachment our study may have.

In sum, the present study supports the idea that in judgement situations victims and third-parties are given different moral norms. Victims, by virtue of being victims, are allowed to make any choice they like and do not have a societally imposed obligation to punish, since they have already suffered at the hand of the one they are punishing. However, third-parties are obligated to punish and risk being sanctioned as immoral if they fail to uphold justice in the public eye. Due to this obligation to punish we propose that third-parties have a rational motivation to blame.

References

- Berman, M. (2015, June 19th), 'I forgive you.' Relatives of Charleston church shooting victims address dylann roof. *The Washington Post*, Retrieved from <https://www.washingtonpost.com>
- Ditto, H., P., Pizarro, A., P., & Tannenbaum, D. (2009). Motivated moral reasoning. In *The psychology of learning and motivation* (pp. 307-334). Burlington: Academic Press.
- Everett, J. A. C., Faber, N. S., Savulescu, J., & Crockett, M. (2018). The costs of being consequentialist: Social inference from instrumental harm and impartial beneficence, *Journal of Experimental Psychology*. Retrieved from <https://doi.org/10.31234/osf.io/a2kx6>
- Everett, J., Pizarro, D., & Crockett, M. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology*. Retrieved from <http://dx.doi.org/10.2139/ssrn.2726330>
- Gonzales, R., & Domonoske, C. (2018, June 5th). Voters recall aaron persky, judge who sentenced brock turner. *npr*, Retrieved from <https://www.npr.org>
- Gromet, M., D., Okimoto, G., T., Wenzel, M., & Darley, M., J. (2012). A victim-centered approach to justice? victim satisfaction effects on third-party punishments. *Law and Human Behavior*, 36(5), 375-389. DOI: 10.1037/h0093922
- Hadit, J. (2008). Morality. *Perspectives on Psychological Science*, 3(10), 65-72. Retrieved from <https://doi.org/10.1111/j.1745-6916.2008.00063.x>
- Hadit, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(40), 55-66.

Jordan, J., J., Hoffman, M., Bloom, P., & Rand, D., G. (2016). Third-party punishments as a costly signal of trustworthiness. *Nature*, 530, 473-476. DOI: 10.1038/nature16981

Monroe, E., A., & Malle, F., B. (2018). People systematically update moral judgements of blame. *Journal of Personality and Social Psychology*. DOI: 10.1037/pspa0000137

Nadler, J., & McDonnell, M. (2012). Moral character, motive, and the psychology of blame. *Cornell Law Review* 97 (2).

Raihani, J., N., & Bshary, R. (2015a). The reputation of punishers. *Trends in Ecology & Evolution*, 30(2). DOI: 10.1016/j.tree.2014.12.003

Raihani, J., N., & Bshary, R. (2015b). Third-party punishers are rewarded – but third- party helpers even more so. *Evolution*, 69(4). DOI: 10.1111/evo.12637

Walsh, A. (2006). Evolutionary psychology and the origins of justice. *Justice Quarterly*, 17(4), 841-865. DOI: 10.1080/07418820000094781

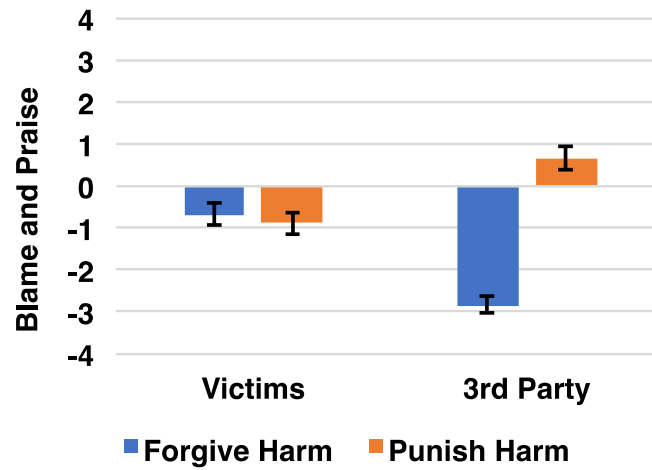


Figure 1. Whereas people blame victims little regardless whether they forgive or punish wrongdoing, people severely blame third-party judges who forgive and praise them if they punish. Error bars = ± 1 SE.

