

Supplementary Document to “Managerial Moral Self-Regulation and the Microfoundations of Corporate Social (Ir)Responsibility”

Table of Contents

Table of Contents.....	1
Appendix A: Supplementary Analyses for Study 1	2
Appendix B: Supplementary Analyses for Study 2A	18
Appendix C: Supplementary Analyses for Study 2B.....	26
Appendix D: Constructive Replication of Study 2B with Revised Decision Tasks	32
Appendix E: Supplementary Analyses for Study 2C.....	41
Appendix F: Supplementary Undergraduate Sample for Study 2C.....	46
Appendix G: Constructive Replication of Study 2C with Revised Moral Self-Image Manipulation and Revised Decision Tasks	51

Appendix A: Supplementary Analyses for Study 1

We conducted additional analyses to examine potential boundary conditions of the moral self-regulatory processes underlying our findings. We also performed several supplementary analyses to further assess the robustness of our findings and enrich our understanding of the proposed relationships.

Moderating Role of CEO Agentic Values

Upper echelons theory suggests that CEOs interpret strategic cues through the lens of their personal values and traits, which subsequently shape the firm's actions and outcomes (Hambrick and Mason 1984, Finkelstein et al. 2009). Building on this perspective, we examine whether CEO agentic values moderate the moral self-regulatory processes that link firms' social performance feedback to subsequent CSR and CSI behavior. Agentic values emphasize achievement, independence, control, and self-advancement (Abele et al. 2008, Abele and Wojciszke 2014).

We argue that highly agentic CEOs place less weight on stakeholder welfare and societal contributions when evaluating organizational outcomes, focusing instead on competence, control, and firm-centered objectives (DesJardine and Shi 2023, Zyung 2025). As a result, CEOs' agentic value is likely to influence their moral self-regulation. Following social underperformance, highly agentic CEOs may be less likely to experience moral threat and therefore engage in weaker moral cleansing responses. Conversely, following social overperformance, they may be more likely to view the firm's positive social standing as validating their leadership and competence, thereby exhibiting stronger moral licensing responses. Consequently, we expect CEO agentic values to weaken the effects of social underperformance on CSR increases and CSI reductions, while strengthening the effects of social overperformance on CSR reductions and CSI increases.

The results, reported in Models 1–4 of Table A1, generally support the moderating role of CEO agentic values. Specifically, in Model 1, the interaction between *social performance below aspiration* and *CEO agentic value* on CSR is negative and significant ($\beta = -0.103, p = 0.028$). In Model 2, however, the interaction term on CSI is not statistically significant ($\beta = -0.189, p = 0.157$). In Model 3, the interaction between *social performance above aspiration* and *CEO agentic value* on CSR is negative and significant

($\beta = -0.042, p = 0.014$). In Model 4, the interaction on CSI is positive and marginally significant ($\beta = 0.127, p = 0.086$). Taken together, these results suggest that CEO agentic values weaken moral cleansing (through the CSR path) but strengthen moral licensing (through both CSR and CSI channels). To facilitate interpretation, Figure A1 plots the marginal effects of social performance feedback on Δ CSR and Δ CSI across the observed range of CEO agentic values.

Moderating Role of Managerial Discretion

Upper echelons theory suggests that the influence of CEO psychological characteristics on firm outcomes depends on the degree of managerial discretion available to the CEO (Hambrick and Finkelstein 1987, Finkelstein et al. 2009). When discretion is high, CEOs can more readily translate their values, judgments, and psychological dispositions into organizational actions. When discretion is constrained by external stakeholders or governance structures, psychological tendencies are less likely to manifest in observable firm behavior (Crossland and Hambrick 2007, Wangrow et al. 2015). Accordingly, we suggest that the extent to which social performance feedback translates into CSR and CSI adjustments through moral self-regulation depends on CEOs' managerial discretion.

We measured constraints on managerial discretion using transient institutional ownership (Bushee 1998, 2001). Transient institutional investors typically have short investment horizons and place strong emphasis on short-term financial performance. As a result, they can constrain managerial latitude over discretionary investments and strategic resource allocation (Bushee 1998, 2001, Cella et al. 2013, Edmans et al. 2017). Under such conditions, CEOs may have less ability to translate moral cleansing or moral licensing motivations into CSR and CSI decisions. Consequently, we expect transient institutional ownership to weaken the effects of social underperformance on CSR increases and CSI reductions, and the effects of social overperformance on CSR reductions and CSI increases.

The results, reported in Models 5–8 of Table A1, support the moderating role of transient institutional ownership. Specifically, in Model 5, the interaction between *social performance below aspiration* and *transient institutional ownership* on CSR is negative and significant ($\beta = -0.728, p < 0.001$). In Model 6, however, the interaction term on CSI is not statistically significant ($\beta = 0.016, p =$

0.979). In Model 7, the interaction between *social performance above aspiration* and *transient institutional ownership* on CSR is positive and significant ($\beta = 0.198, p = 0.016$). In Model 8, the interaction on CSI is negative and significant ($\beta = -0.776, p = 0.021$). Taken together, these results suggest that the weaker managerial discretion is, as reflected in higher levels of transient institutional ownership, the weaker both moral cleansing (through the CSR path) and moral licensing (through both CSR and CSI channels). To facilitate interpretation, Figure A2 plots the marginal effects of social performance feedback on ΔCSR and ΔCSI across the observed range of transient institutional ownership.

Robustness Tests

First, we tested our hypotheses with alternative measures of social performance aspiration. In the main analysis, the social performance aspiration is a combined measure of a firm's historical aspirations and peers' (size-weighted) performance. To ensure the robustness of our findings, we tested our hypotheses (H1 and H2) with alternative measures of social performance aspiration. We reran all the analyses using historical aspiration and peers' performance separately instead of using the combined measure. The results, as reported in Tables A2 and A3, are consistent with our main findings, supporting our predictions. In addition, we re-estimated our main models using alternative peer performance-based aspiration measures constructed at the 3-digit and 4-digit SIC levels, in addition to our baseline 2-digit SIC specification. The results, as reported in Table A4, are generally consistent with our main findings.

Second, for CSR and CSI outcomes, in the main analyses, we gauged *change in CSR (CSI)* as the difference between the CSR (CSI) score at time $t + 1$ and the score at time t for a given firm. In additional analyses, we constructed industry-adjusted change measures defined as a firm's CSR (CSI) change minus the average change among other firms in the same two-digit SIC industry (excluding the focal firm). The results, as reported in Table A5, are consistent with our main findings.

Third, we tested the sensitivity of our findings to alternative fixed-effects configurations, including: (1) firm and industry-by-year fixed effects, (2) year and industry fixed effects, and (3) firm and year fixed effects. The results, as reported in Table A6, show that the signs and significance of our main effects are consistent across these specifications.

