

Supplementary Results

Analytical Plan

Consistent with our analyses in the main manuscript, we conducted a linear mixed model (LMM) analysis on each continuous variable (i.e., moralization judgment, emotional reaction, reputational concern) for moral and immoral actions, respectively, with culture (US vs. Chinese), age group (6-7 years vs. 8-9 years), and time (Past vs. Future) being fixed, between-subjects factors and subject (child) being a random factor. However, no covariates were included in the models.

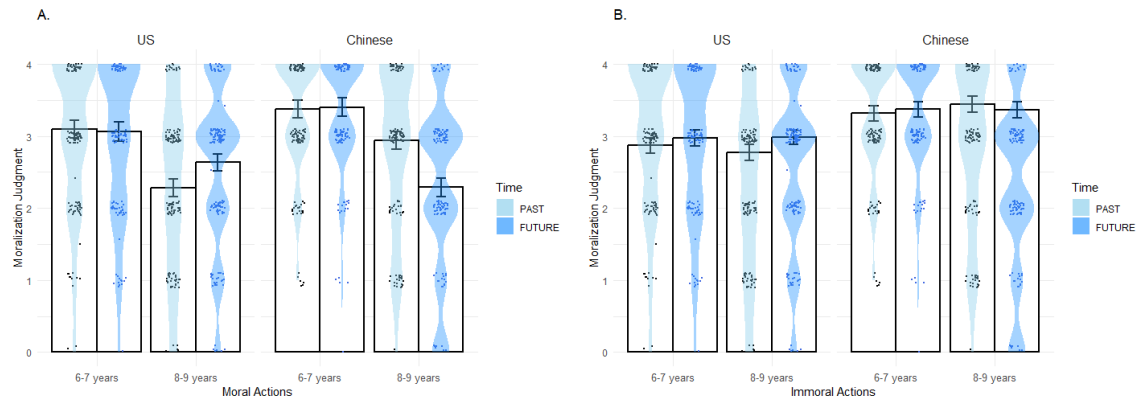
Across all analyses, the variance explained by the random effect of subject ranged between .28 to .65, and none of the 95% CI for the random effect overlapped with 0, indicating the random effect in all models was statistically significant. Model intraclass correlation coefficients (ICC) ranged between .31 to .44.

Moralization Judgment

A 2 (culture) x 2 (age group) x 2 (time condition) LMM analysis was first conducted on moralization judgment for moral actions, which yielded main effects of culture, $F(1, 338) = 6.84$, $p = .009$, $\eta_p^2 = .02$, and age group, $F(1, 338) = 62.00$, $p < .001$, $\eta_p^2 = .15$, as well as an interaction between culture and time, $F(1, 338) = 7.10$, $p = .008$, $\eta_p^2 = .02$, and a three-way interaction between culture, time, and age, $F(1, 338) = 9.14$, $p = .003$, $\eta_p^2 = .03$. The adjusted means from the model, in addition to the distribution of the raw data, are shown in Figure S1A.

Figure S1

Children's Moralization Judgment for Moral (A) and Immoral Actions (B)



Note. The violin plots represent the distribution of raw data. Individual data point was jittered for better visual presentation. The bar plots represent the adjusted means, with error bars representing corresponding standard errors.

To interpret the effect of culture in the context of three-way interactions, we conducted a 2 (culture) x 2 (time condition) LMM analysis within each age group. Among 6-7-year-olds, a main culture effect emerged, $F(1, 169) = 8.68, p = .004, \eta_p^2 = .05, 95\%CI = [.10, .51]$, whereby Chinese children ($M = 3.39, SE = .07$) called for more praises for their moral actions than did US children ($M = 3.08, SE = .07$) across time conditions. Among the 8-9-year-olds, a significant Culture x Time interaction emerged, $F(1, 169) = 12.33, p < .001, \eta_p^2 = .07$. Whereas Chinese children ascribed more praises than did US children to their past moral actions, $t(169) = 3.24, p = .001, d = .77, 95\%CI = (.26, 1.06)$, US children ascribed marginally more praises than did Chinese children to their future moral actions, $t(169) = 1.72, p = .087, d = .4195\%, CI = [-.75, .05]$. Interpreted alternatively, whereas Chinese 8-9-year-olds ascribed more praises to their past than future moral actions, $t(169) = 3.09, p = .002, d = .76, 95\%CI = [.24, 1.07]$, US 8-

9-year-olds ascribed marginally more praises to their future than past moral actions, $t(169) = -1.83$, $p = .069$, $d = .42$, 95% CI = [-.74, .03].

