

The Centrality of Reasoning in Moral Judgments: First- and Third-Party Evaluations of Cheating Supplementary Online Materials (SOM)

In these supplementary materials, we report details about materials, data analyses, additional findings, and one supplementary study not included in the main text. For additional information, visit the Open Science Framework page for this project: <https://osf.io/fn8ys/>

Research Setting

The university in which this research was conducted is a non-honor code secular research university. The university is in Northern California and has an undergraduate enrollment of over 18,000 students. The academic policy outlines cheating as: “Cheating is defined as fraud, deceit, or dishonesty in an academic assignment, or using or attempting to use materials, or assisting others in using materials, which are prohibited or inappropriate in the context of the academic assignment in question.” (UCSC Division of Undergraduate Education, 2019).

Study 1: First-Party Judgments About Past Cheating Events

Method

Participants

As mentioned in the main text, participants were college undergraduates ($N = 60$, 43 women, 16 men, 1 non-binary; $M_{\text{age}} = 19.63$; $SD_{\text{age}} = 1.40$). About half of participants were first-generation college students (45%) and learned English as their first language (53%). Most participants (63%) were born in the United States.

Materials and Procedure

In the main text, we focused on the part of the interview in which participants discussed a personal past cheating event. Before this, participants were interviewed about hypothetical scenarios. These data were omitted from the main text due to space constraints. Here, we provide further details about the methods and data analyses related to these hypothetical scenarios.

ORIGINAL	STEVEN
It was expected (a) that children's judgments of moral events would be based on factors intrinsic to actions and not on the presence of a rule, and (b) that their judgments of social conventional events would be based on whether or not they are in violation of a rule. Second, it was hypothesized that moral events would produce different types of social interactions from those produced by social conventional events.	One proposal would be that children's judgments of moral events would be based on aspects intrinsic to actions and not on the existence of a rule, and another is that their judgments of social conventional events would be based on whether or not they are in violation of a rule. Secondly, it is expected that moral situations would produce different types of social interactions from those produced by social conventional situations.

Figure S1. Example similar pair of text shown to participants in Study 1.

In the first part of the interview, participants were presented with a hypothetical scenario involving potential cheating (Waltzer & Dahl, 2021). In the scenario, a student reads a source

text while working on a graded assignment for a class. The student then turns in work that is either *highly similar* to ($n = 30$) or *different* from ($n = 30$) the source text (Figure S1). Participants were randomly assigned to view similar or different pairs of text. After reading the scenario, participants were asked to judge whether the student's action was OK or not OK, rate it on a scale from 0 ("really bad") to 10 ("really good"), indicate whether they thought the student had plagiarized, and decide whether the act would be OK or not OK had there been no rules against plagiarism.

Data Analysis

We used Generalized Linear Mixed-Effects Models (GLMMs, Hox, 2010) to predict participants' judgments. Fixed effects included participants' perspective (first-party, third-party) and perception of whether the act constituted cheating. Because all participants responded to both hypothetical events and personal events, participants were included as random intercepts in the models. Hypotheses were tested using likelihood ratio tests.

Results

Do Students Evaluate Their Own Acts More Favorably?

There was no interaction between perspective (first-party, third-party) and perception (cheating, not cheating) in predicting *judgments that the act was okay*, $D(1) = 0.65$, $p = .420$. There was an interaction between perspective and perception in predicting *evaluative ratings* of the act, $D(1) = 6.75$, $p = .009$. Follow-up analyses revealed that when they perceived the acts as *not cheating*, participants were more likely to rate *third-party* acts more positively (7.91) compared to *first-party* acts (6.25), $D(1) = 9.45$, $p = .002$. However, they did not rate acts differently when they perceived them as *cheating* ($M = 2.49$), $D(1) = 0.88$, $p = .347$.

Study 2: First- and Third-Party Judgments of Cheating Events

Method

Participants

As mentioned in the main text, participants were college undergraduates ($N = 60$, 40 women, 18 men, 2 non-binary; $M_{\text{age}} = 19.85$; $SD_{\text{age}} = 1.98$). Half of the participants (50%) were first-generation college students, 48% learned English as their first language, and 73% were born in the United States. A requirement for participating was that students had not previously participated in Study 1 or related studies conducted by our lab.

Materials and Procedure

In Study 2, we developed new scenarios based on the real events described in Study 1. To generate these scenarios, we used descriptive information about the event as well as the reasons mentioned by the first-party respondent. To ensure anonymity of participants, protagonists were given pseudonyms in the scenarios. To create the pseudonyms, we searched for names that would resemble those of the original first-party respondents in gender and cultural-linguistic origin. Specifically, we used demographic information about participants' country of origin, or that of their family, to search online databases that listed the most common names of people in those countries. We then selected one of the top ten most common names listed for the participant's gender (either male, female, or unisex names were used). To choose among the top ten names, we selected names so that they were distinct from those in the other scenarios.

Differences Between Study 1 and 2 Reasoning

As mentioned in the main text, third-party respondents generally mentioned more reasons than first-party respondents. We believe that these differences were primarily driven by methodological differences between the two studies as opposed to a substantive finding, evidenced by the common patterns that emerge between the two studies.

To build on the role of contextual circumstances that prevailed in participant reasoning from Study 1, the interviewing procedure was refined for Study 2 to capture more detail. One such change was that the prompting process for reasoning was substantially more extensive, and participants were asked to provide much more detailed reasoning in Study 2. Interviews were also transcribed with more detail, including every exact statement made by participants. This provided a much more rich set of data than what was available for Study 1.

While not a planned methodological change, due to the design of Study 1, which included participants both evaluating and reasoning about decisions “back then” and “now”, the reasoning data presented in the paper was the second time participants in Study 1 were prompted to reason about their acts. Participants in Study 2 were only provided to evaluate and reason about acts once. As such, the authors believe that in Study 1 the reasons for “back then” and “now” evaluations might have overlapped, leading participants to provide less reasoning when judging their acts the second time.