Fourth, we used RepRisk as an alternative dataset to measure CSI. The RepRisk Index (RRI) is a dynamic, machine-learning– and analyst-driven metric that captures a firm’s exposure to ESG-related reputational risks based on over 100,000 public sources (e.g., media, regulators, NGOs). Unlike Refinitiv ESG Scores, which cover a broad range of ESG dimensions, RRI focuses specifically on negative ESG events and risk exposure related to a firm’s operations, business relationships, and investments. We constructed an annual firm-level CSI measure by averaging the monthly RRI scores (0–100) within each fiscal year and then computed *Change in CSI* as the difference between the averages at time $t+1$ and t . Because RepRisk provides only CSI (no CSR), this analysis focuses on whether firms’ social performance relative to their CSI aspiration predicts subsequent CSI engagement (H1b and H2b). Using the same aspiration formulas applied to Refinitiv ESG, we generated a CSI aspiration level for each firm and decomposed feedback into two spline components: *performance above CSI aspiration*—the level of a firm’s CSI exceeds its aspiration (i.e., underperformance/high risk), and *performance below CSI aspiration*—the level of a firm’s CSI falls below its aspiration (i.e., overperformance/low risk).

Table A7 reports the supplementary results using RepRisk. Importantly, RRI reflects risk, not virtue: performing below CSI aspiration indicates *better-than-expected* social performance (low risk), and performing above aspiration indicates *worse-than-expected* performance (high risk). The results strongly support our hypotheses and replicate our main findings. We find (1) a positive and significant coefficient for RRI below aspiration ($\beta = 1.057, p < 0.001$), indicating that firms performing better/safer than peers subsequently experience higher risk exposure—consistent with H2b (moral licensing), and (2) a negative and significant coefficient for RRI above aspiration ($\beta = -0.511, p < 0.001$), indicating that firms facing higher-than-expected irresponsibility actively reduce future risk—supporting H1b (moral cleansing). Overall, the RepRisk analysis provides robust validation for our theorized mechanisms using an independent, event-based measure of CSI.

Lastly, we conducted an additional analysis to test the distinctiveness of CSR and CSI. We examined whether the effects of social performance feedback on CSR and CSI are equal in magnitude but opposite in direction. Specifically, we used a stacked regression model with clustered standard errors to

account for the interdependence of the dependent variables and performed Wald tests on the sum of the coefficients (testing $H_0: \beta_{CSR} + \beta_{CSI} = 0$). We found strong evidence of asymmetry. For *social performance below aspiration*, the sum of the coefficients is significantly positive ($\beta_{sum} = 0.13, p = 0.034$), indicating that the magnitude of the effect on CSR is significantly larger than the magnitude of the effect on CSI. Regarding *social performance above aspiration*, the sum of the coefficients is also significantly positive ($\beta_{sum} = 0.31, p < 0.001$), indicating that the magnitude of the effect on CSI is significantly larger than that on CSR. These results suggest that social performance feedback triggers asymmetrical reactions in firms' CSR and CSI strategies. This test provides additional evidence that CSR and CSI are not simply "two sides of the same coin." Combined with our earlier finding that CEO agentic values weaken moral cleansing through CSR but not through CSI, these results further support the view that CSR and CSI represent distinct pathways through which firms restore moral balance.

References not included in the main manuscript

- Abele AE, Rupperecht T, Wojciszke B (2008) The influence of success and failure experiences on agency. *European J. Soc. Psych.* 38: 436–448.
- Bushee BJ (1998) The influence of institutional investors on myopic R&D investment behavior. *Accounting Rev.* 73(3): 305–333.
- Cella C, Ellul A, Giannetti M (2013) Investors' horizons and the amplification of market shocks. *Rev. Financial Stud.* 26(7): 1607–1648.
- Edmans A, Fang VW, Lewellen KA (2017) Equity vesting and investment. *Rev. Financial Stud.* 30(7): 2229–2271.
- Finkelstein S, Hambrick DC, Cannella AA (2009). *Strategic leadership: Theory and research on executives, top management teams, and boards*. (Oxford University Press, New York).
- Hambrick DC, Finkelstein S (1987) Managerial Discretion: A Bridge between Polar Views of Organizational Outcomes, Staw B, Cummings LL, eds. *Research in Organizational Behavior*, (JAI Press, Greenwich, CT), 369–406.
- Zyung JD (2025) Expectations and exits: The effect of CEO agency and communion on boards' expectancy violations and dismissal of CEOs following breaches of stakeholder rights. *Organ. Sci.* 36(6): 2372–2398.

Table A1. Boundary Conditions of Moral Self-regulatory Processes in Study 1

Dependent Variable	(1) Change in CSR	(2) Change in CSI	(3) Change in CSR	(4) Change in CSI	(5) Change in CSR	(6) Change in CSI	(7) Change in CSR	(8) Change in CSI
Social performance below aspiration	0.082 (0.107)	-0.145 (0.363)			0.528*** (0.000)	-0.785*** (0.000)		
Social performance above aspiration			-0.061*** (0.004)	0.279*** (0.002)			-0.260*** (0.000)	1.091*** (0.000)
Social performance below aspiration X CEO agentive value	-0.103** (0.028)	-0.189 (0.157)						
Social performance above aspiration X CEO agentive value			-0.042** (0.014)	0.127* (0.086)				
Social performance below aspiration X Transient institutional ownership					-0.728*** (0.000)	0.016 (0.979)		
Social performance above aspiration X Transient institutional ownership							0.198** (0.016)	-0.776** (0.021)
CEO agentive value	-0.352 (0.218)	1.175 (0.203)	0.545 (0.121)	-1.315 (0.284)	-0.178 (0.672)	0.604 (0.707)	-0.192 (0.647)	0.619 (0.706)
CEO communal value	-0.760** (0.016)	0.139 (0.910)	-0.582* (0.057)	-0.199 (0.879)	-0.617 (0.209)	-1.267 (0.538)	-0.612 (0.207)	-1.230 (0.554)
Firm size	0.098 (0.206)	0.255 (0.188)	0.219*** (0.009)	-0.726*** (0.003)	0.989*** (0.003)	0.934 (0.372)	0.479 (0.127)	2.490** (0.025)
Firm age	-0.321** (0.016)	-0.908*** (0.005)	-0.287** (0.025)	-1.482*** (0.000)	2.168** (0.016)	-13.271*** (0.000)	0.936 (0.253)	-10.414*** (0.000)
Firm performance	1.378 (0.176)	-0.537 (0.836)	1.465 (0.137)	-2.238 (0.413)	3.018* (0.053)	-1.809 (0.750)	3.132** (0.044)	-4.961 (0.406)
Financial leverage	-0.012 (0.648)	0.013 (0.898)	-0.035 (0.196)	0.114 (0.290)	-0.007 (0.882)	0.120 (0.497)	0.018 (0.696)	0.050 (0.791)
Market-to-book ratio	-0.003 (0.836)	0.051 (0.348)	0.008 (0.602)	0.004 (0.951)	0.008 (0.753)	-0.047 (0.691)	-0.010 (0.717)	0.027 (0.830)
Independent director ratio	3.657*** (0.008)	-0.822 (0.809)	2.656** (0.042)	-2.684 (0.463)	-1.791 (0.510)	4.140 (0.637)	-3.197 (0.221)	5.991 (0.511)
Female director ratio	0.021 (0.981)	-8.233*** (0.003)	-0.314 (0.723)	-10.232*** (0.001)	0.061 (0.969)	-12.794** (0.017)	-0.157 (0.913)	-17.752*** (0.001)
Transient institutional ownership	2.331** (0.041)	1.699 (0.647)	2.125* (0.052)	2.363 (0.533)	5.465*** (0.007)	-6.538 (0.354)	1.141 (0.546)	0.041 (0.995)
Dedicate institutional ownership	-0.965 (0.542)	5.324 (0.319)	-1.071 (0.508)	5.916 (0.273)	-0.694 (0.831)	0.586 (0.945)	-0.101 (0.975)	-3.003 (0.730)
Female CEO	0.108 (0.771)	0.183 (0.864)	0.187 (0.608)	-0.005 (0.996)	-0.665 (0.265)	0.696 (0.778)	-0.291 (0.622)	-0.190 (0.943)
CEO tenure	-0.404*** (0.000)	-0.013 (0.969)	-0.444*** (0.000)	0.094 (0.792)	-0.641*** (0.000)	0.155 (0.773)	-0.708*** (0.000)	0.154 (0.779)
CEO duality	-0.038 (0.873)	0.521 (0.414)	-0.005 (0.982)	0.157 (0.821)	-0.754** (0.047)	2.231* (0.079)	-0.717* (0.062)	1.724 (0.170)
Year FE # Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-0.228 (0.872)	3.963 (0.332)	0.242 (0.857)	14.123*** (0.001)	-11.891*** (0.008)	38.359*** (0.002)	2.265 (0.582)	2.246 (0.863)
Observations	7,412	7,412	7,412	7,412	7,412	7,412	7,412	7,412
R-squared	0.078	0.060	0.086	0.085	0.054	-0.059	0.058	0.008
F	12.038	4.163	20.453	9.388	21.961	5.219	22.216	16.099