To delve into the effect of age groups, we also conducted a 2 (age) x 2 (time condition) LMM analysis within each cultural group. For US children, the analysis yielded a main age effect, $F(1, 175) = 25.95$, $p < .001$, $\eta_p^2 = .13$, 95% CI = (-.86, -.38), whereby 6-7-year-olds ($M = 3.08$, $SE = .09$) attributed more praises to their moral actions than did 8-9-year-olds ($M = 2.46$, $SE = .08$) across time conditions. For Chinese children, there were main effects of age group, $F(1, 163) = 36.18$, $p < .001$, $\eta_p^2 = .18$, and time condition, $F(1, 163) = 5.91$, $p = .016$, $\eta_p^2 = .03$, qualified by an Age x Time interaction, $F(1, 163) = 6.96$, $p = .009$, $\eta_p^2 = .04$. Chinese 6-7-year-olds made similar moralization judgements in the past and future conditions ($p = .88$), whereas 8-9-year-olds attributed more praises to their past than future moral actions, $t(163) = 3.51$, $p < .001$, $d = .98$, 95% CI = [.29, 1.02]. Alternatively, although Chinese 6-7-year-olds attributed more praises to their moral actions than did 8-9-year-olds in both time conditions, the age effect was more salient for past moral actions $t(163) = 2.42$, $p = .017$, $d = .65$, 95% CI = [.08, .79], than for future moral actions, $t(163) = 6.03$, $p < .001$, $d = 1.67$, 95% CI = [.75, 1.47].

A similar 2 (culture) x 2 (age group) x 2 (time condition) LMM analysis was conducted for immoral actions, which yielded only a main effect of culture, $F(1, 338) = 38.49$, $p < .001$, $\eta_p^2 = .10$. Across time conditions and age groups, Chinese children ($M = 3.37$, $SE = .05$) called for more punishments for their immoral actions than did US children ($M = 2.90$, $SE = .05$), 95%CI = [.32, .62]. See Figure S1B.

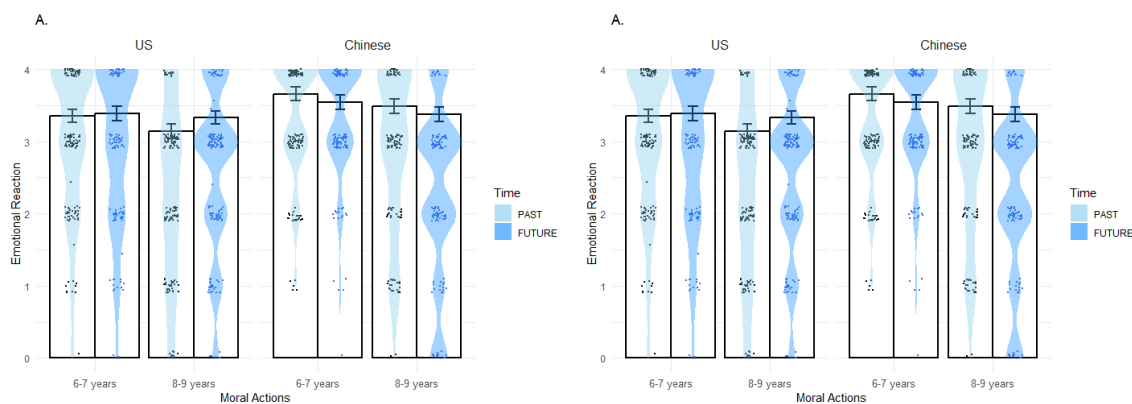
Emotional Reaction

A 2 (culture) x 2 (age group) x 2 (time condition) LMM analysis on emotional reaction for moral actions yielded a main effect of culture, $F(1, 338) = 9.48$, $p = .002$, $\eta_p^2 = .03$, whereby

