

Supplemental Online Materials

Study 1

All populations and domains together. When further collapsing across the full sample and all domains ($N = 620$, $\alpha = .88$), the effect remained robust ($M_{\text{Past}} = 3.72$, $SD_{\text{Past}} = 1.22$; $M_{\text{Future}} = 4.24$, $SD_{\text{Future}} = 1.30$), $t(618) = -5.07$, $p < .001$, $d = 0.41$, 95% $CI_{\text{diff}} = [-0.71, -0.31]$.

Results by population and domain. Students perceived a friend's dream vacation as stinging less after versus before they take it, $t(201) = -4.55$, $p < .001$, $d = 0.64$, 95% $CI_{\text{diff}} = [-1.38, -0.55]$; a friend's dream date as stinging less after versus before they go on it, $t(201) = -3.42$, $p = .001$, $d = 0.48$, 95% $CI_{\text{diff}} = [-1.31, -0.35]$; and a friend's new dream house stinging less after versus before they pick up the keys, $t(201) = -2.18$, $p = .030$, $d = 0.31$, 95% $CI_{\text{diff}} = [-0.92, -0.05]$. Students did not envy a friend significantly less before versus after they started a dream job, $t(201) = -1.05$, $p = .297$, $d = 0.15$, 95% $CI_{\text{diff}} = [-0.67, 0.21]$, or before versus after they bought a dream car, $t(201) = -1.27$, $p = .206$, $d = 0.18$, 95% $CI_{\text{diff}} = [-0.64, 0.14]$.

Community adults perceived the same vacations as less enviable after versus before they are taken, $t(215) = -2.11$, $p = .036$, $d = 0.28$, 95% $CI_{\text{diff}} = [-0.83, -0.03]$; the same job as less enviable after versus before they start, $t(215) = -3.26$, $p = .001$, $d = 0.44$, 95% $CI_{\text{diff}} = [-1.08, -0.27]$; the same house as less enviable after versus before they are moved into, $t(215) = -2.85$, $p = .005$, $d = 0.39$, 95% $CI_{\text{diff}} = [-0.99, -0.18]$; and the same cars as less enviable after versus before they are acquired, $t(215) = -3.13$, $p = .002$, $d = 0.43$, 95% $CI_{\text{diff}} = [-1.04, -0.24]$. Community adults did not envy a date significantly less after versus before it occurs, $t(215) = -1.47$, $p = .144$, $d = 0.20$, 95% $CI_{\text{diff}} = [-0.78, 0.12]$.

For M-Turkers too, past vacations were less enviable than future vacations, $t(198) = -4.01, p < .001, d = 0.57, 95\% CI_{diff} = [-1.23, -0.42]$; past dates less than future dates, $t(198) = -2.66, p = .009, d = 0.38, 95\% CI_{diff} = [-1.08, -0.16]$; and past car purchases less than future car purchases, $t(198) = -2.16, p = .032, d = 0.31, 95\% CI_{diff} = [-0.81, -0.04]$. M-Turkers did not view a friend's dream job as significantly less enviable after versus before starting, $t(198) = -0.67, p = .504, d = 0.09, 95\% CI_{diff} = [-0.53, 0.26]$, or after versus before moving into a dream house, $t(198) = -1.25, p = .214, d = 0.18, 95\% CI_{diff} = [-0.62, 0.14]$.