Note. *p*-value in parentheses; standard errors are clustered by firms; * *p* < .1; ** *p* < .05; *** *p* < .01.

Table A2. Robustness Check for Study 1: Alternative Measure of Social Performance Aspiration—Historical Aspiration

	(1) H1a (+)	(2) H1b (-)	(3) H2a (-)	(4) H2b (+)
Dependent Variable	Change in CSR	Change in CSI	Change in CSR	Change in CSI
Social performance below aspiration	0.113*** (0.002)	-1.138*** (0.000)		
Social performance above aspiration			-0.063*** (0.000)	0.248*** (0.000)
CEO agentic value	-0.063 (0.805)	0.719 (0.390)	-0.036 (0.891)	0.555 (0.527)
CEO communal value	-0.595** (0.045)	0.089 (0.943)	-0.597** (0.048)	-0.086 (0.946)
Firm size	-0.163** (0.020)	0.865*** (0.000)	0.102 (0.209)	-0.285 (0.185)
Firm age	-0.540*** (0.000)	-0.448 (0.171)	-0.352*** (0.007)	-1.240*** (0.000)
Firm performance	0.795 (0.402)	-0.405 (0.877)	1.209 (0.229)	-1.353 (0.613)
Financial leverage	-0.010 (0.705)	0.038 (0.709)	-0.035 (0.198)	0.113 (0.281)
Market-to-book ratio	-0.005 (0.737)	0.048 (0.372)	0.008 (0.557)	0.001 (0.993)
Independent director ratio	1.127 (0.374)	2.223 (0.504)	1.844 (0.164)	0.529 (0.880)
Female director ratio	-1.479* (0.078)	-5.733** (0.034)	-1.030 (0.243)	-7.350*** (0.010)
Transient institutional ownership	2.297** (0.034)	1.423 (0.700)	2.404** (0.030)	1.249 (0.741)
Dedicate institutional ownership	-0.845 (0.603)	4.392 (0.421)	-1.282 (0.430)	6.653 (0.218)
Female CEO	0.203 (0.570)	-0.387 (0.716)	0.064 (0.861)	0.459 (0.671)
CEO tenure	-0.425*** (0.000)	-0.040 (0.905)	-0.455*** (0.000)	0.127 (0.717)
CEO duality	-0.129 (0.564)	0.507 (0.455)	-0.112 (0.627)	0.575 (0.401)
Year FE # Industry FE	YES	YES	YES	YES
Constant	5.556*** (0.000)	-5.581 (0.151)	2.921** (0.025)	4.149 (0.280)
Observations	7,412	7,412	7,412	7,412
R-squared	0.057	0.071	0.072	0.072
F	7.398	4.538	14.493	9.127

Note. *p*-value in parentheses; standard errors are clustered by firms; * $p < .1$; ** $p < .05$; *** $p < .01$.

Table A3. Robustness Check for Study 1: Alternative Measure of Social Performance Aspiration—Peers' Size-weighted Performance

	(1) H1a (+)	(2) H1b (–)	(3) H2a (–)	(4) H2b (+)
Dependent Variable	Change in CSR	Change in CSI	Change in CSR	Change in CSI
Social performance below aspiration	0.135*** (0.000)	-0.164*** (0.000)		
Social performance above aspiration			-0.119*** (0.000)	0.423*** (0.000)
CEO agentic value	0.058 (0.827)	0.490 (0.563)	0.115 (0.655)	0.023 (0.979)
CEO communal value	-0.684** (0.029)	-0.070 (0.955)	-0.450 (0.143)	-0.625 (0.637)
Firm size	0.139* (0.067)	0.352* (0.065)	0.194** (0.015)	-0.515** (0.030)
Firm age	-0.262** (0.048)	-0.854*** (0.007)	-0.340*** (0.008)	-1.211*** (0.001)
Firm performance	1.418 (0.163)	-0.160 (0.951)	1.125 (0.239)	-0.833 (0.752)
Financial leverage	-0.009 (0.723)	0.006 (0.949)	-0.026 (0.334)	0.072 (0.497)
Market-to-book ratio	-0.003 (0.806)	0.053 (0.327)	0.001 (0.927)	0.031 (0.598)
Independent director ratio	3.813*** (0.005)	0.617 (0.853)	3.038** (0.018)	-3.368 (0.344)
Female director ratio	0.406 (0.648)	-7.791*** (0.005)	0.209 (0.810)	-11.573*** (0.000)
Transient institutional ownership	2.202* (0.054)	1.895 (0.608)	1.943* (0.076)	2.942 (0.437)
Dedicate institutional ownership	-1.056 (0.516)	5.433 (0.318)	-0.933 (0.574)	5.306 (0.336)
Female CEO	0.071 (0.846)	0.200 (0.852)	0.247 (0.488)	-0.228 (0.835)
CEO tenure	-0.441*** (0.000)	0.049 (0.885)	-0.437*** (0.000)	0.053 (0.880)
CEO duality	-0.039 (0.866)	0.589 (0.351)	0.058 (0.803)	-0.019 (0.977)
Year FE # Industry FE	YES	YES	YES	YES
Constant	-1.588 (0.268)	2.104 (0.596)	0.492 (0.703)	11.743*** (0.004)
Observations	7,408	7,408	7,408	7,408
R-squared	0.079	0.057	0.088	0.081
F	12.869	3.513	21.371	9.748

Note. *p*-value in parentheses; standard errors are clustered by firms; * $p < .1$; ** $p < .05$; *** $p < .01$.