Chinese children ($M = 3.51$, $SE = .05$) reported stronger emotional reactions to their moral actions than did US children ($M = 3.30$, $SE = .05$), $95\%CI = [.08, .35]$. There was also a significant age effect, $F(1, 338) = 4.88$, $p = .028$, $\eta_p^2 = .01$, whereby 6-7-year-olds ($M = 3.48$, $SE = .05$) reported more emotional reactions to their moral actions than did 8-9-year-olds ($M = 3.33$, $SE = .05$), $95\%CI = [.02, .29]$. There was no significant effect pertaining to time condition. A similar LMM analysis for immoral actions yielded no significant main effects or interactions ($ps > .24$).

Figure S2

Children's Emotional Reaction to Moral (A) and Immoral (B) Actions



Note. The violin plots represent the distribution of the raw data. Individual data point was jittered for better visual presentation. The bar plots represent the adjusted means, with error bars representing corresponding standard errors.

Reputational Concern

A 2 (culture) x 2 (age group) x 2 (time condition) LMM analysis on reputational concern for moral actions yielded a main effect of culture, $F(1, 338) = 4.41$, $p = .037$, $\eta_p^2 = .01$, $95\%CI = [.01, .31]$, whereby Chinese children ($M = 3.19$, $SE = .06$) reported greater reputational concern

to their moral actions than did US children ($M = 3.03$, $SE = .05$). There was also a main effect of age, $F(1, 338) = 24.43$, $p < .001$, $\eta_p^2 = .07$, $95\%CI = [.23, .53]$, whereby 6-7-year-old children ($M = 3.30$, $SE = .05$) reported stronger reputational concern to their moral actions than did 8-9-year-olds ($M = 2.92$, $SE = .05$). There was no significant effect pertaining to the time condition. The adjusted means from the model, in addition to the distribution of the raw data, are shown in Figure S3A.

Consistent with our analysis in the main manuscript, we also conducted additional analysis to explore the Culture x Age x Time three-way interaction here, although the effect itself was not significant, $p = .134$.

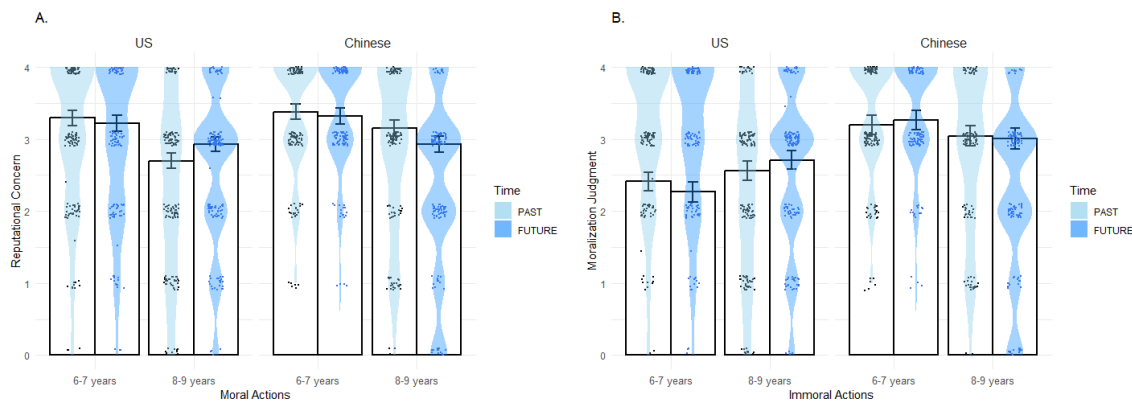
Follow-up 2 (culture) x 2 (time condition) LMM analyses within each age group showed that among 6-7-year-olds, a significant effect of culture emerged $F(1, 169) = 8.68$, $p = .004$, $\eta_p^2 = .05$, such that Chinese children ($M = 3.39$, $SE = .07$) expressed significantly great reputational concern than did US children ($M = 3.08$, $SE = .087$) across time conditions, $95\% CI = [.10, .51]$. Among 8-9-year-olds, a significant Culture x Time condition interaction emerged for 8-9 year-olds, $F(1, 169) = 12.33$, $p < .001$, $\eta_p^2 = .07$. Chinese 8-9 year-olds expressed stronger reputational concern than did US children for their past moral actions, $t(169) = 3.24$, $p = .001$, $d = .77$, $95\% CI = [.26, 1.06]$, but the cultural difference was not significant for future moral actions ($p = .087$; marginally significant in the reversed direction with $M_{US} = 2.64$, $SE_{US} = .13$ and $M_{CN} = 2.29$, $SE_{CN} = .15$).