Data Analysis

We used Generalized Linear Mixed-Effects Models (GLMMs, Hox, 2010) to assess the predictive relations between first-party and third-party responses. GLMMs were an ideal choice in this case because GLMMs are versatile statistical modeling tools capable of predicting an outcome variable while handling a wide range of response distributions (e.g., dichotomous judgments, continuous ratings), missing responses, and hierarchical structures (e.g., repeated measures within subjects). We used first-party responses (e.g., whether a student thought their act was cheating) to predict aggregated third-party responses to the same scenarios (e.g., how many third-party respondents thought the act was cheating). We ran a GLMM for each response variable: perception of cheating, judgment of whether the act was okay, evaluative rating, and mention of each coded reason category. Each model included random intercepts for (Study 2) participants. Hypotheses were tested using likelihood ratio tests on the change in model deviance (D) accounted for by the predictor.

Results

Predictive Relations

First- and third-party perceptions of described cheating events. Third-party perceptions of cheating aligned with first-party perceptions. Third-party respondents perceived the acts as cheating more often when the first-party respondent considered it cheating (74 %, 95% CI: [69%, 78%]) than when they did not (53%, 95% CI: [44%, 62%]), $D(1) = 17.06$, $p < .001$. Table S1 summarizes the first- and third-party responses broken down by each scenario.

First- and third-party evaluations of described cheating events. First-party and third-party respondents' judgments of whether the act was okay aligned in the majority of cases. Third-party respondents were more likely to judge the act as okay when the first-party respondent did too (45% of third-party responses saying okay, 95% CI: [37%, 54%]) than when the first-party respondent said it was not okay (35% of third-party responses saying okay, 95% CI: [30%, 40%]), $D(1) = 5.15$, $p = .023$.

Table S1

Summary of First- and Third-Party Responses to All 60 Scenarios

name	cheating perception			okay judgment			evaluative rating		
	1st	3rd	3rd 95%CI	1st	3rd	3rd 95%CI	1st	3rd	3rd 95%CI
Adriena	1	0.75	[0.36, 0.96]	0	0.13	[0.01, 0.53]	3.5	2.94	[2.12, 3.76]
Alexander	1	1.00	[0.60, 1.00]	0	0.00	[0.00, 0.40]	1.3	1.38	[0.38, 2.37]
Allie	1	0.88	[0.47, 0.99]	0	0.13	[0.01, 0.53]	3.4	3.53	[1.11, 5.95]
Alyssa	0	0.75	[0.36, 0.96]	1	0.25	[0.04, 0.64]	3.5	4.19	[2.61, 5.77]
Anastasia	1	0.88	[0.47, 0.99]	1	0.13	[0.01, 0.53]	6	2.63	[0.58, 4.67]
Andrea	1	0.50	[0.22, 0.78]	0	0.88	[0.47, 0.99]	1.5	5.11	[4.61, 5.61]
Angela	1	0.88	[0.47, 0.99]	0	0.38	[0.10, 0.74]	3.5	3.63	[2.31, 4.94]
Anna	1	0.88	[0.47, 0.99]	0	0.38	[0.10, 0.74]	NA	4.00	[2.74, 5.26]
Ashley	1	0.88	[0.47, 0.99]	0	0.50	[0.22, 0.78]	NA	4.63	[3.86, 5.39]
Bhadra	0	1.00	[0.60, 1.00]	1	0.13	[0.01, 0.53]	6	3.33	[1.49, 5.17]
Camila	0	0.13	[0.01, 0.53]	1	0.50	[0.22, 0.78]	5.5	4.63	[3.16, 6.09]
Catherine	1	0.88	[0.47, 0.99]	0	0.13	[0.01, 0.53]	1.3	3.13	[1.68, 4.57]
Chen	1	1.00	[0.60, 1.00]	0	0.00	[0.00, 0.40]	0.9	3.19	[2.03, 4.35]
Chip	0	0.13	[0.01, 0.53]	1	0.88	[0.47, 0.99]	8	5.31	[2.97, 7.66]
Cho	1	0.88	[0.47, 0.99]	0	0.25	[0.04, 0.64]	0	3.28	[1.72, 4.84]
Christina	0	0.00	[0.00, 0.40]	1	0.88	[0.47, 0.99]	4.7	6.44	[4.96, 7.91]
Daniel	1	0.13	[0.01, 0.53]	0	0.63	[0.26, 0.90]	4	4.88	[3.66, 6.09]
Danilo	0	0.50	[0.22, 0.78]	1	0.50	[0.22, 0.78]	10	4.19	[2.96, 5.41]
Diana	1	1.00	[0.60, 1.00]	0	0.13	[0.01, 0.53]	3.6	2.94	[1.65, 4.23]
Fatima	1	1.00	[0.60, 1.00]	0	0.13	[0.01, 0.53]	1	1.94	[0.63, 3.25]
Fei	1	1.00	[0.60, 1.00]	0	0.00	[0.00, 0.40]	4	2.44	[1.31, 3.56]
Felicia	1	0.63	[0.26, 0.90]	0	0.38	[0.10, 0.74]	2.5	4.29	[2.47, 6.10]
Gabriel	0	0.63	[0.26, 0.90]	1	0.50	[0.22, 0.78]	8.1	3.66	[2.44, 4.89]
Gao	1	0.63	[0.26, 0.90]	1	0.75	[0.36, 0.96]	4.8	5.00	[3.09, 6.91]
Gloria	1	0.75	[0.36, 0.96]	0	0.50	[0.22, 0.78]	3.6	4.50	[2.83, 6.17]
Grace	1	1.00	[0.60, 1.00]	0	0.13	[0.01, 0.53]	2.1	3.69	[2.44, 4.93]
Isabella	1	1.00	[0.60, 1.00]	1	0.00	[0.00, 0.40]	4.6	2.88	[1.81, 3.94]
Jackie	0	0.88	[0.47, 0.99]	1	0.50	[0.22, 0.78]	3.1	3.66	[2.08, 5.24]
Jasmine	1	0.88	[0.47, 0.99]	1	0.50	[0.22, 0.78]	1.5	5.00	[3.64, 6.36]
Jeff	1	0.75	[0.36, 0.96]	0	0.25	[0.04, 0.64]	0.3	5.44	[4.15, 6.73]
Jessica	1	0.38	[0.10, 0.74]	0	0.50	[0.22, 0.78]	4.3	5.63	[3.54, 7.71]
Jimena	1	0.38	[0.10, 0.74]	0	0.50	[0.22, 0.78]	5	5.13	[3.75, 6.50]
Joel	1	0.88	[0.47, 0.99]	0	0.13	[0.01, 0.53]	4.5	2.94	[1.59, 4.29]
Josefina	1	0.25	[0.04, 0.64]	0	0.88	[0.47, 0.99]	0	6.88	[4.85, 8.90]
Lin	1	0.88	[0.47, 0.99]	0	0.50	[0.22, 0.78]	2	3.98	[2.24, 5.71]
Luciana	1	1.00	[0.60, 1.00]	0	0.00	[0.00, 0.40]	0	2.38	[0.97, 3.78]
Luis	0	0.25	[0.04, 0.64]	0	0.88	[0.47, 0.99]	8	5.00	[4.55, 5.45]
Maggie	1	1.00	[0.60, 1.00]	0	0.00	[0.00, 0.40]	4	2.29	[1.20, 3.37]
Mariana	1	0.88	[0.47, 0.99]	0	0.38	[0.10, 0.74]	0.3	3.94	[3.09, 4.79]
Martina	1	0.63	[0.26, 0.90]	0	0.88	[0.47, 0.99]	0.1	5.19	[3.55, 6.83]
Mary	1	0.50	[0.22, 0.78]	1	0.63	[0.26, 0.90]	9.5	5.13	[3.04, 7.21]
Mateo	1	1.00	[0.60, 1.00]	0	0.00	[0.00, 0.40]	0	2.13	[0.99, 3.26]
Mei	0	0.25	[0.04, 0.64]	1	0.50	[0.22, 0.78]	7	5.13	[3.83, 6.42]
Michelle	0	0.63	[0.26, 0.90]	1	0.38	[0.10, 0.74]	5	4.31	[2.85, 5.78]
Minji	0	0.38	[0.10, 0.74]	1	0.75	[0.36, 0.96]	6.3	5.76	[4.47, 7.06]