Table A4. Robustness Check for Study 1: Alternative Measure of Social Performance Aspiration—using 3- and 4-digit SIC levels for peers' performance

	(1) H1a (+)	(2) H1b (-)	(3) H2a (-)	(4) H2b (+)	(5) H1a (+)	(6) H1b (-)	(7) H2a (-)	(8) H2b (+)
	3-digit SIC				4-digit SIC			
Dependent Variable	Change in CSR	Change in CSI	Change in CSR	Change in CSI	Change in CSR	Change in CSI	Change in CSR	Change in CSI
Social performance below aspiration	0.181*** (0.000)	-0.359*** (0.000)			0.090** (0.036)	-0.201 (0.178)		
Social performance above aspiration			-0.105*** (0.000)	0.381*** (0.000)			-0.044** (0.026)	0.281*** (0.001)
CEO agentic value	-0.010 (0.968)	0.539 (0.526)	0.086 (0.743)	0.116 (0.897)	-0.320 (0.260)	1.076 (0.247)	0.677** (0.048)	-0.860 (0.488)
CEO communal value	-0.661** (0.037)	-0.023 (0.985)	-0.542* (0.076)	-0.298 (0.821)	-0.708** (0.026)	0.058 (0.963)	-0.623** (0.041)	-0.163 (0.900)
Firm size	0.116 (0.136)	0.178 (0.361)	0.228*** (0.006)	-0.663*** (0.006)	0.124 (0.113)	0.164 (0.399)	0.237*** (0.005)	-0.678*** (0.005)
Firm age	-0.315** (0.017)	-0.957*** (0.003)	-0.255** (0.049)	-1.535*** (0.000)	-0.313** (0.018)	-0.961*** (0.003)	-0.252* (0.051)	-1.540*** (0.000)
Firm performance	0.787 (0.428)	0.518 (0.841)	1.491 (0.136)	-2.212 (0.413)	0.872 (0.381)	0.371 (0.886)	1.629 (0.102)	-2.440 (0.369)
Financial leverage	-0.014 (0.611)	0.017 (0.869)	-0.033 (0.213)	0.100 (0.343)	-0.013 (0.613)	0.017 (0.870)	-0.033 (0.215)	0.100 (0.345)
Market-to-book ratio	-0.005 (0.753)	0.054 (0.327)	0.009 (0.520)	0.002 (0.977)	-0.005 (0.742)	0.054 (0.326)	0.009 (0.529)	0.002 (0.975)
Independent director ratio	3.662*** (0.006)	-1.297 (0.705)	3.150** (0.016)	-3.935 (0.276)	3.800*** (0.005)	-1.536 (0.656)	3.037** (0.020)	-3.748 (0.304)
Female director ratio	0.323 (0.720)	-9.099*** (0.001)	-0.062 (0.944)	-10.721*** (0.000)	0.253 (0.780)	-8.978*** (0.001)	-0.171 (0.847)	-10.540*** (0.001)
Transient institutional ownership	2.440** (0.033)	1.468 (0.696)	2.018* (0.068)	2.699 (0.479)	2.398** (0.035)	1.542 (0.681)	2.039* (0.064)	2.665 (0.484)
Dedicate institutional ownership	-1.113 (0.490)	5.628 (0.309)	-1.167 (0.479)	6.123 (0.252)	-1.024 (0.516)	5.474 (0.312)	-1.008 (0.534)	5.862 (0.267)
Female CEO	0.121 (0.739)	0.167 (0.876)	0.136 (0.708)	0.167 (0.877)	0.146 (0.690)	0.124 (0.908)	0.160 (0.664)	0.128 (0.906)
CEO tenure	-0.436*** (0.000)	0.045 (0.893)	-0.462*** (0.000)	0.143 (0.686)	-0.419*** (0.000)	0.015 (0.965)	-0.445*** (0.000)	0.115 (0.744)
CEO duality	0.001 (0.995)	0.424 (0.505)	0.016 (0.945)	0.120 (0.862)	0.011 (0.963)	0.408 (0.522)	0.027 (0.907)	0.102 (0.882)
Year FE # Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-0.843 (0.548)	6.208 (0.124)	0.049 (0.971)	13.727*** (0.001)	-0.671 (0.632)	5.910 (0.147)	-0.571 (0.673)	14.750*** (0.001)
Observations	7,412	7,412	7,412	7,412	7,412	7,412	7,412	7,412
R-squared	0.081	0.061	0.088	0.082	0.082	0.061	0.089	0.082
F	14.393	5.042	22.970	10.154	13.738	4.837	22.598	9.549

Note. *p*-value in parentheses; standard errors are clustered by firms; * *p* < .1; ** *p* < .05; *** *p* < .01.

Table A5. Robustness Check for Study 1: Alternative Measure of Change in CSR/CSI—Industry-Adjusted Measurement

	(1) H1a (+)	(2) H1b (-)	(3) H2a (-)	(4) H2b (+)
Dependent Variable	Change in CSR	Change in CSI	Change in CSR	Change in CSI
Social performance below aspiration	0.209*** (0.000)	-0.398*** (0.000)		
Social performance above aspiration			-0.117*** (0.000)	0.461*** (0.000)
CEO agentic value	0.001 (0.998)	0.505 (0.591)	0.073 (0.790)	0.093 (0.926)
CEO communal value	-0.757** (0.024)	-0.238 (0.882)	-0.553* (0.088)	-0.722 (0.666)
Firm size	0.102 (0.209)	0.228 (0.269)	0.232*** (0.008)	-0.809*** (0.002)
Firm age	-0.331** (0.017)	-1.004*** (0.004)	-0.301** (0.027)	-1.591*** (0.000)
Firm performance	1.387 (0.194)	-0.217 (0.937)	1.469 (0.158)	-1.831 (0.522)
Financial leverage	-0.011 (0.698)	0.024 (0.822)	-0.036 (0.208)	0.135 (0.247)
Market-to-book ratio	-0.005 (0.744)	0.050 (0.368)	0.007 (0.627)	-0.002 (0.972)
Independent director ratio	3.792*** (0.009)	-1.073 (0.769)	2.914** (0.036)	-3.334 (0.392)
Female director ratio	0.106 (0.911)	-9.209*** (0.002)	-0.238 (0.800)	-11.412*** (0.000)
Transient institutional ownership	2.367* (0.055)	1.614 (0.691)	2.094* (0.076)	2.460 (0.552)
Dedicate institutional ownership	-1.337 (0.444)	6.108 (0.314)	-1.414 (0.423)	6.826 (0.264)
Female CEO	0.102 (0.794)	0.171 (0.883)	0.183 (0.632)	0.004 (0.998)
CEO tenure	-0.455*** (0.000)	-0.025 (0.945)	-0.491*** (0.000)	0.096 (0.801)
CEO duality	-0.051 (0.838)	0.411 (0.546)	-0.010 (0.967)	0.015 (0.984)
Year FE # Industry FE	YES	YES	YES	YES
Constant	-2.815* (0.060)	4.132 (0.346)	-1.491 (0.283)	12.437*** (0.005)
Observations	7,412	7,412	7,412	7,412
R-squared	0.031	0.058	0.023	0.031
F	12.963	4.379	20.559	9.800

Note. *p*-value in parentheses; standard errors are clustered by firms; * $p < .1$; ** $p < .05$; *** $p < .01$.