Condition checks. There were no differences in difficulty among students ($M_{Past} = 2.70, SD = 1.78$ vs. $M_{Future} = 2.60, SD = 1.72$), the community ($M_{Past} = 2.10, SD = 1.57$ vs. $M_{Future} = 1.96, SD = 1.53$), or M-Turk ($M_{Past} = 2.18, SD = 1.45$ vs. $M_{Future} = 2.34, SD = 1.60$), $ts \leq 0.74, ps \geq .459, ds \leq 0.11$; nor in realism among students ($M_{Past} = 4.30, SD = 1.81$ vs. $M_{Future} = 3.89, SD = 1.81$), the community ($M_{Past} = 4.51, SD = 1.89$ vs. $M_{Future} = 4.90, SD = 1.67$), or M-Turk ($M_{Past} = 5.26, SD = 1.53$ vs. $M_{Future} = 5.28, SD = 1.51$), $ts \leq 1.61, ps \geq .108, ds \leq 0.23$; nor in confusion among students ($M_{Past} = 1.73, SD = 1.37$ vs. $M_{Future} = 1.79, SD = 1.36$), the community ($M_{Past} = 1.43, SD = 0.94$ vs. $M_{Future} = 1.60, SD = 1.28$), or M-Turk ($M_{Past} = 1.49, SD = 1.00$ vs. $M_{Future} = 1.71, SD = 1.28$), $ts \leq 1.35, ps \geq .177, ds \leq 0.19$; nor in detail among students ($M_{Past} = 4.34, SD = 1.77$ vs. $M_{Future} = 4.33, SD = 1.69$), the community ($M_{Past} = 5.14, SD = 1.66$ vs. $M_{Future} = 5.09, SD = 1.65$), or M-Turk ($M_{Past} = 5.39, SD = 1.53$ vs. $M_{Future} = 5.47, SD = 1.38$), $ts \leq 0.39, ps \geq .698, ds \leq 0.06$; and nor in closeness among students ($M_{Past} = 6.16, SD = 1.17$ vs. $M_{Future} = 6.24, SD = 1.17$), the community ($M_{Past} = 5.77, SD = 1.48$ vs. $M_{Future} = 5.91, SD = 1.31$), or M-Turk ($M_{Past} = 6.03, SD = 0.93$ vs. $M_{Future} = 6.02, SD = 0.88$), $ts \leq 0.74, ps \geq .461, ds \leq 0.10$.

Study 2a

Attention check. Four percent of participants (111 of 2,824) failed the attention check. All participants were retained in analyses.

Condition checks. Throughout the month the task was similarly easy ($M_{\text{overall}} = 2.10$, $SD = 1.44$), realistic ($M_{\text{overall}} = 5.53$, $SD = 1.31$), non-confusing ($M_{\text{overall}} = 1.71$, $SD = 1.26$), and elicited similarly detailed mental imagery ($M_{\text{overall}} = 5.16$, $SD = 1.37$). These null trends are captured via the same regression analyses we conducted for envy: there were main effects of tense, $\beta_s \leq 0.03$, $ts \leq 0.77$, $ps \geq .440$, no main effects of distance, $\beta_s \leq 0.04$, $ts \leq 0.73$, $ps \geq .467$, and no interactions, $\beta_s \leq 0.10$, $ts \leq 1.47$, $ps \geq .141$. Likewise when specifically comparing February 13, February 14, and February 15 via the same ANOVA analyses, there were only null effects: difficulty was no different before ($M = 2.08$, $SD = 1.49$), during ($M = 2.00$, $SD = 1.45$), and after ($M = 2.07$, $SD = 1.49$) Valentine's Day, $F(2, 300) = 0.09$, $p = .918$, $\eta^2 = .001$; realism was no different before ($M = 5.80$, $SD = 1.22$), during ($M = 5.67$, $SD = 1.46$), and after ($M = 5.63$, $SD = 1.30$) Valentine's Day, $F(2, 300) = 0.47$, $p = .625$, $\eta^2 = .003$; confusion was no different before ($M = 1.66$, $SD = 1.27$), during ($M = 1.66$, $SD = 1.28$), and after ($M = 1.74$, $SD = 1.34$) Valentine's Day, $F(2, 300) = 0.13$, $p = .880$, $\eta^2 = .001$; and detail was no different ($M = 5.33$, $SD = 1.32$), during ($M = 5.23$, $SD = 1.52$), and after ($M = 5.22$, $SD = 1.36$) Valentine's Day, $F(2, 300) = 0.21$, $p = .814$, $\eta^2 = .001$.

Selection variables. The kinds of participants who signed up from day to day did not systematically differ (see Table S1 for summary statistics over key points of interest). All envy effects remain when controlling these variables (main effect of tense, $\beta = -0.26$, $p < .001$; main effect of distance, $\beta = -0.23$, $p = .001$; interaction, $\beta = 0.25$, $p = .004$).