Ofelia	0	0.63	[0.26, 0.90]	1	0.25	[0.04, 0.64]	8	3.75	[1.04, 6.46]
Peter	1	0.75	[0.36, 0.96]	0	0.25	[0.04, 0.64]	2.9	3.88	[2.50, 5.25]
Rachel	1	0.63	[0.26, 0.90]	0	0.63	[0.26, 0.90]	2.4	5.44	[4.00, 6.88]
Rosa	1	0.38	[0.10, 0.74]	0	0.63	[0.26, 0.90]	3	5.75	[3.70, 7.80]
Samantha	1	0.75	[0.36, 0.96]	0	0.38	[0.10, 0.74]	2.2	4.69	[3.19, 6.19]
Samuel	1	0.50	[0.22, 0.78]	0	0.50	[0.22, 0.78]	2	4.88	[2.28, 7.47]
Santiago	1	0.88	[0.47, 0.99]	0	0.25	[0.04, 0.64]	3	3.50	[1.77, 5.23]
Sarah	1	0.75	[0.36, 0.96]	0	0.38	[0.10, 0.74]	3	3.94	[2.09, 5.79]
Teresa	1	0.88	[0.47, 0.99]	0	0.13	[0.01, 0.53]	3.1	3.13	[1.09, 5.16]
Valentina	1	0.50	[0.22, 0.78]	1	0.50	[0.22, 0.78]	5	5.31	[3.25, 7.37]
Valeria	1	0.38	[0.10, 0.74]	0	0.63	[0.26, 0.90]	3.5	5.69	[3.60, 7.77]
Wei	0	1.00	[0.60, 1.00]	1	0.13	[0.01, 0.53]	7.5	3.13	[1.83, 4.42]
Ying	0	0.88	[0.47, 0.99]	0	0.00	[0.00, 0.40]	3	3.38	[2.64, 4.11]
Zahra	1	0.25	[0.04, 0.64]	0	0.63	[0.26, 0.90]	0	5.41	[4.15, 6.67]
Zoe	1	0.75	[0.36, 0.96]	0	0.38	[0.10, 0.74]	1.9	3.63	[2.29, 4.96]

Note. First-party responses are a single value from the participant who engaged in the action. Third-party responses and 95% confidence intervals are based on averages across eight participants who read and responded to each scenario.

Reasons Mentioned by First- and Third-Party Respondents. As we saw in the main text, first-party reasoning was similar to third-party reasoning. Not only that, but when we examined predictive relations, we also found that first-party reasoning in each scenario predicted third-party reasoning in those same scenarios. Overall, third-party respondents were twice as likely to use a reason to support their evaluation of a given scenario if the first-party respondent had also given the same reason (32%, vs. only 16% when they did not), $D(1) = 46.80$, $p < .001$. When reasoning categories were analyzed separately, this predictive relation held for *labeling as academic misconduct* (e.g., whether the act constitutes cheating/plagiarism), *learning*, *affect others*, *affect agent*, *honesty*, and *rules*. However, *features of the assignment* was the opposite case, where third-party respondents were less likely to mention it when first-party respondents had mentioned it for a given scenario.

Table S2 and Table S3 provide a detailed breakdown of the reasons given by first- and third-party respondents in favor of (Table S2) or against (Table S3) the action in each of the 60 scenarios. First-party responses reflect one response per scenario, and third-party responses are the averages across all eight participants who read and responded to each scenario.

Differences Between Types of Acts

An additional goal of Study 2 was to assess whether richer contextual information would inform students' evaluations of cheating acts. Thus, we compared the responses to hypothetical scenarios from Study 1 (which lacked context) to judgments about the new contextualized third-party scenarios and first-party judgments of personal events.