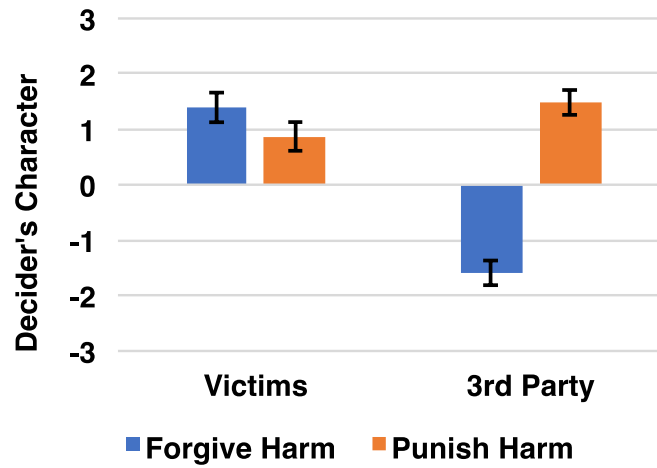


Figure 2. People perceived punishing and forgiving victims as having positive moral character. Punishing third-parties were similarly viewed as morally good, but forgiving third-parties were perceived as having morally bad character. Error bars = ± 1 SE.

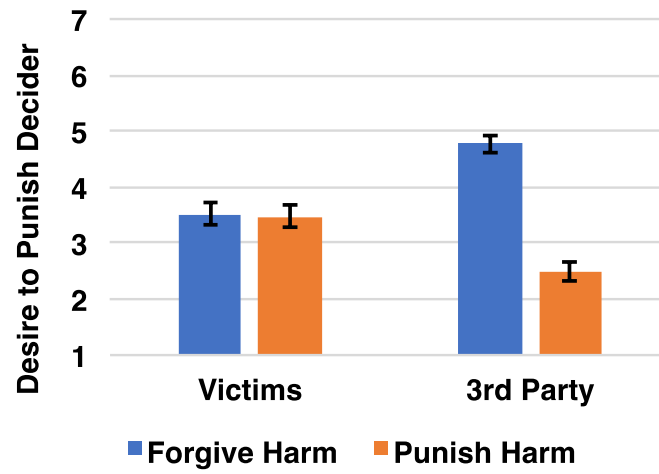


Figure 3. People did want to punish victims regardless of their decision, but saw third-parties who forgave as deserving of punishment. Error bars = $\pm 1 SE$.

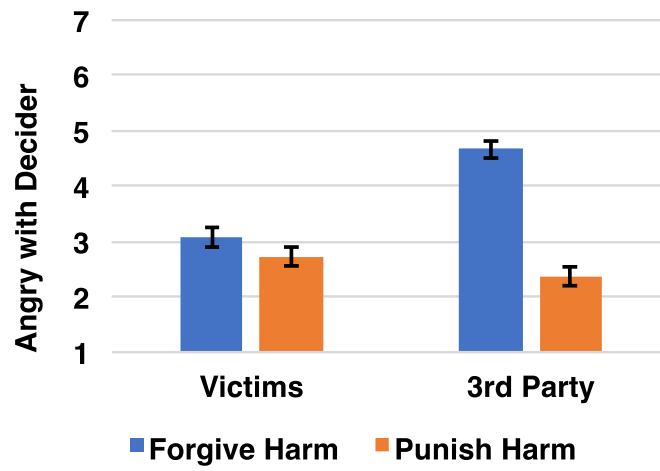


Figure 4. Whether they punished or forgave people were not very angry with victims, similarly they were not angry with punishing third-parties. However, they were extremely angry with third-parties who forgave. Error bars = ± 1 SE.

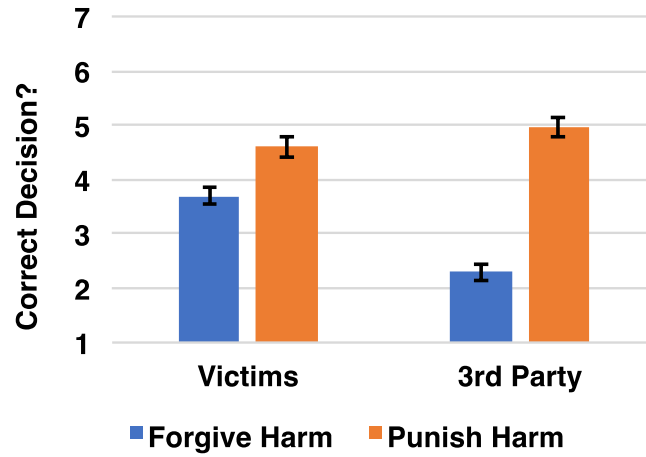


Figure 5. Victims who punished and third-parties who punished were seen as having made the correct decision. While victims who forgave were seen as having made an okay decision, third-parties who forgave were seen as having made the wrong decision. Error bars = $\pm 1 SE$.