Table S1. Participant characteristics in Study 2a. Percent agreement and means values.

	Feb 1-7 (n = 709)	Feb 8-13 (n = 605)	Feb 14 (n = 101)	Feb 15-21 (n = 704)	Feb 22-28 (n = 705)
1. Female	47.80%	53.70%	56.40%	56.80%	56.60%
2. White	78.70%	78.50%	78.20%	77.00%	77.90%
3. Age	35.78	36.13	37.21	35.46	36.11
4. Similar to target	42.90%	44.10%	46.50%	44.50%	43.00%
5. Have V-Day plans	23.60%	31.10%	39.60%	41.60%	44.30%
6. Currently single	34.60%	31.90%	28.70%	31.10%	31.50%
7. Income, do well (1-7)	3.98	4.11	4.29	4.08	4.13
8. Income, provide own	84.20%	82.10%	80.20%	83.40%	81.80%
9. Live alone	24.40%	19.70%	19.80%	18.60%	18.40%
10. Homeowner	47.00%	46.10%	54.50%	46.00%	43.30%
11. Eastern time zone	49.40%	54.20%	60.40%	52.10%	49.40%
12. Have children	49.40%	43.60%	34.70%	41.80%	40.00%
13. College degree	63.30%	57.85%	53.47%	59.09%	58.30%
14. Used M-Turk before	92.50%	94.70%	97.00%	96.00%	94.60%
15. Typical time to use	86.60%	85.80%	80.20%	84.80%	81.60%
16. Have other HITs open	78.00%	78.00%	76.20%	78.40%	80.60%
17. How social today (1-7)	3.20	3.10	3.54	2.92	3.09
18. How busy today (1-7)	3.86	3.52	4.04	3.40	3.51
19. Money for fun (1-7)	3.68	3.77	3.92	3.59	3.61
20. Using phone	3.53%	6.28%	6.93%	6.25%	8.94%
21. Minutes to complete	5.47	5.66	5.29	5.39	5.95

Study 2b

Attention check. Only 5.30% of participants (9 of 170) failed any of the daily attention checks. All participants are retained in analyses.

Condition checks. We found no systematic differences when repeating the analyses with the condition check variables. Over the course of the study, the task was rated as similarly easy ($M_{\text{Feb13}} = 1.94$, $SD = 1.34$; $M_{\text{Feb14}} = 1.76$, $SD = 1.07$; $M_{\text{Feb15}} = 1.93$,

$SD = 1.35$), $F(2, 168) = 1.79$, $p = .171$, $\eta^2 = .02$; similarly non-confusing ($M_{\text{Feb13}} = 1.54$, $SD = 1.16$; $M_{\text{Feb14}} = 1.45$, $SD = 1.10$; $M_{\text{Feb15}} = 1.41$, $SD = 1.05$), $F(2, 168) = 1.24$, $p = .293$, $\eta^2 = .02$; and eliciting similarly detailed mental imagery ($M_{\text{Feb13}} = 5.61$, $SD = 1.16$; $M_{\text{Feb14}} = 5.64$, $SD = 1.22$; $M_{\text{Feb15}} = 5.58$, $SD = 1.22$), $F(2, 168) = 0.24$, $p = .787$, $\eta^2 = .003$. We did observe a significant effect of day on task realism, such that Day 3 seemed more realistic than Days 1-2 ($M_{\text{Feb13}} = 5.91$, $SD = 1.06$; $M_{\text{Feb14}} = 5.94$, $SD = 1.17$; $M_{\text{Feb15}} = 6.07$, $SD = 0.95$), $F(2, 168) = 3.42$, $p = .035$, $\eta^2 = .04$.

Attrition analyses. Eighty-five percent of participants (170 of 200) completed all days. Logistic regressions predicting attrition (0 = no, 1 = yes) suggest that attrition was not significantly related to Day 1 envy ($B = -0.10$, $p = .519$), Day 1 difficulty ($B = -0.09$, $p = .659$), Day 1 realism ($B = -0.29$, $p = .177$), Day 1 confusion ($B = 0.08$, $p = .707$), Day 1 detail ($B = -0.09$, $p = .653$), sex (0 = female, 1 = male: $B = 0.16$, $p = .703$), age ($B = 0.003$, $p = .868$), ethnicity (0 = White, 1 = Other: $B = 0.01$, $p = .984$), similarity to target (0 = similar, 1 = dissimilar: $B = -0.15$, $p = .775$), own plans (0 = no, 1 = yes: $B = 0.15$, $p = .797$), or relationship status (0 = single, 1 = other: $B = 0.75$, $p = .163$).