In the main text, we presented comparisons only between first- and third-party evaluations of the realistic cheating events. Recall, though, that participants in Study 1 also read and evaluated hypothetical scenarios involving text pairs. Here, we present analyses designed to compare perceptions and evaluations across all four act types (third-party similar text [Study 1], third-party different text [Study 1], first-party judgment of real event [Study 1], and third-party judgment of real event [Study 2]).

Table S2

Detailed Summary of First- and Third-Party Reasons Given for Why the act was Okay

name	academic misconduct		affect agent		affect others		effort		evaluative label		fairness		features of assignment		honesty		learning		on purpose		rules		stealing	
	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd
Adriana	0.00	0.13		0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00
Alexander	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Allie	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00
Alyssa		0.63		0.38	0.00	0.00	0.00	0.38	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.38		0.13	0.00
Anastasia	0.00	0.13		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Andrea	0.00	0.38	0.00	0.25	0.00	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00
Angela	0.00	0.25	0.00	0.25		0.13	0.00	0.13	0.00	0.00	0.00	0.50	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.00	0.00	0.00
Anna	0.00	0.25		0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.13	0.00	0.00		0.13	0.00	0.00	0.00
Ashley	0.00	0.13	0.00	0.38	0.00	0.13	0.00	0.25	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00
Bhadra		0.13		0.25	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00
Camila		0.50	0.00	0.13	0.00	0.13		0.13	0.00	0.13		0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.00	0.00
Catherine	0.00	0.13		0.13	0.00	0.13		0.13	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.00		0.00	0.00	0.25	0.00	0.00	0.00	0.00
Chen	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00
Chip		0.63		0.50	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.50	0.00	0.25		0.63	0.00	0.00		0.38	0.00	0.13	0.00
Cho	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00
Christina		0.75		0.38	0.00	0.25	0.00	0.63	0.00	0.00		0.00	0.00	0.25		0.00	0.00	0.63	0.00	0.00		0.75		0.00
Daniel	0.00	0.63	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.75		0.25	0.00	0.13
Danilo	0.00	0.38		0.50	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.50	0.00	0.00		0.38	0.00	0.00
Diana	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fatima	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fei	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.00
Felicia	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.25
Gabriel		0.25	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.13	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.25	0.00	0.00
Gao	0.00	0.50	0.00	0.50		0.50	0.00	0.13		0.13	0.00	0.13		0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00
Gloria	0.00	0.25		0.38	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.38		0.00	0.00	0.00	0.25	0.00	0.25	0.00	0.13	0.00
Grace		0.25		0.13	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Isabella	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Jackie	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00
Jasmine	0.00	0.38	0.00	0.38	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00

Jeff	0.00	0.50	0.00	0.13	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.25	0.00	0.00	
Jessica	0.00	0.63	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.13	
Jimena	0.00	0.50	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.38	0.00	0.25	0.00	0.13	0.00	0.00	
Joel	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00		0.00	0.00	0.00	
Josefina	0.00	0.63	0.00	0.00	0.00	0.88	0.00	0.25	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.38	0.00	0.00	
Lin	0.00	0.13	0.00	0.50	0.00	0.13	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	
Luciana	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Luis		0.75	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.88	0.00	0.00		0.75	0.00	0.25	0.00	0.38	0.00	0.00	
Maggie	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	
Mariana	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	
Martina	0.00	0.25	0.00	0.50	0.00	0.50	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.88	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.25	0.00	0.00
Mary		0.38	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.38	0.00	0.25	0.00	0.00	0.13		0.25		0.13	
Mateo	0.00	0.00		0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Mei	0.00	0.00		0.50	0.00	0.00		0.38	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.13	0.00	0.00	
Michelle	0.00	0.38	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.25	0.00	0.13	
Minji	0.00	0.50	0.00	0.13	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.13	0.00	0.25	0.00	0.00	0.00	0.50	0.00	0.00	0.38	0.00	0.13	
Ofelia		0.38		0.00	0.00	0.13	0.00	0.13	0.00	0.00		0.00	0.00	0.13		0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.13	
Peter	0.00	0.13		0.13	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	
Rachel	0.00	0.25	0.00	0.25	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.25	0.00	0.00
Rosa	0.00	0.75	0.00	0.25	0.00	0.00	0.00	0.50	0.00	0.13	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.50	0.00	0.00	0.13	0.00	0.25	0.00
Samantha	0.00	0.00	0.00	0.38	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	
Samuel	0.00	0.38	0.00	0.25	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.50	0.00	0.13	0.00	0.38	0.00	0.00
Santiago		0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.00	
Sarah	0.00	0.25	0.00	0.38	0.00	0.13	0.00	0.38	0.00	0.13	0.00	0.00	0.00	0.38	0.00	0.13	0.00	0.13	0.00	0.00	0.25	0.00	0.13	
Teresa	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Valentina		0.63		0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.00	
Valeria	0.00	0.63	0.00	0.25	0.00	0.00	0.00	0.38	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	
Wei		0.38	0.00	0.13	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	
Ying	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.00	
Zahra	0.00	0.88		0.25	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	
Zoe	0.00	0.00	0.00	0.50	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.13	

Note. Third-party reasons reflect proportion of eight respondents who mentioned the type of reason.