Table A6. Robustness Check for Study 1: Different Fixed Effects Configuration

Panel A: Firm and industry-by-year fixed effects				
	(1)	(2)	(3)	(4)
	H1a (+)	H1b (-)	H2a (-)	H2b (+)
Dependent Variable	Change in CSR	Change in CSI	Change in CSR	Change in CSI
Social performance below aspiration	0.430*** (0.000)	-0.786*** (0.000)		
Social performance above aspiration			-0.240*** (0.000)	1.013*** (0.000)
CEO agentic value	-0.191 (0.651)	0.581 (0.718)	-0.201 (0.631)	0.680 (0.678)
CEO communal value	-0.523 (0.287)	-1.269 (0.537)	-0.569 (0.241)	-1.392 (0.502)
Firm size	1.006*** (0.003)	0.951 (0.366)	0.496 (0.114)	2.384** (0.032)
Firm age	2.480** (0.005)	-13.306*** (0.000)	1.106 (0.174)	-11.044*** (0.000)
Firm performance	3.087** (0.048)	-1.855 (0.745)	3.131** (0.045)	-4.910 (0.411)
Financial leverage	-0.006 (0.895)	0.120 (0.496)	0.018 (0.705)	0.052 (0.781)
Market-to-book ratio	0.006 (0.836)	-0.047 (0.688)	-0.011 (0.686)	0.031 (0.799)
Independent director ratio	-1.444 (0.597)	4.204 (0.632)	-3.110 (0.235)	5.595 (0.541)
Female director ratio	0.159 (0.918)	-12.747** (0.018)	-0.024 (0.987)	-18.279*** (0.001)
Transient institutional ownership	0.874 (0.590)	-3.995 (0.507)	1.384 (0.372)	-3.733 (0.536)
Dedicate institutional ownership	-0.694 (0.834)	0.384 (0.964)	-0.235 (0.942)	-2.427 (0.781)
Female CEO	-0.612 (0.299)	0.641 (0.796)	-0.260 (0.658)	-0.287 (0.916)
CEO tenure	-0.648*** (0.000)	0.164 (0.759)	-0.713*** (0.000)	0.168 (0.761)
CEO duality	-0.754** (0.048)	2.243* (0.078)	-0.716* (0.062)	1.709 (0.176)
Firm FE	YES	YES	YES	YES
Year FE # Industry FE	YES	YES	YES	YES
Constant	-12.951*** (0.004)	38.261*** (0.002)	1.334 (0.743)	6.378 (0.624)
Observations	7,389	7,389	7389	7,389
R-squared	0.075	0.059	0.058	0.008
F	22.444	5.489	23.536	16.694

Note. *p*-value in parentheses; standard errors are clustered by firms; * *p* < .1; ** *p* < .05; *** *p* < .01.

“Year FE # Industry FE” represents industry-by-year fixed effects.

Panel B: Year and industry fixed effects				
	(1) H1a (+)	(2) H1b (-)	(3) H2a (-)	(4) H2b (+)
Dependent Variable	Change in CSR	Change in CSI	Change in CSR	Change in CSI
Social performance below aspiration	0.221*** (0.000)	-0.433*** (0.000)		
Social performance above aspiration			-0.128*** (0.000)	0.513*** (0.000)
CEO agentic value	0.053 (0.804)	0.363 (0.621)	0.123 (0.566)	-0.038 (0.962)
CEO communal value	-0.671*** (0.009)	-0.297 (0.798)	-0.458* (0.068)	-0.830 (0.519)
Firm size	0.118* (0.074)	0.174 (0.279)	0.249*** (0.000)	-0.932*** (0.000)
Firm age	-0.298*** (0.008)	-1.084*** (0.000)	-0.265** (0.016)	-1.758*** (0.000)
Firm performance	1.541* (0.050)	-2.840 (0.223)	1.623** (0.035)	-4.066* (0.098)
Financial leverage	0.004 (0.866)	-0.069 (0.448)	-0.015 (0.524)	0.023 (0.816)
Market-to-book ratio	-0.010 (0.462)	0.056 (0.236)	0.001 (0.917)	0.007 (0.891)
Independent director ratio	3.570*** (0.003)	-2.798 (0.393)	2.841** (0.014)	-5.343 (0.118)
Female director ratio	-0.168 (0.823)	-8.402*** (0.001)	-0.346 (0.643)	-11.315*** (0.000)
Transient institutional ownership	1.714* (0.065)	0.168 (0.957)	1.183 (0.186)	2.009 (0.529)
Dedicate institutional ownership	-0.636 (0.641)	10.087* (0.056)	-0.549 (0.698)	10.742** (0.042)
Female CEO	0.101 (0.733)	-1.233 (0.163)	0.139 (0.638)	-1.236 (0.190)
CEO tenure	-0.485*** (0.000)	-0.115 (0.679)	-0.499*** (0.000)	-0.087 (0.765)
CEO duality	0.182 (0.359)	0.663 (0.261)	0.187 (0.354)	0.370 (0.556)
Year FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Constant	-1.061 (0.374)	9.455*** (0.009)	0.251 (0.824)	18.339*** (0.000)
Observations	8,794	8,794	8,794	8,794
R-squared	0.065	0.016	0.074	0.046
F	20.478	6.830	32.134	14.127

Note. *p*-value in parentheses; standard errors are clustered by firms; * *p* < .1; ** *p* < .05; *** *p* < .01.