At the end of the Day 3 study, we asked participants to re-rate the Valentine's Day and relationship items, if anything had changed over the holiday. We included the re-rating of these items for curiosity and do not discuss them further (see data file).

Studies 3a-3b-3c

Study 3a. Thirteen percent of participants (40 of 302) failed the attention check. All participants were retained in analyses. A standard principle component analysis using a Varimax rotation confirmed two and only two distinct components ($r = -.27$, $p < .001$):

The five malicious items comprised Component 1, accounting for 39.62% of total variance (each item loaded above .64). The five benign items comprised Component 2, accounting for 22.44% of total variance (each item loaded above .44).

Study 3b. We requested 300 participants, yielding 307 (46.60% female; $M_{\text{age}} = 34.88$, $SD = 11.28$). Participants were randomly assigned into past ($n = 153$) or future ($n = 154$) conditions. All participants were retained in analyses (13.03% (40 of 307) failed the attention check). The principle component analysis extracted two and only two components ($r = -.23$, $p < .001$), reflecting the malicious envy items (39.62% of total variance, each item loading above .64; $\alpha = .84$) and the benign envy items (22.45% of total variance, each item loading above .44; $\alpha = .81$).

Study 3c. We requested 400 participants, yielding 403 (46.20% female; $M_{\text{age}} = 37.11$, $SD = 11.88$). Participants were randomly assigned into past ($n = 200$) or future ($n = 203$) conditions. All participants were retained in analyses (10.70% (43 of 403) failed the attention check). The principle component analysis extracted two and only two components ($r = -.27$, $p < .001$), reflecting the malicious envy items (22.87% of total variance, each item loading above .65; $\alpha = .84$) and the benign envy items (40.47% of total variance, each item loading above .63; $\alpha = .85$).

Condition checks. There were no differences in difficulty in Study 3a ($M_{\text{Past}} = 2.06$, $SD = 1.48$ vs. $M_{\text{Future}} = 2.42$, $SD = 1.75$), Study 3b ($M_{\text{Past}} = 2.48$, $SD = 1.71$ vs. $M_{\text{Future}} = 2.37$, $SD = 1.58$), or Study 3c ($M_{\text{Past}} = 2.28$, $SD = 1.63$ vs. $M_{\text{Future}} = 2.22$, $SD = 1.54$), $ts \leq 1.95$, $ps \geq .052$, $ds \leq 0.22$; no differences in realism in Study 3a ($M_{\text{Past}} = 6.09$, $SD = 1.18$ vs. $M_{\text{Future}} = 5.99$, $SD = 1.20$), Study 3b ($M_{\text{Past}} = 5.89$, $SD = 1.27$ vs. $M_{\text{Future}} = 5.99$, $SD = 1.31$), or Study 3c ($M_{\text{Past}} = 6.05$, $SD = 1.16$ vs. $M_{\text{Future}} = 6.04$, $SD = 0.95$), $ts \leq$