Table S3

Detailed Summary of First- and Third-Party Reasons Given for Why the act was not Okay

	academic misconduct		affect agent		affect others		effort		evaluative label		fairness		features of assignment		honesty		learning		on purpose		rules		stealing	
name	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd	1st	3rd
Adriena		0.25	0.00	0.25	0.00	0.25	0.00	0.50	0.00	0.00	0.00	0.50	0.00	0.00		0.75	0.00	0.38	0.00	0.00	0.00	0.25	0.00	0.25
Alexander	0.00	0.75		0.50	0.00	0.63	0.00	0.25		0.13	0.00	0.25	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.38	0.00	0.00
Allie		0.63	0.00	0.25	0.00	0.00		0.00	0.00	0.00		0.25	0.00	0.13		0.38	0.00	0.25	0.00	0.00	0.00	0.25		0.25
Alyssa	0.00	0.50	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.25	0.00	0.00	0.00
Anastasia	0.00	0.88	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.63	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.00	0.13	0.00	0.63	0.00
Andrea	0.00	0.25	0.00	0.13	0.00	0.38	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00		0.13	0.00
Angela	0.00	0.38		0.63		0.38	0.00	0.13	0.00	0.13	0.00	0.38	0.00	0.13	0.00	0.25		0.00	0.00	0.13		0.63	0.00	0.00
Anna	0.00	0.38	0.00	0.38	0.00	0.38	0.00	0.50	0.00	0.25	0.00	0.13	0.00	0.13	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.38	0.00	0.00
Ashley	0.00	0.63	0.00	0.38	0.00	0.25	0.00	0.38	0.00	0.00	0.00	0.38	0.00	0.13	0.00	0.13	0.00	0.63	0.00	0.00	0.00	0.25	0.00	0.13
Bhadra	0.00	0.38	0.00	0.38	0.00	0.38	0.00	0.25	0.00	0.13	0.00	0.88	0.00	0.13	0.00	0.13	0.00	0.25	0.00	0.25	0.00	0.25	0.00	0.00
Camila	0.00	0.75	0.00	0.25	0.00	0.00		0.38	0.00	0.00	0.00	0.25	0.00	0.25	0.00	0.25	0.00	0.00	0.00	0.00	0.25	0.00	0.13	0.00
Catherine		0.88		0.50	0.00	0.13	0.00	0.13		0.25	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.25	0.00	0.38	0.00	0.38
Chen		0.88	0.00	0.50	0.00	0.13	0.00	0.50		0.13	0.00	0.13	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.38		0.25		0.25
Chip	0.00	0.63	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.13	0.00
Cho	0.00	0.50		0.63	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.63	0.00	0.13	0.00	0.38	0.00	0.25	0.00	0.13	0.00	0.38	0.00	0.00
Christina	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.25	0.00
Daniel	0.00	0.63	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.25		0.50	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.13	0.00	0.25		0.13
Danilo	0.00	0.50	0.00	0.50	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.25	0.00	0.13	0.00	0.13	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.00
Diana	0.00		0.00	0.38	0.00	0.13	0.00	0.75	0.00	0.13	0.00	0.38	0.00	0.25	0.00	0.38	0.00	0.25	0.00	0.00		0.38	0.00	0.38
Fatima	0.00	0.88	0.00	0.38	0.00	0.13	0.00	0.00		0.13	0.00	0.50	0.00	0.13		0.50	0.00	0.38	0.00	0.00	0.00	0.63		0.00
Fei		0.88	0.00	0.25	0.00	0.38		0.50	0.00	0.00		0.25	0.00	0.00	0.00	0.13	0.00	0.25	0.00	0.50	0.00	0.38		0.25
Felicia	0.00	0.75	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.50	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.13	0.00	0.50		0.38
Gabriel	0.00	0.75	0.00	0.13	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.38	0.00	0.38
Gao	0.00	0.38	0.00	0.25	0.00	0.50	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.00	0.00	0.00
Gloria		0.38		0.13	0.00	0.00		0.13	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00		0.13	0.00	0.00	0.00	0.25	0.00	0.25
Grace	0.00	0.63	0.00	0.88	0.00	0.25	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.13	0.00	0.75	0.00	0.00	0.00	0.50	0.00	0.13
Isabella		0.75		0.25	0.00	0.88	0.00	0.00	0.00	0.25	0.00	0.38	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.13		0.38	0.00
Jackie		0.63	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00

Jasmine		0.63	0.00	0.25	0.00	0.13	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.25		0.00		0.13	0.00	0.00	0.00	0.25	0.00	0.25
Jeff		0.63	0.00	0.25	0.00	0.00	0.00	0.38		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.25	0.00	0.13	
Jessica	0.00	0.63	0.00	0.38	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	
Jimena	0.00	0.50	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.25	0.00	0.13	0.00	0.38	0.00	0.25	0.00	0.50	0.00	0.00
Joel	0.00	0.50	0.00	0.25	0.00	0.13	0.00	0.25	0.00	0.25	0.00	0.50	0.00	0.00	0.00	0.38	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.50
Josefina	0.00	0.25	0.00	0.13		0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lin		0.75		0.63	0.00	0.13	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00		0.38	0.00	0.00	
Luciana	0.00	0.75		0.38	0.00	0.38	0.00	0.13	0.00	0.13	0.00	0.63	0.00	0.13	0.00	0.25	0.00	0.25	0.00	0.38	0.00	0.50	0.00	0.00
Luis	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	
Maggie	0.00	0.75		0.75	0.00	0.25	0.00	0.25	0.00	0.25	0.00	0.25	0.00	0.00	0.00	0.25		0.38	0.00	0.00	0.00	0.38	0.00	0.25
Mariana		0.63		0.38	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.38	0.00	0.00	0.00	0.00	0.63	0.00	0.25		0.25	0.00	0.00	
Martina		0.38	0.00	0.13	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Mary	0.00	0.38	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	
Mateo	0.00			0.75	0.00	0.13	0.00	0.25	0.00	0.13	0.00	0.50	0.00	0.13	0.00	0.50		0.63	0.00	0.00		0.50		0.00
Mei	0.00	0.13	0.00	0.25	0.00	0.25	0.00	0.13	0.00	0.13	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	
Michelle	0.00	0.88	0.00	0.25	0.00	0.13	0.00	0.38	0.00	0.13		0.63	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.00	0.38	0.00	0.63	
Minji	0.00	0.38	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	
Ofelia	0.00	0.75	0.00	0.13	0.00	0.13	0.00	0.38	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.13	0.00	0.38	0.00	0.13
Peter	0.00	0.88	0.00	0.63	0.00	0.75	0.00	0.25	0.00	0.25		0.25	0.00	0.13	0.00	0.25	0.00	0.25	0.00	0.00	0.50	0.00	0.00	
Rachel		0.13	0.00	0.25	0.00	0.13	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00	
Rosa		0.38	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.13		0.25	0.00	0.00	
Samantha	0.00	0.50		0.13	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00		0.50	0.00	0.13	0.00	0.63	0.00	0.13
Samuel		0.50	0.00	0.38	0.00	0.13	0.00	0.38	0.00	0.13		0.25	0.00	0.25	0.00	0.13	0.00	0.25	0.00	0.00	0.38	0.00	0.00	
Santiago	0.00	0.63	0.00	0.00	0.00	0.13	0.00	0.13	0.00	0.00		0.50	0.00	0.13	0.00	0.50	0.00	0.13	0.00	0.00	0.25	0.00	0.38	
Sarah	0.00	0.63	0.00	0.38	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.38	0.00	0.13		0.38	0.00	0.25	0.00	0.00	0.50	0.00	0.25	
Teresa	0.00	0.75		0.25		0.00		0.00	0.00	0.13	0.00	0.25	0.00	0.50	0.00	0.38		0.00	0.00	0.13	0.00	0.50	0.00	0.25
Valentina		0.88	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.38	0.00	0.00	0.00	0.13	0.00	0.63	0.00	0.38
Valeria		0.38	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.13		0.25
Wei	0.00	0.75	0.00	0.25	0.00	0.00	0.00	0.38	0.00	0.13	0.00	0.13	0.00	0.13	0.00	0.38	0.00	0.25	0.00	0.00	0.50	0.00	0.38	
Ying	0.00		0.00	0.25	0.00	0.00	0.00	0.13	0.00	0.25	0.00	0.50	0.00	0.00	0.00	0.25	0.00	0.13	0.00	0.00		0.50	0.00	0.25
Zahra		0.38		0.13	0.00	0.13	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	
Zoe	0.00	0.50		0.63	0.00	0.38	0.00	0.38	0.00	0.13		0.13	0.00	0.25	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.50	0.00	0.13

