Replication of Study 7 by Exline, Baumeister, Zell, Kraft & Witvliet (2008, *Journal of Personality and Social Psychology*)

Stephanie Lin

Stanford University

Introduction

Exline, Baumeister, Zell, Kraft, and Witvliet’s (2008) show in seven studies that people are more likely to feel forgiving after imagining committing similar offenses themselves. In studies 4-7, an interaction of gender emerges such that only men become more forgiving after imagining their own capability for wrongdoing than women. In study 6, the vengefulness subscale of the unforgiveness scale (TRIM: Transgression Related Interpersonal Motivations Inventory-Revised) drives the interaction such that men felt less vengeful after imagining their own capability for wrongdoing than women. Study 7, which is the study chosen for replication here, shows that even subtle priming (simply asking participants questions about their empathic understanding and personal capability of similar transgressions) can reduce men’s feelings of vengefulness.

Methods

**Power Analysis**  
 The power analysis indicates that we need 128 subjects in order to capture an ANOVA interaction effect size η2 = .06 with 4 groups and a numerator df = 1; this will give us a 80% chance of detecting the effect if it indeed exists.

**Planned Sample**

The study will be administered via Stanford Graduate School of Business’s “Studies in Mass.” In “Studies in Mass,” approximately 140 Stanford affiliates (mostly students, but some staff also) take part in approximately 12 five-minute studies in an hour in exchange for $20.   
  
**Materials**

Participants will first be asked to recall a time that they were offended, harmed or hurt by someone else. Participants will self-report their personal capability of wrongdoing and empathic understanding via three measures.

One of the measures will assess personal capability for committing an offense similar in severity to the one that they recalled. Participants will rate responses to four questions from 0 (no, definitely not) to 10 (yes, definitely): “Given the right circumstances, do you think that you could be capable of doing something just as bad (i.e., just as harmful or wrong) as what the other person did?”; “Can you imagine a situation in which you could do something as bad as what the other person did?”; “Do you think it’s possible that you could ever do something as bad as what the other person did?”; and “Thinking back over your entire life, do you think that you have ever done anything as bad as what the other person did?”

As in the original report, a second set of personal capability questions will emphasize similarity in offense type rather than severity (e.g., “Given the right circumstances, could you be capable of doing something similar in type to what the other person did?”). This and the severity scale will be counterbalanced.

After the personal capability questions, they will be asked the four empathic understanding items. Participants will be asked, “To what extent can you . . .” followed by four prompts: “understand why the other person acted as s/he did”; “see the situation from the other person’s perspective”; “see his/her behavior as making sense;” and “think of valid reasons why s/he acted as s/he did.” Responses will be rated from 0 (not at all) to 10 (totally).

Participants will also be administered the 18-item Transgression Related Interpersonal Motivations Inventory–Revised (TRIM-18-R; McCullough & Hoyt, 2002) to assess current motivations toward the offender. The TRIM consists of three subscales: avoidance motivations, revenge motivations, and benevolence motivations. The revenge motivations subscale (e.g. I’ll make him/her pay; I want to see him/her hurt and miserable) will be the dependent variable of interest.

The TRIM will be administered either before or after the capability—empathy questions above, which serves as the manipulation for this study; that is, men who have answered the capability—empathy questions before the TRIM should report less vengefulness on the TRIM than those who answer the questions after the TRIM.

**Procedure**

Participants will first be asked: “Please take a few minutes to recall a time in which you were deeply offended, harmed, or hurt by another person, and as a result still have some anger or resentment towards that person.  You must write at least 200 characters (about 30 words) to move on.” In one condition, participants will answer capability—empathy questions first, and then answer the TRIM, which contains the vengeance measures. In the other, participants will answer the TRIM first, followed by the capability—empathy questionnaire.

**Analysis Plan**  
 We will conduct a 2 (gender) by 2 (order of capability-empathy items) ANOVA on participants’ TRIM scores. We expect a significant interaction for the vengefulness subscale such that men who completed the capability—empathy measures before the vengefulness questions should score lower on this subscale than those who complete the measures afterward. Women’s scores should not be affected by the order of the questions.

**Differences from Original Study**

Although participants will still be run in the lab, the study will be administered in the form of an online survey rather than paper-and-pencil. However, this should not affect the results, as the authors used online versions of their questionnaires in study 1. Although study 1 did not use the exact same procedure as study 7, it was similar in that it required the participants to recall an instance when they were hurt or offended. No significant interactions emerged between student sample and internet sample, student status and non-student status, age, gender (TRIM was not measured here), ethnicity, or nationality in their analyses. Therefore, using an online format in the lab should be sufficient to replicate the results.