0.71, $ps \geq .476$, $ds \leq 0.08$; no differences in confusion in Study 3a ($M_{\text{Past}} = 1.66$, $SD = 1.16$ vs. $M_{\text{Future}} = 1.92$, $SD = 1.46$), Study 3b ($M_{\text{Past}} = 1.89$, $SD = 1.49$ vs. $M_{\text{Future}} = 1.89$, $SD = 1.37$), or Study 3c ($M_{\text{Past}} = 1.64$, $SD = 1.13$ vs. $M_{\text{Future}} = 1.57$, $SD = 1.15$), $ts \leq 1.75$, $ps \geq .081$, $ds \leq 0.20$; no differences in detail in Study 3a ($M_{\text{Past}} = 5.86$, $SD = 1.15$ vs. $M_{\text{Future}} = 5.83$, $SD = 1.25$), Study 3b ($M_{\text{Past}} = 5.71$, $SD = 1.31$ vs. $M_{\text{Future}} = 5.82$, $SD = 1.15$), or Study 3c ($M_{\text{Past}} = 5.77$, $SD = 1.21$ vs. $M_{\text{Future}} = 5.79$, $SD = 1.05$), $ts \leq 0.80$, $ps \geq .424$, $ds \leq 0.09$; and no differences in closeness in Study 3a ($M_{\text{Past}} = 4.96$, $SD = 1.55$ vs. $M_{\text{Future}} = 4.68$, $SD = 1.72$), Study 3b ($M_{\text{Past}} = 4.58$, $SD = 1.67$ vs. $M_{\text{Future}} = 4.83$, $SD = 1.56$), or Study 3c ($M_{\text{Past}} = 4.88$, $SD = 1.53$ vs. $M_{\text{Future}} = 4.72$, $SD = 1.63$), $ts \leq 1.51$, $ps \geq .132$, $ds \leq 0.17$.

Study 4

Full pairwise comparisons for envy. Pairwise comparisons reveal that participants who re-construed an envied event as having already occurred in the past reported significantly less envy ($M = 2.80$, $SD = 1.43$) than they felt originally ($M = 4.13$, $SD = 1.18$), $F(1, 319) = 138.31$, $p < .001$, $\eta^2 = .31$ ($d = 1.02$), 95% $CI_{\text{diff}} = [1.11, 1.55]$. This drop was also observed among all conditions: Participants who re-construed the event as occurring in a distant *future* also reported significantly less envy ($M = 3.29$, $SD = 1.46$) than they felt originally ($M = 4.26$, $SD = 1.23$), $F(1, 319) = 73.46$, $p < .001$, $\eta^2 = .19$ ($d = 0.72$), 95% $CI_{\text{diff}} = [0.75, 1.19]$; and even *control* participants felt significantly less envy after reflecting on the event a second time ($M = 3.73$, $SD = 1.52$) compared to how they felt originally ($M = 4.05$, $SD = 1.48$), $F(1, 319) = 7.35$, $p = .007$, $\eta^2 = .02$ ($d = 0.21$), 95% $CI_{\text{diff}} = [0.09, 0.55]$.

Full pairwise comparisons for stress. Pairwise comparisons reveal that control participants did *not* feel significantly less stressed upon reflecting a second time ($M = 2.13$, $SD = 0.77$) compared to how they felt originally ($M = 2.20$, $SD = 0.80$), $F(1, 319) = 1.40$, $p = .238$, $\eta^2 = .004$ ($d = 0.09$), 95% $CI_{diff} = [-0.05, 0.19]$; whereas participants who took a past perspective felt significantly less stressed ($M = 1.85$, $SD = 0.67$) than they felt originally ($M = 2.32$, $SD = 0.76$), $F(1, 319) = 63.13$, $p < .001$, $\eta^2 = .17$ ($d = 0.66$), 95% $CI_{diff} = [0.35, 0.58]$. Distant-future participants again felt significantly less stressed ($M = 2.07$, $SD = 0.75$) than they felt originally ($M = 2.35$, $SD = 0.69$), $F(1, 319) = 22.68$, $p < .001$, $\eta^2 = .07$ ($d = 0.39$), 95% $CI_{diff} = [0.16, 0.39]$.

Full pairwise comparisons for self-esteem. Pairwise comparisons reveal that control participants felt no better about themselves after reflecting a second time ($M = 3.55$, $SD = 1.17$) compared to how they felt originally ($M = 3.45$, $SD = 1.17$), $F(1, 319) = 1.73$, $p = .190$, $\eta^2 = .005$ ($d = 0.09$), 95% $CI_{diff} = [-0.05, 0.26]$; while past participants felt significantly better about themselves ($M = 3.90$, $SD = 0.99$) compared to how they felt originally ($M = 3.42$, $SD = 1.04$), $F(1, 319) = 39.62$, $p < .001$, $\eta^2 = .11$ ($d = 0.47$), 95% $CI_{diff} = [0.33, 0.63]$. Distant-future participants also felt significantly better about themselves ($M = 3.49$, $SD = 1.11$) compared to how they felt originally ($M = 3.17$, $SD = 1.15$), $F(1, 299) = 17.85$, $p < .001$, $\eta^2 = .05$ ($d = 0.28$), 95% $CI_{diff} = [0.17, 0.47]$.