Note. Third-party reasons reflect proportion of eight respondents who mentioned the type of reason.

Perceptions of different types of acts. Overall, perceptions of whether acts were cheating significantly differed by the type of act, $D(3) = 41.56, p < .001$. *Third-party different texts* were perceived as cheating significantly less (20%) than all others: *third-party similar texts* (93%), *first-party judgments of real events* (75%), and *third-party judgments of real events* (69%), all $ps < .001$. *Similar texts* were also perceived as cheating significantly more often than *different texts* and *third-party judgments of real events*, $ps < .014$, but not *first-party judgments of real events*, $p = .056$.

Evaluations of different types of acts. We next focused specifically on the events perceived as *cheating*, modeling participants' evaluations only on this subset of data. When participants perceived an event *as cheating*, their okay judgments significantly differed based on the type of act (similar text, different text, first-party real event, third-party real event), $D(3) = 14.98, p = .002$. Post-hoc analyses revealed that *different texts* were judged as more okay (50%) than *first-party real events* (13%), $p = .034$.

Participants' evaluative ratings also significantly differed based on the type of act, $D(3) = 27.63, p < .001$. Crucially, suggesting that additional contextual details softened participants' evaluations, post-hoc analyses revealed that *third-party real events* were rated significantly more positively ($M = 3.32$) than *similar texts* (1.71), $p < .001$. *Different texts* were also rated more positively (4.63) than *first-party real events* (2.68) and *similar texts*, $ps < .012$.

Study S1: Using Open-Ended Labels to Develop Measures for Study 3

Due to space constraints, Study S1 is not described in the main text. In Study S1, we presented students with cheating scenarios from Study 2 to generate measures for Study 3.

Method

Participants

Undergraduate students ($N = 44$, 31 women, 9 men, 1 non-binary, 3 not reporting; $M_{\text{age}} = 20.44$; $SD_{\text{age}} = 1.67$) were recruited using a university subject pool and compensated for their time with class credit. About half of participants were first-generation college students (51%) and learned English as their first language (54%). Most participants (72%) were born in the U.S. A requirement for participating was that students had not previously participated in Study 1, Study 2, or related studies conducted by our lab.

Materials and Procedure

Participants took an online survey via Qualtrics. To develop scenarios, the 5 stories with the highest and lowest average ratings from Study 2 were chosen, yielding 10 scenarios. The 5 most frequently mentioned reasons in Study 2 (learning, effort, rule, stealing, deceit) were used to manipulate every scenario either *for* or *against* cheating (for an overview of reasoning coding in Study 2, see main paper). Participants were randomly presented 1 out of the 10 versions of each scenario, resulting in 10 scenarios per participant. Below is an example scenario to highlight the difference between a *for* and *against* manipulation used in Study S1:

Example Scenario Manipulated *For Cheating*.

“When Rachel was in her second year of college, she was in an environmental studies class. In this class, students were sometimes given worksheet assignments for their section. The worksheets included fill-in-the-blank questions, and the worksheets were graded for completion. After completing the worksheet, they would go over the answers

to the worksheet in class. Rachel and her friend would sometimes work on these assignments at the same time together. One day, Rachel's friend forgot to do the assignment because she had many commitments and did not have time. Rachel agreed to help her by sending her friend pictures of her completed assignment. Rachel's friend used the pictures in order to fill out her own worksheet and hand it in. *Rachel tried very hard and put in her best effort while sending the pictures to her friend.*"

Example Scenario Manipulated Against Cheating.

"When Rachel was in her second year of college, she was in an environmental studies class. In this class, students were sometimes given worksheet assignments for their section. The worksheets included fill-in-the-blank questions, and the worksheets were graded for completion. After completing the worksheet, they would go over the answers to the worksheet in class. Rachel and her friend would sometimes work on these assignments at the same time together. One day, Rachel's friend forgot to do the assignment because she had many commitments and did not have time. Rachel agreed to help her by sending her friend pictures of her completed assignment. Rachel's friend used the pictures in order to fill out her own worksheet and hand it in. *Rachel did not try hard at all and did not put in any effort in sending the pictures to her friend.*"

After reading the scenario, participants were asked: "If you had to choose one word or phrase to describe [character's] action, what would it be?". They were then asked to explain "why do you think what [character] did was [answer to first question]?". Finally, they were asked to evaluate the character's action as OK or Not OK and to provide an evaluative rating on a Likert-type scale from really bad (0) – really good (10).

