

Supplementary Materials

Apology Cost Is More Strongly Associated With Perceived Sincerity Than Forgiveness

Yohsuke Ohtsubo & Miyu Higuchi

1. Participant Exclusion Criteria

We had a total of 1668 accesses to our study site. Of the 1668 accesses, 105 did not agree to participate in the study. Of the remaining 1563 accesses, 77 were discarded because multiple responses were recorded from the same IP address, 41 were discarded for incomplete responses, and 130 were discarded due to employment status and age-based restrictions. Additional responses were excluded because they did not meet the pre-registered inclusion criteria: 466 participants failed to pass attention check items, and 59 participants completed the study within 3 minutes or took longer than 1 hour. Accordingly, 790 participants (406 women, 380 men, and 4 no-reported their sex; mean age $\pm$ SD = 31.23 $\pm$ 5.61 years) were retained for subsequent analyses.

2. Manipulation Checks

The first manipulation check item asked participants whether the transgressor was aware of the harmful impact of their action. The scenarios in the no intention condition explicitly described that the transgressor was not aware of it. Therefore, this score should be lower in the no intention condition than in the other two conditions. The responses to the four scenarios were aggregated (Cronbach's α coefficient = .63) and submitted to a 2 (apology cost) $\times$ 3 (intention) ANOVA. The result indicated that the main effect of intention was significant: $F(2, 784) = 67.34, p < .001, \eta^2 = .146$, while the main effect of apology cost and their interaction were not significant: $F(1, 784) = 2.89, p = .090, \eta^2 = .003$ and $F(2, 784) = 2.11, p = .122, \eta^2 = .005$ for the main effect of apology cost and the apology cost $\times$ intention interaction, respectively. As predicted, Tukey's HSD test revealed that the mean score in the no intention condition (2.29, $SD = 0.70$) was significantly smaller than the mean scores in the ambiguous intention condition (2.85, $SD = 0.74$) and the malicious intention condition (2.96, $SD = 0.72$) at the .001 level.

The second manipulation check item asked participants whether the transgressor prioritized their own business despite their awareness of the foreseeable harmful impact of their act. The scenarios in the malicious intention condition explicitly described that the transgressor was aware of it and did not attempt to avoid it (i.e., prioritized their own business). The scenario in the ambiguous intention condition did not explicitly describe the prioritization part. However, participants may infer that the transgressor prioritized their own business because they were aware of the possible harmful impact. Therefore, this score should be highest in the malicious intention condition, which is followed by the ambiguous intention condition, and lowest in the no intention condition. The responses to the four scenarios were aggregated (Cronbach's α coefficient = .67) and submitted to a 2 (apology cost) $\times$ 3 (intention) ANOVA. The result indicated that the main effect of intention, $F(2, 784) = 61.17, p < .001, \eta^2 = .133$, and the main

effect of apology cost, $F(1, 784) = 21.46, p < .001, \eta^2 = .023$, were significant. However, the interaction effect was not significant, $F(2, 784) = 0.68, p = .509, \eta^2 = .001$. As predicted, Tukey's HSD tests revealed that the mean score in the malicious intention condition (2.99, $SD = 0.68$) was significantly higher than the means in the other two conditions at the .001 level. The mean score in the ambiguous intention condition (2.66, $SD = 0.74$) was higher than the mean in the no intention condition (2.30, $SD = 0.76$). Unexpectedly, the mean score in the costly apology condition (2.52, $SD = 0.78$) was lower than the mean in the non-costly apology condition (2.76, $SD = 0.77$).

The third manipulation check item asked whether the apology that each scenario described was costly in some way (e.g., financially costly, time-consuming, psychologically costly). Therefore, it should be higher in the costly apology condition than in the non-costly apology condition. The responses to the four scenarios were aggregated (Cronbach's α coefficient = .81) and submitted to a 2 (apology cost) $\times$ 3 (intention) ANOVA. The results indicated that the main effects of apology cost, $F(1, 784) = 248.84, p < .001, \eta^2 = .234$, and intention, $F(2, 784) = 12.52, p < .001, \eta^2 = .025$, were significant, but their interaction was not significant, $F(2, 784) = 1.91, p = .148, \eta^2 = .004$. As predicted, the mean score was higher in the costly apology condition (3.64, $SD = 0.79$) than in the non-costly apology condition (2.75, $SD = 0.84$). Unexpectedly, the mean score in the no intention condition (3.39, $SD = 0.90$) was significantly higher than the mean scores in the other two conditions: 3.19, $SD = 0.97$ and 3.05, $SD = 0.87$ in the ambiguous and malicious intention conditions, respectively.

Based on the responses to the manipulation check items, we excluded participants whose responses did not meet the expected pattern in their assigned conditions. Applying these strict inclusion criteria, we retained 298 participants in the data set. Analyses with this smaller subset of participants did not change the conclusion reported in the main text. The results of these analyses are available in R Markdown format (analysis2.html) from OSF (<https://osf.io/sfyzq/>).