In addition, exploratory 2 (age) x 2 (time condition) LMM analysis conducted within each cultural group showed that among US children, a significant main effect of age emerged, $F(1, 175) = 25.95$, $p < .001$, $\eta_p^2 = .13$, $95\%CI = [.38, .86]$, whereby US 6-7-year-olds ($M = 3.08$, $SE = .09$) exhibited a stronger reputational concern than US 8-9-year-olds ($M = 2.46$, $SE = .08$)

across time conditions. Among Chinese children, a significant main effect of age emerged, $F(1, 163) = 36.18, p < .001, \eta_p^2 = .18$, as well as a significant effect of time condition, $F(1, 163) = 5.91, p = .016, \eta_p^2 = .03$, qualified by a significant Age x Time interaction, $F(1, 163) = 6.96, p = .009, \eta_p^2 = .04$. While there was no effect of time condition among Chinese 6-7-year-olds ($p = .88$), Chinese 8-9-year-olds expressed significantly greater reputational concern for their past moral actions ($M = 2.94, SE = .13$) than their future moral actions ($M = 2.29, SE = .13$), $t(163) = 3.51, p < .001, d = .98, 95\% = CI = [.29, 1.02]$. Alternatively, while Chinese 6-7-year-olds expressed greater reputational concerns for their moral actions than did Chinese 8-9-year-olds across time conditions, the age effect was more salient for future actions, $t(163) = 6.03, p < .001, d = 1.67, 95\% = CI = [.75, 1.47]$, than for past actions, $t(163) = 2.42, p = .017, d = .65, 95\% = CI = [.08, .79]$.

Figure S3

Children's Reputational Concern for Moral (A) and Immoral (B) Actions



Note. The violin plots represent the distribution of the raw data. Individual data point was jittered for better visual presentation. The bar plots represent the adjusted means, with error bars representing corresponding standard errors.

For immoral actions, a similar LMM analysis showed a main effect of culture, $F(1, 338) = 44.06, p < .001, \eta_p^2 = .12$, qualified by an interaction between culture and age group, $F(1, 338) = 6.63, p = .010, \eta_p^2 = .02$ (See Figure 3B). Follow-up comparisons revealed that although Chinese children expressed greater reputational concern than did US children regardless of age, the cultural difference was more salient among 6-7-year-olds ($M_{US} = 2.34, SE_{US} = .10; M_{CN} = 3.23, SE_{CN} = .10$), $t = 6.52, p < .001, d = 1.07, 95\%CI = [.62, 1.16]$, than among 8-9-year-olds ($M_{US} = 2.63, SE_{US} = .09; M_{CN} = 3.02, SE_{CN} = .10$), $t = 2.87, p = .004, d = .47, 95\%CI = [.12, .66]$. Alternatively, whereas US 8-9-year-olds exhibited stronger reputational concern than did 6-7-year-olds across time conditions, $t = 2.19, p = .029, d = .35, 95\%CI = [.03, .57]$, the age difference was not significant among Chinese children ($p = .14$). There was no significant effect pertaining to time condition.