Data Coding

Two coders independently classified statements based on a coding scheme for *short phrase labels*. This categorization scheme was developed through a mix of bottom-up and top-down approaches. Members of the research team reviewed a subset of the data to deductively generate categories that captured common types of responses (bottom-up approach). Meanwhile, theoretically relevant categories (e.g., *clear negative moral*) were also added (top-down approach). Reliability was assessed by computing Cohen's kappa scores (κ) for both coders' categorization of a random subset of the data (20% of all responses) (McHugh, 2012).

Participant responses to the first question were organized into groups of similar words, generating eight different categories: *trait (positive)*, *trait (negative)*, *clear moral (positive)*, *clear moral (negative)*, *academic misconduct*, *acceptable*, and *conflicted*. Examples of *trait* words include "smart" and "resourceful" (positive), as well as "foolish" and "lazy" (negative). Example of *clear moral* words include "right" and "fair" (positive), as well as "bad" and "unethical" (negative). Words marked as *acceptable* included "okay" and "fine", while words marked as *conflicted* included "ok but not ideal" and "maybe wrong". Agreement was high ($\kappa > .95$).

Data Analysis

One main goal for this study was to identify the words and phrases students use to describe cheating acts. We created a word cloud to identify descriptive trends by reducing all the

[illegible]

Results

Overall, 25% of participants described the scenarios as *acceptable*, 16% described it as *academic misconduct*, 11% described it in terms of a *negative trait*, 17% described it in terms of a *positive trait*, 22% described it as *morally wrong*, 2% described it as *morally right*, and 6% were *conflicted* in their short phrase answer. See Figure S2 for a word cloud of the most common standardized phrases.

Method

As mentioned in the main text, we recruited 98 undergraduate students (59 women, 28 men, 11 non-binary; $M_{\text{age}} = 21.05$; $SD_{\text{age}} = 3.15$). About half of participants were first-generation students (54%) and learned English as their first language (57%), and most were born in the United States (77%). A requirement for participating was that students had not previously participated in any of the studies described here or any related studies conducted by our lab.

Study 3 expanded on the dependent measures used in Study 2 to capture a wider range of evaluative responses beyond judgments of “OK” or “not OK” and evaluative ratings of good or bad. The dependent measures used in Study 3 came from the work outlined in Study S1, specifically the evaluative labels people gave for acts. In addition to the dependent measures for Study 3 already mentioned in the main text, we also asked participants to evaluate each scenario

on different dimensions: *trait-based*, *clear moral*, and *conflicted* (Table S4). These questions were not used in any analyses in the main text due to space constraints. The full list of prompts is available in Table S4.

Table S4

Additional Evaluative Prompts Used in Study 3, Presented Following Each Scenario

Prompt	Type	Response Format
What [protagonist] did was clever and resourceful	Trait-based	5-point Likert (Disagree, Somewhat disagree, Neutral, Somewhat agree, Agree)
What [protagonist] did was foolish and lazy	Trait-based	5-point Likert (Disagree, Somewhat disagree, Neutral, Somewhat agree, Agree)
What [protagonist] did was absolutely fine	Clear moral	5-point Likert (Disagree, Somewhat disagree, Neutral, Somewhat agree, Agree)
What [protagonist] did was absolutely wrong	Clear moral	5-point Likert (Disagree, Somewhat disagree, Neutral, Somewhat agree, Agree)
What [protagonist] did was technically wrong, but I can see where [he/she]’s coming from	Conflicted	5-point Likert (Disagree, Somewhat disagree, Neutral, Somewhat agree, Agree)
What [protagonist] did was technically okay, but it’s not the best thing to do	Conflicted	5-point Likert (Disagree, Somewhat disagree, Neutral, Somewhat agree, Agree)

Data Analysis

To assess whether the features we manipulated in the Study 3 scenarios would influence judgments and reasoning, we tested our key hypotheses using GLMMs to model participants’ evaluative responses. To test whether the manipulated features influenced participants’ judgments, models predicted evaluations (okay, rating) as a function of feature type (for, against) for each type of scenario. To see whether participants reasoned about the relevant features, models predicted whether each type of reason (e.g., learning) was selected as a function of scenario type (e.g., learning scenario). Like in Study 2, we used random intercepts for participants and tested hypotheses using likelihood ratio tests on model deviance (*D*).

Results

Effect of Manipulated Features on Evaluations

Okay Judgments. Overall, participants more often judged the act as okay when they saw the *for cheating* version (61%, 95% CI: [58%, 65%]) than when they saw the *against cheating* version (33%, 95% CI: [30%, 37%]), $D(1) = 151.26$, $p < .001$. Comparing *for*- and *against*-manipulations of the same reason revealed that *effort*, *fairness*, *honesty*, *learning*, and *rules* for-scenarios were all evaluated significantly more positively than their against-scenario counterparts, $Ds(1) > 11.46$, $ps < .001$ (Figure S3). Evaluations of *affect agent* for-scenarios did

not differ significantly from evaluations of their respective against-scenarios, $D(1) = 2.55$, $p = .11$.