Full pairwise comparisons for life satisfaction. Pairwise comparisons reveal that control participants felt no more satisfied upon reflecting a second time ($M = 4.35$, $SD = 1.63$) compared to how they felt originally ($M = 4.34$, $SD = 1.59$), $F(1, 319) = 0.02$, $p = .893$, $\eta^2 < .001$ ($d = 0.01$), 95% $CI_{diff} = [-0.20, 0.23]$; while past participants felt significantly more satisfied ($M = 4.58$, $SD = 1.72$) than they felt originally ($M = 3.96$, SD

= 1.58), $F(1, 319) = 34.09$, $p < .001$, $\eta^2 = .10$ ($d = 0.38$), 95% $CI_{diff} = [0.41, 0.82]$. Again, distant-future participants also felt significantly more satisfied ($M = 4.07$, $SD = 1.71$) than they felt originally ($M = 3.77$, $SD = 1.60$), $F(1, 319) = 8.40$, $p = .004$, $\eta^2 = .03$ ($d = 0.18$), 95% $CI_{diff} = [0.10, 0.51]$.

Condition checks. There were no effects of condition on task difficulty ($M_{control} = 2.75$, $SD_{control} = 1.71$; $M_{past} = 2.73$, $SD_{past} = 1.77$; $M_{future} = 2.68$, $SD_{future} = 1.79$), $F(2, 319) = 0.04$, $p = .964$, $\eta^2 < .001$; nor task realism ($M_{control} = 4.62$, $SD_{control} = 1.65$; $M_{past} = 4.49$, $SD_{past} = 1.57$; $M_{future} = 4.68$, $SD_{future} = 1.69$), $F(2, 319) = 0.39$, $p = .679$, $\eta^2 = .002$; nor task detail ($M_{control} = 4.74$, $SD_{control} = 1.69$; $M_{past} = 4.50$, $SD_{past} = 1.56$; $M_{future} = 4.85$, $SD_{future} = 1.66$), $F(2, 319) = 1.27$, $p = .281$, $\eta^2 = .01$; nor closeness to target ($M_{control} = 5.39$, $SD_{control} = 1.59$; $M_{past} = 5.22$, $SD_{past} = 1.72$; $M_{future} = 5.19$, $SD_{future} = 1.70$), $F(2, 319) = 0.45$, $p = .641$, $\eta^2 = 0.003$; nor closeness of event ($M_{control} = 3.07$, $SD_{control} = 1.58$; $M_{past} = 2.67$, $SD_{past} = 1.67$; $M_{future} = 2.59$, $SD_{future} = 1.64$), $F(2, 319) = 2.57$, $p = .078$, $\eta^2 = .02$).

Unlike our other studies, we did observe an effect on task confusion such that control participants found their task more confusing than the other participants ($M_{control} = 2.35$, $SD_{control} = 1.60$; $M_{past} = 1.75$, $SD_{past} = 1.27$; $M_{future} = 2.13$, $SD_{future} = 1.53$), $F(2, 319) = 4.65$, $p = .010$, $\eta^2 = .03$. Critically, task confusion was unrelated to control participants' envy and wellbeing at Time 1 ($rs \leq .12$, $ps \geq .248$), at Time 2 ($rs \leq .15$, $ps \geq .143$), and their degree of change over time ($rs \leq .14$, $ps \geq .155$), and all of our main analyses hold when controlling task confusion (main effects of condition, $F_s \geq 15.64$, $ps \leq .001$, $\eta_s^2 \geq .05$; main effects of time, $F_s \leq 2.88$, $ps \geq .058$, $\eta_s^2 \leq .02$; critical interactions, $F_s \geq 4.80$, $ps \leq .009$, $\eta_s^2 \geq .03$).