Furthermore, it will be administered as one of a series of studies, which may lower concentration on the study (the order of all twelve studies is randomized in “Studies in Mass”). We are asking participants to write at least 200 characters (about 30 words) in order to ensure compliance with the manipulation. Otherwise, procedures and analysis should be close to the original study. Participants who clearly do not complete the manipulation properly will be excluded.

**(Post Data Collection) Methods Addendum**

**Actual Sample**  
 Participants were 135 Stanford affiliates (55 men, 80 women). All participants completed the manipulation sufficiently. Thirty-three participants identified themselves as Caucasian, 84 as Asian/Pacific Islander, 11 as African-American, 12 as Hispanic, 2 as Native American or Alaskan Native, and 1 did not respond. The mean age was 21.64 (SD = 3.23). Actual power with original effect size of η2 = .06 is 82.2%

**Differences from pre-data collection methods plan**  
 Compared to the original study, there seem to be many more Asians and fewer Caucasians. Cultural differences may cause some variance in the data.

Furthermore, when analyzing the data, it seemed that the vengeance questions (and therefore the overall TRIM) differed slightly from the one used by the original authors. We used the TRIM-R as appeared in McCullogh & Hoyt, 2002; the vengefulness items from those were as follows: “I’ll make him/her pay,” “I wish that something bad would happen to him/her,” “I want him/her to get what he/she deserves,” “I’m going to get even,” and “I want to see him/her hurt and miserable.” The original items reported by the authors were: “I want them to suffer,” “I want revenge,” “I wish that something bad would happen to them,” “I hate them,” “I feel cold toward them,” “I feel bitter toward them,” and “I want to see them hurt and miserable” (Exline et al., 2008, p. 503). When contacted, the authors indicated that they used the McCullogh & Hoyt, 2002 TRIM-R that was identical to the one used in this replication attempt; therefore, it is unclear why the scale items differed, but they should be testing the same construct.

Results

**Data preparation**  
 Vengeance scores were calculated as indicated by McCullough, Root, and Cohen (2006) (M = 1.95, SD = 0.87). Overall TRIM scores were calculated using the mean across all scale items (M = 2.68, SD = 0.76), with the benevolence items reverse coded. Items that measured one’s capability for *equally bad* offense were averaged (M = 5.29 SD = 2.93) into a single score. Items that measured one’s capability to commit a *similar* offense were also averaged (M = 4.71, SD = 2.73). As in the original study, capability to commit an equally bad and a similar offense were correlated r(133) = .77, p < .001, so all eight items from the two scales were averaged to form a personal capability score (M = 5.00, SD = 2.66). Empathy was calculated using the mean of the four empathy questions (M = 5.33, SD = 2.53).

**Confirmatory analysis**  
 I conducted a 2 (gender) by 2 (condition) ANOVA on vengefulness scores. There was a main effect on gender (F(1,131) = 4.51, p = .04) such that males (M = 2.15, SD = .86 were more vengeful than females (M = 1.81, SD = .87). The main effect of condition was not significant (F(1,131) = .08, p = .77). The interaction, was also not significant (F(1, 131) = 0.018 , p = .89, η2 <.001).

For the overall TRIM score, the main effect of gender approached significance (F(1, 131) = 3.52 , p = .06) such that men (M = 2.55, SD = .71) were less forgiving than women (M = 2.79, SD = .79). The effect of condition was significant (F(1,131) = 4.63, p = .03) in the predicted direction such that those who viewed the capability questions after the TRIM had higher TRIM scores (M = 2.79, SD = .77) than those who viewed the capability questions before the TRIM scale (M = 2.55, SD = .72). The interaction between gender and condition was not significant for the overall TRIM score (F(1, 131) = 0.15 , p = .70).  
  
**Exploratory analyses**

There were no gender effects on personal capability (t(116) = -.59, p = .56) or empathic understanding (t(116) = -.81, p = .42). As expected, those who scored higher on empathy and capability scored lower on TRIM (unforgiveness) and vengefulness. There was a significant correlation between capability and vengefulness (r(133) = -.18, p = .04) such that those who felt more capable of committing a similar or equally bad offense felt less vengeful. Furthermore, the correlations between capability and TRIM score (r(133) = -.29, p < .001), empathy and vengefulness score (r(133) = -.15, p = .08), and empathy and TRIM (r(133) = -.32, p <.001) score were significant or marginally significant in the expected directions.