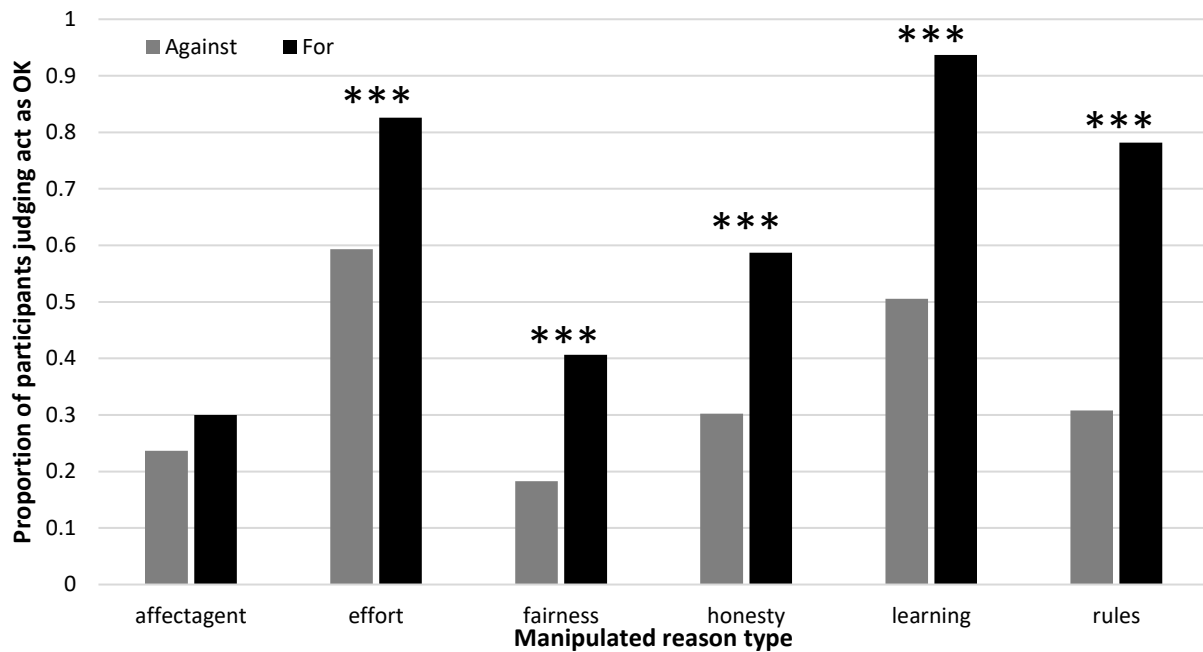


Figure S3. Judgments about whether the action was okay, grouped by the manipulated types of reasons *for* or *against* the act in the scenario. *** $p < .001$

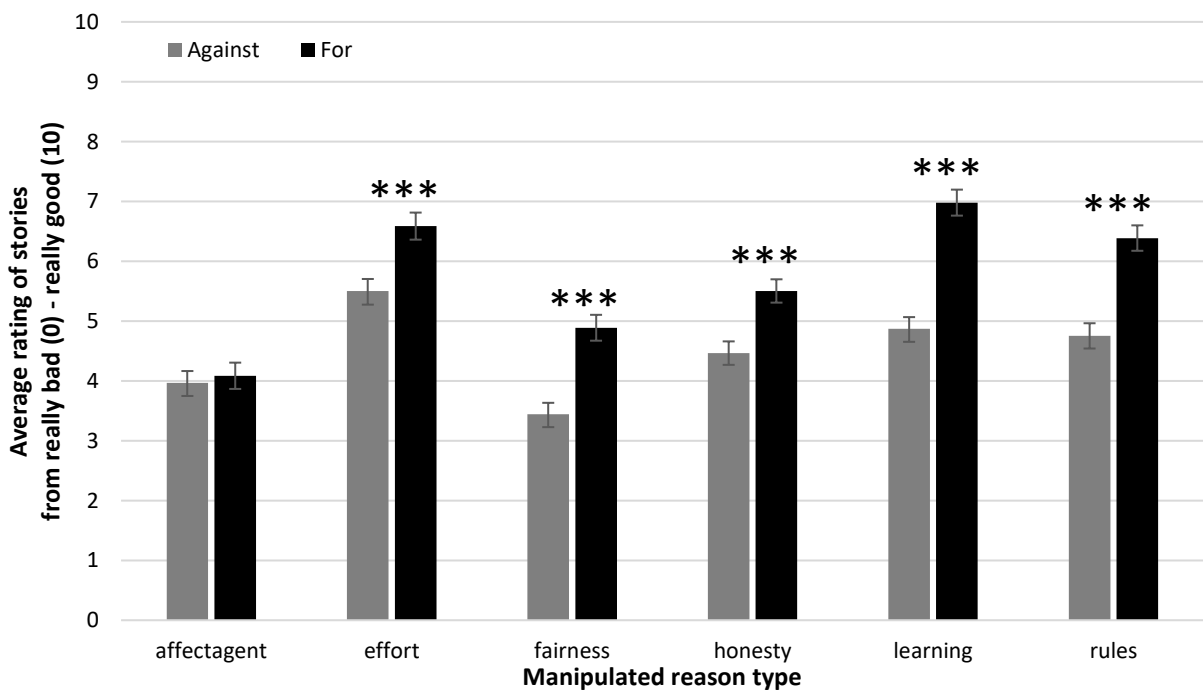


Figure S4. Evaluative ratings of scenarios, grouped by the manipulated types of reasons. Error bars represent standard errors of the mean. *** $p < .001$.

Evaluative Ratings. Similarly, the *for cheating* versions elicited more positive ratings ($M = 5.62$) than the *against cheating* versions ($M = 4.46$), $D(1) = 104.75$, $p < .001$. Similar to okay judgments, the same manipulated features (*effort, fairness, honesty, learning, and rules*) influenced ratings in the hypothesized direction, $Ds(1) > 13.89$, $ps < .001$ (Figure S4). Again, the *affect agent* manipulation did not yield significant differences in ratings, $D(1) = 0.05$, $p = .460$.

References

- Hox, J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22, 276-282.
- UCSC Division of Undergraduate Education (2019). *Code of student conduct*. Santa Cruz, CA. <https://deanofstudents.ucsc.edu/student-conduct/student-handbook/>
- Waltzer, T., & Dahl, A. (2021). Students' perceptions and evaluations of plagiarism: Effects of text and context. *Journal of Moral Education*, 50(4), 436-451. <http://doi.org/10.1080/03057240.2020.1787961>