Panel C: Firm and year fixed effects				
	(1) H1a (+)	(2) H1b (-)	(3) H2a (-)	(4) H2b (+)
Dependent Variable	Change in CSR	Change in CSI	Change in CSR	Change in CSI
Social performance below aspiration	0.390*** (0.000)	-0.765*** (0.000)		
Social performance above aspiration			-0.233*** (0.000)	0.955*** (0.000)
CEO agentic value	-0.077 (0.798)	0.079 (0.946)	-0.053 (0.856)	0.040 (0.973)
CEO communal value	-0.292 (0.397)	-1.244 (0.450)	-0.299 (0.392)	-1.157 (0.547)
Firm size	0.633** (0.016)	1.067 (0.147)	0.301 (0.198)	1.849** (0.023)
Firm age	1.190* (0.074)	-11.490*** (0.000)	0.179 (0.771)	-9.964*** (0.000)
Firm performance	2.517** (0.021)	-5.505 (0.174)	2.369** (0.021)	-6.216 (0.144)
Financial leverage	0.016 (0.651)	-0.084 (0.513)	0.012 (0.726)	-0.054 (0.682)
Market-to-book ratio	-0.012 (0.585)	0.032 (0.693)	-0.014 (0.532)	0.055 (0.524)
Independent director ratio	0.402 (0.847)	-1.355 (0.845)	-0.542 (0.784)	-1.278 (0.860)
Female director ratio	-1.077 (0.347)	-13.323*** (0.001)	-0.879 (0.429)	-18.740*** (0.000)
Transient institutional ownership	0.952 (0.444)	-4.530 (0.317)	0.237 (0.846)	-1.064 (0.819)
Dedicate institutional ownership	-0.486 (0.848)	12.210* (0.093)	0.845 (0.741)	9.054 (0.226)
Female CEO	-0.331 (0.506)	-1.627 (0.323)	-0.074 (0.887)	-2.196 (0.235)
CEO tenure	-0.648*** (0.000)	0.022 (0.955)	-0.652*** (0.000)	-0.152 (0.709)
CEO duality	-0.128 (0.656)	1.410 (0.150)	-0.174 (0.546)	1.311 (0.182)
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Constant	-7.674** (0.024)	38.663*** (0.000)	3.130 (0.312)	15.563 (0.120)
Observations	8,775	8,775	8,775	8,775
R-squared	0.067	0.036	0.081	0.023
F	25.832	8.089	30.878	20.899

Note. *p*-value in parentheses; standard errors are clustered by firms; * *p* < .1; ** *p* < .05; *** *p* < .01.

Table A7. Robustness Check for Study 1: Alternative Measure of CSI—RepRisk RRI Measure

	(1)	(2)
	H2b	H1b
Dependent Variable	Change in RRI	
RRI below aspiration	1.057*** (0.000)	
RRI above aspiration		-0.511*** (0.000)
CEO agentic value		
CEO communal value		
Firm size	0.923*** (0.000)	1.740*** (0.000)
Firm age	-0.040 (0.764)	0.096 (0.524)
Firm performance	1.168 (0.247)	1.801 (0.110)
Financial leverage	-0.016 (0.508)	-0.049** (0.046)
Market-to-book ratio	0.007 (0.744)	0.024 (0.239)
Independent director ratio	0.140 (0.872)	-0.371 (0.699)
Female director ratio	1.632* (0.088)	2.589** (0.017)
Transient institutional ownership	1.213 (0.222)	-0.340 (0.761)
Dedicate institutional ownership	0.819 (0.499)	0.384 (0.797)
Female CEO	-0.547 (0.288)	-0.619 (0.215)
CEO tenure	-0.008 (0.922)	0.003 (0.974)
CEO duality	0.130 (0.669)	0.227 (0.468)
Year FE	YES	YES
Industry FE	YES	YES
Constant	6276 0.191	6276 0.247
Observations	47.838	63.719
R-squared	6276	6276
F	0.191	0.247

Note. *p*-value in parentheses; standard errors are clustered by firms; * $p < .1$; ** $p < .05$; *** $p < .01$

Figure A1. Marginal Effect of Social Performance Feedback on Change in CSR and CSI across CEO Agentic Values

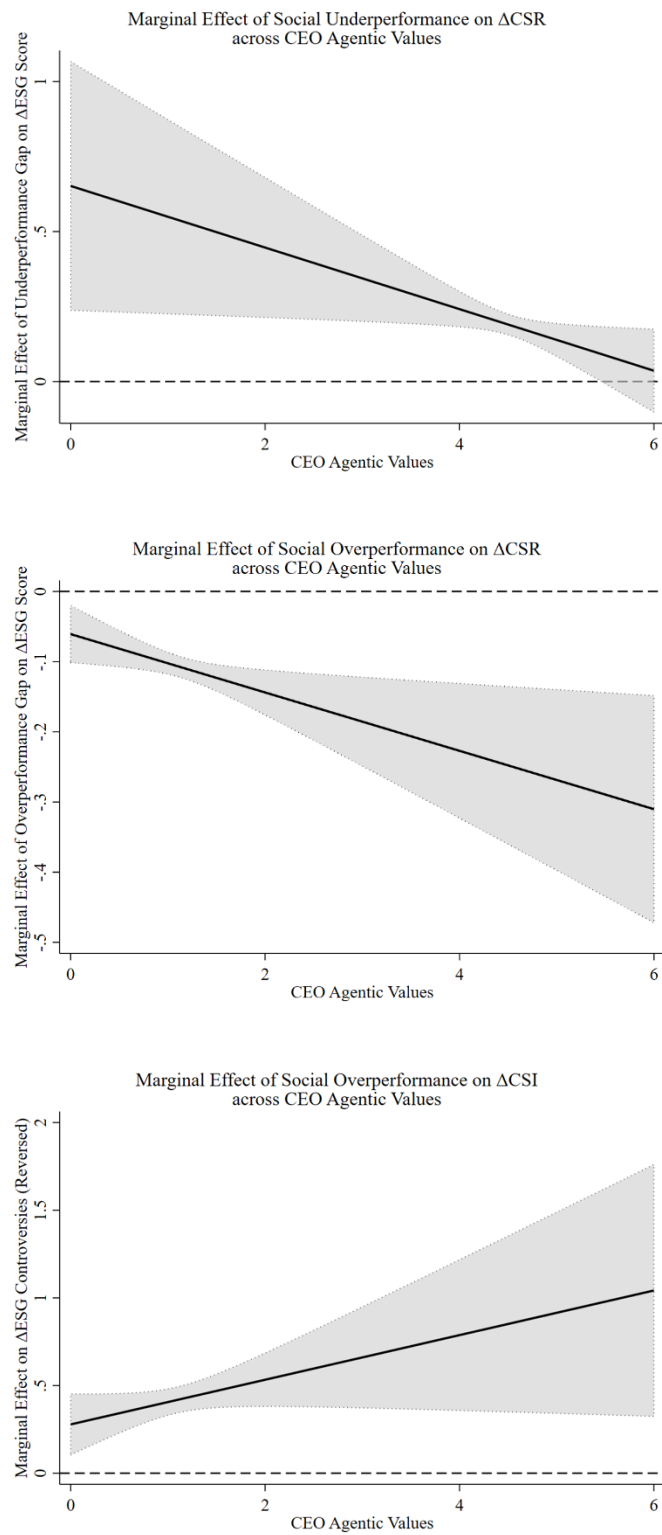
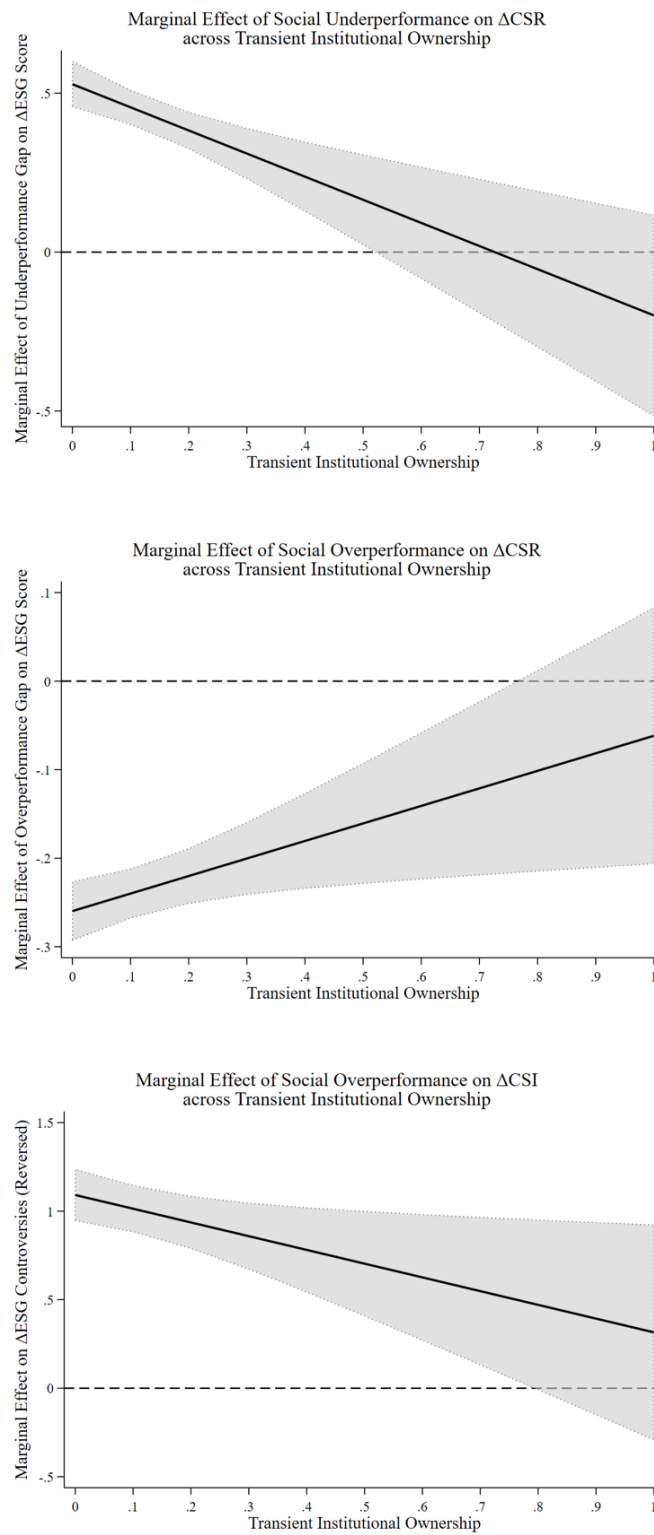