3. Perceived Exploitation Risk and Valuation

Participants were asked whether the transgressor may commit a similar transgression again in the future (perceived exploitation risk). The perceived exploitation risk scores for the four scenarios were aggregated (Cronbach's α coefficient = .68) and submitted to a 2 (apology cost) $\times$ 3 (intention) ANOVA. The results showed that the main effects of apology cost, $F(1, 784) = 49.31, p < .001, \eta^2 = .055$, and intention, $F(2, 784) = 31.46, p < .001, \eta^2 = .072$, were significant, but their interaction was not significant (see Figure 2c in the main text). The mean exploitation risk was lower in the costly apology condition (2.98, $SD = 0.78$) than in the non-costly apology condition (3.33, $SD = 0.70$). The mean exploitation risk was 2.91 ($SD = 0.70$), 3.17 ($SD = 0.79$), and 3.40 ($SD = 0.71$) in the no, ambiguous, and malicious intention conditions, respectively. These three means were significantly different from each other at the .001 level.

Participants also reported how much they think the transgressor values the relationship with them. The responses to the four scenarios were aggregated (Cronbach's α coefficient = .77) and submitted to a 2 (apology cost) $\times$ 3 (intention) ANOVA. The results showed that the main effects of apology cost, $F(1, 784) = 152.60, p < .001, \eta^2 = .151$, and intention, $F(2, 784) = 35.12, p < .001, \eta^2 = .073$, were significant, but their interaction was not significant (see Figure 2d in the

main text). The mean valuation was higher in the costly apology condition (3.83, $SD = 0.78$) than in the non-costly apology condition (3.19, $SD = 0.74$). The mean valuation was 3.77 ($SD = 0.90$), 3.54 ($SD = 0.80$), and 3.24 ($SD = 0.79$) in the no, ambiguous, and malicious intention conditions, respectively. These three means were significantly different from each other at the .001 level.

4. A 2 (Apology Cost) $\times$ 2 (Intention: No vs. Ambiguous Intention) ANOVA

The predictions derived from Ohtsubo and Watanabe's (2009) results were concerning a 2 (apology cost) $\times$ 2 (intention: no vs. ambiguous intention) factorial design (see Figure 1 in the main text). Therefore, we ran a set of 2 $\times$ 2 ANOVAs on forgiveness and perceived sincerity to confirm that the apology cost $\times$ intention interaction is not significant.

When forgiveness was submitted to a 2 $\times$ 2 ANOVA, the main effects of apology cost, $F(1, 531) = 35.36, p < .001, \eta^2 = .061$, and intention, $F(1, 531) = 9.45, p = .002, \eta^2 = .016$, were significant, but their interaction was not significant, $F(1, 531) = 1.18, p = .278, \eta^2 = .002$. When perceived sincerity was submitted to a 2 $\times$ 2 ANOVA, the main effects of apology cost, $F(1, 531) = 154.46, p < .001, \eta^2 = .222$, and intention, $F(1, 531) = 10.35, p = .001, \eta^2 = .014$, were significant, but their interaction was not significant, $F(1, 531) = 0.30, p = .585, \eta^2 < .001$.

The same is true for perceived exploitation risk and valuation. For perceived exploitation risk, the main effects of apology cost, $F(1, 531) = 33.41, p < .001, \eta^2 = .057$, and intention, $F(1, 531) = 17.27, p < .001, \eta^2 = .029$, were significant, but their interaction was not significant, $F(1, 531) = 1.40, p = .237, \eta^2 = .002$. For valuation, the main effects of apology cost, $F(1, 531) = 115.52, p < .001, \eta^2 = .175$, and intention, $F(1, 531) = 13.32, p < .001, \eta^2 = .020$, were significant, but their interaction was not significant, $F(1, 531) = 0.22, p = .638, \eta^2 < .001$.

5. Comparison of the Effect Sizes Between the No vs. Ambiguous Intention Conditions

Although we reported non-significant interaction between apology cost and intention conditions in the main text (in a 2 $\times$ 3 design) and in Section 4 of this Supplementary Material (in a 2 $\times$ 2 design), we did not directly compare the effect sizes of the no intention condition and the ambiguous intention condition. We admit that the effect of apology cost was significant in both conditions by post hoc tests and the apology cost $\times$ intention interaction failed to reach the statistically significant level. Nonetheless, the visual inspections of Figure 3 (in the main text) suggest that the effect of apology cost was larger in the ambiguous intention condition than in the no intention condition (see the upper two data points in Figure 3). Although the 95% CIs of the two correlations substantially overlap, to confirm the non-significant difference between these two correlations, we ran a test of independent correlations. The results showed that the apology cost $\times$ forgiveness correlation was not significantly smaller in the no intention condition (.21, $df = 271, p < .001$) than in the ambiguous intention condition (.30, $df = 260, p < .001$): $z = 1.11, p = .267$. Parenthetically, the apology cost $\times$ forgiveness correlation in both the no and ambiguous intention conditions was not significantly different from that in the malicious intention condition (.27, $df = 253, p < .001$): $z = 0.32, ns$, and 0.78, ns , for the no-malicious and ambiguous-malicious comparisons, respectively.