In addition, throughout these analyses, a pattern emerged such that capability and empathy were more correlated with vengefulness and TRIM (total unforgiveness) when capability/empathy questions were asked after the TRIM than when they were asked before the TRIM (see appendix for summary of results). See below for discussion.

Discussion

**Summary of Replication Attempt**

Overall,the original findings that thinking about one’s own capability for wrongdoing decreases vengefulness feelings in men failed to replicate. However, there was evidence that thinking about one’s own capability for wrongdoing does lead to more forgiveness. Specifically, being primed with the capability/empathy questions did lead to the predicted pattern of people scoring lower on the unforgiveness scale (TRIM) relative to those who viewed the capability/empathy questions after completing the TRIM scale. This did not hold for the vengefulness subscale. Although not a direct replication of study 7, this serves as a theoretical replication of the germane findings of Exline et al. (2007). Interestingly, the original authors did not observe this effect (F = 2.34, η2 = .02).

Also of note, there were interactions between condition (capability/empathy questions before and after) and personal capability for wrongdoing on overall unforgiveness (TRIM), and a similar pattern (with nonsignificant interaction) on vengefulness. That is, people who completed capability and empathy questions *after* completing the TRIM scale showed a negative correlation between their own capability for wrongdoing and their level of vengefulness and unforgiveness. However, people who completed the questions *beforehand* did not show this correlation. Similarly, they also showed negative correlations between empathy and vengefulness as well as empathy and unforgiveness when asked the capability/empathy questions after the TRIM, but not when asked beforehand. Although it is a completely speculative account, one plausible explanation is that those who were not primed with personal capability used severity of the offense as a heuristic to answer both the TRIM and the capability questions, whereas those primed with personal capability took a more sensitive approach to answering the TRIM scale.

**Commentary**

The population did differ between this replication attempt and the original study, especially in ethnic makeup. This sample was 62% Asian (62%) and only 24% Caucasian whereas the original sample had 79% Caucasian and 16% Asian. These proportions are significantly different (χ2(2) = 250, p < .001) and therefore it is possible that this contributed to the different findings. For instance, it has been argued that collectivist cultures have different patterns of forgiveness due to motivation of promoting social harmony (Hook, Worthington & Utsey, 2009), and that collectivists are more willing to forgive and have less lasting resentment (Suwartono, Prawasti & Mullt, 2006) among other findings. As many of this sample were of Asian descent, it could be that they simply have different conceptualizations of forgiveness that might be worth investigating in the future. The actual scale used for vengefulness different also, although the constructs tested are the same.

**Appendix**

When looking at the correlations between conditions, there was a correlation between capability and vengefulness when the capability/empathy questions were asked after the vengefulness questions (r(69) = -.28, p = .02) such that those who deemed themselves as more capable of similar or equally bad wrongdoings were also less vengeful. The correlation was not significant when the capability/empathy questions were asked before (r(69) = -.11, p = .39). However, this interaction of condition by personal capability was not significant (B = .02, p = .33).

The correlation between capability and TRIM score was marginally moderated by condition (B = -.04, p = .075). When capability questions were asked before the TRIM measure, there was no correlation (r(62) = -.16, p = .19) but there was a significant correlation when the TRIM scale preceded the capability/empathy questions (r(69) = -.39, p <.001).

Empathy was marginally negatively correlated with vengefulness overall (r(133) = -.15, p = .08). As above, this was marginally moderated by condition (B = -.05, p = .09). Empathy and vengefulness were not correlated when the empathy/capability questions preceded the TRIM (r(62) = .002, p = .99). However, empathy and vengefulness were negatively correlated when the TRIM preceded the capability/empathy questions TRIM (r(69) = -.31, p = .007).

The empathy and TRIM correlation was significant with a similar pattern to the capability findings. This was significantly moderated by condition (B = -.06, p = .015). That is, the correlation was significant for people who completed the empathy measures *after* the TRIM (r(69) = -.48, p <.001) but nonsignificant for people who completed those measures before the TRIM (r(62) = -.11, p = .39).

References

**Exline, J.J.,** Baumeister, R.F., Zell, A.L., Kraft, A., & Witvliet, C.V.O. (2008). Not so innocent: Does seeing one's own capability for wrongdoing predict forgiveness? Journal of Personality and Social Psychology, 94, 495-515.

Hook, J.N., Worthington, E.L., & Utsey, S.O. (2009). Collectivism, Forgiveness, and Social Harmony. *The Counseling Psychologist*, *37* 821-847.

Suwartono, C., Prawasti, C.Y., & Mullet, E. (2007). Effect of culture on forgivingness: A Southern Asia-Western Europe comparison. *Personality and Individual differences, 42*(3), 513-523.