Figure A2. Marginal Effect of Social Performance Feedback on Change in CSR and CSI across Transient Institutional Ownership



Appendix B: Supplementary Analyses for Study 2A

Attention Check Analyses

We included two attention check questions in Study 2A. No participant in our sample failed both attention checks; however, 10 participants failed one of them. We first examined whether the participants who failed an attention check differed significantly on demographic or key study variables compared to those who passed all attention checks. Independent sample *t*-tests and chi-square tests revealed that those who failed one of the attention checks did not significantly differ from those who passed in age ($p = .555$), race ($p = .972$), sex ($p = .186$), education ($p = .834$), work experience ($p = .821$), managerial experience ($p = .885$), or company size ($p = .409$). They also did not statistically differ on their responses to moral identity ($p = .069$) or perceptions of sustainability image ($p = .508$). However, those who failed one attention check donated less ($M = 3.05$, $SD = 2.41$) than those who passed ($M = 5.21$, $SD = 2.43$, $t(413) = 2.77$, $p = .006$) and engaged in more CSI ($M = 57.90$, $SD = 26.33$) than those who passed ($M = 36.08$, $SD = 23.35$, $t(413) = -2.91$, $p = .004$).

To test the robustness of our findings, we conducted the same analyses on a trimmed sample that excluded these 10 participants (after exclusions, $N = 405$). Table B1 contains the means and standard deviations of key variables by condition for both the full and trimmed samples, and Table B2 provides the corresponding descriptive statistics and correlations of key variables. Because Levene's test again indicated unequal variances for CSR in this trimmed sample, $F(2,402) = 9.20$, $p < .001$, we conducted Welch's ANOVA. Results revealed a significant effect of social performance feedback on CSR, $F(2, 261.88) = 5.02$, $p = .007$, $\eta^2 = .03$. Planned contrasts demonstrated the same pattern as in our analysis of the full sample. Participants in the underperformance condition reported higher CSR ($M = 5.78$; $SD = 2.77$) than those in the control condition ($M = 5.00$, $SD = 2.39$; $p = .015$), consistent with moral cleansing (H1a), and CSR in the overperformance condition ($M = 4.86$, $SD = 2.01$) did not differ from that in the control condition ($p = .605$). In summary, underperformance increased CSR (supporting moral cleansing; H1a), whereas outperformance did not reduce CSR (providing no support for licensing; H2a).

Because the homogeneity of variance assumption was met for CSI, $F(2,402) = 2.28$, $p = .104$,

planned contrasts were conducted within the framework of a traditional ANOVA, $F(2, 402) = 12.90, p < .001, \eta^2 = .06$. Planned contrasts showed that participants in the underperformance condition reported slightly lower CSI ($M = 29.58, SD = 24.31$) than those in the control condition ($M = 34.96, SD = 22.13; p = .054$), consistent with moral cleansing (H1b). Consistent with the full sample results, participants in the overperformance condition reported higher CSI ($M = 43.45, SD = 21.59$) than those in the control condition ($p = .002$), again supporting moral licensing through increased CSI (H2b).

Due to unequal variances in perceived firm sustainability image, $F(2, 402) = 7.37, p < .001$, planned contrasts were conducted within the framework of Welch's ANOVA, $F(2, 258.09) = 499.71, p < .001, \eta^2 = .73$. Consistent with our hypotheses and the results from the full sample, participants in the underperformance condition perceived a lower organizational sustainability image ($M = 2.32; SD = 1.22$) than those in the control ($M = 4.78; SD = 0.92, p < .001$), and participants in the overperformance condition ($M = 6.24; SD = 0.77$) perceived a higher organizational sustainability image than those in the control condition, $p < .001$. Regression analysis also supported the negative effect of underperformance ($b = -2.47; p < .001$) and positive effect of overperformance ($b = 1.46; p < .001$) on perceived firm sustainability image.

Mediation analysis results showed that perceived sustainability image was negatively related to donation ($b = -0.30; p = .012$) and positively related to pollution ($b = 1.99; p = 0.08$). Consistent with the full sample results, the indirect effect of *underperformance* via perceived sustainability image on CSR was significant ($Estimate = 0.75; SE = .33; 95\% CI [0.08, 1.38]$), and the indirect effect on CSI was not significant ($Estimate = -4.91; SE = 3.22; 95\% CI [-11.12, 1.50]$). Likewise, the effect of *overperformance* via perceived sustainability image on CSR was significant ($Estimate = -.44; SE = .20; 95\% CI [-.84, -.04]$), and the indirect effect on CSI was not significant ($Estimate = 2.90; SE = 1.94; 95\% CI [-.69, 6.79]$). Thus, excluding participants who failed the attention checks did not change the results.

CSI Outlier Analyses

Given the substantial variance in CSI, which may have attenuated tests of mediation, we explored the potential influence of outliers in CSI. Following recommended practices for identifying and handling

outliers (Aguinis et al. 2013), we identified 13 outlier cases (all above +2.24 SDs) in our CSI measure. Independent sample *t*-tests and chi-square tests revealed that these outliers did not differ from non-outliers in age ($p = .331$), race ($p = .999$), sex ($p = .157$), education ($p = .957$), managerial experience ($p = .732$), or company size ($p = .613$). They also did not differ in their responses to perceptions of sustainability image ($p = .696$). However, outliers showed greater work experience ($M = 26.23$, $SD = 12.24$) than non-outliers ($M = 20.23$, $SD = 11.07$, $t(413) = -1.92$, $p = .028$). Extreme cases tended to show lower donations ($M = 3.13$, $SD = 3.21$) than non-outlier cases ($M = 5.22$, $SD = 2.40$, $t(413) = 3.06$, $p = .038$) and higher CSI ($M = 94.92$, $SD = 4.52$) than non-extreme cases ($M = 34.72$, $SD = 21.49$, $t(413) = -10.08$, $p < .001$). A chi-square analysis further confirmed that the outliers were evenly distributed across experimental conditions (2.1% of control, 3.6% of positive feedback, and 3.6% of negative feedback, $\chi^2 = .68$, $p = .711$), suggesting that they were not a result of an ineffective manipulation.

We next conducted all main analyses excluding these 13 extreme cases. Table B3 contains the means and standard deviations of key variables by condition for both the full and trimmed samples, and Table B4 provides the corresponding descriptive statistics and correlations of key variables. Similarly, we first found that the homogeneity of variance assumption for ANOVAs was not met for both CSR, $F(2,399) = 10.86$, $p < .001$, and CSI, $F(2,399) = 3.41$, $p = .034$. Results from Welch's ANOVA showed a main effect of social performance feedback on CSR, $F(2, 260.31) = 5.024$, $p = .007$, $\eta^2 = .03$, and CSI, $F(2, 265.06) = 13.93$, $p < .001$, $\eta^2 = .06$. Planned contrasts revealed that participants in the underperformance condition reported higher CSR ($M = 5.80$; $SD = 2.75$) than those in the control condition ($M = 4.92$; $SD = 2.36$; $p = .005$), and lower CSI ($M = 28.07$; $SD = 22.01$) than those in the control condition ($M = 34.62$; $SD = 21.14$; $p = .014$), supporting the moral cleansing hypotheses (H1a and H1b). On the other hand, CSR for those in the overperformance condition ($M = 4.96$; $SD = 1.93$) was not significantly different from those in the control condition ($p = 0.892$; not supporting H2a), while CSI in the overperformance condition ($M = 41.45$; $SD = 19.29$) was significantly higher than the control condition ($p = .006$; supporting H2b). Thus, the direct effects of social performance feedback on CSR and

CSI mirror those reported in the main manuscript, except for the effect of underperformance on CSI, which became statistically significant after removing outliers.

In examining perceived firm sustainability image, we again found unequal variances, $F(2,399) = 7.49, p < .001$, requiring planned contrasts to be conducted within the framework of Welch's ANOVA, $F(2, 257.54) = 484.19, p < .001, \eta^2 = .73$. In support of the predicted effects on perceived firm sustainability image, participants in the underperformance condition perceived a lower organizational sustainability image ($M = 2.34; SD = 1.23$) than those in the control ($M = 4.78; SD = 0.91, p < .001$), and participants in the overperformance condition ($M = 6.25; SD = 0.77$) perceived a higher organizational sustainability image than those in the control condition, $p < .001$. Regression analysis also supported the negative effect of social underperformance ($b = -2.43; p < .001$) and positive effect of social overperformance ($b = 1.47; p < .001$) on perceived firm sustainability image.

Perceived sustainability image was negatively related to CSR ($b = -.37; p = .002$) and positively related to CSI ($b = 2.72; p = .010$). Accordingly, the indirect effect of underperformance through perceived sustainability image was both significant for CSR (*Estimate* = .90; *SE* = .32; 95% CI [0.26, 1.53]) and CSI (*Estimate* = -6.63; *SE* = 2.97; 95% CI [-12.31, -.68]). The overperformance indirect effects were also significant for both CSR (*Estimate* = -.55; *SE* = .21; 95% CI [-.96, -.16]) and CSI (*Estimate* = 4.01; *SE* = 1.86; 95% CI [.42, 7.69]). These results mirror the full-sample findings, except that, after removing outliers, the CSI indirect effects provide evidence of both moral cleansing and moral licensing, via perceived firm sustainability image.

Table B1. Means and Standard Deviations for Key Variables by Condition in Study 2A: Before and After Excluding Attention Check Failures

	Full Sample ($N = 415$)				Trimmed Sample ($N = 405$)			
	N	Firm sustainability image	CSR	CSI	N	Firm sustainability image	CSR	CSI
Social underperformance	137	2.33 (1.22)	5.68 (2.80)	30.50 (25.00)	133	2.32 (1.22)	5.78 (2.77)	29.58 (24.31)
No social performance feedback	140	4.79 (0.91)	4.93 (2.41)	35.84 (22.54)	134	4.78 (0.92)	5.00 (2.39)	34.96 (22.13)
Social overperformance	138	6.24 (0.77)	4.86 (2.01)	43.45 (21.59)	138	6.24 (0.77)	4.86 (2.01)	43.45 (21.59)

Note. Means are reported in each cell, with standard deviations reported in parentheses.

Table B2. Descriptive Statistics and Correlations of Key Variables in Study 2A: Before and After Excluding Attention Check Failures

	Full Sample (<i>N</i> = 415)				Trimmed Sample (<i>N</i> = 405)			
	<i>M</i>	<i>SD</i>	Firm sustainability image	CSR	<i>M</i>		Firm sustainability image	CSR
Firm sustainability image	4.46	1.89	--	--	4.47	1.90	--	--
CSR	5.16	2.45	-.193***	--	5.21	2.43	-.201***	--
CSI	36.61	23.63	.237***	-.286***	36.08	23.35	.246***	-.265***

Note. *M* = Mean. *SD* = Standard deviation. * $p < .1$; ** $p < .05$; *** $p < .01$

Table B3. Means and Standard Deviations for Key Variables by Condition in Study 2A: Before and After Excluding Outliers

	Full Sample ($N = 415$)				Trimmed Sample ($N = 402$)			
	N	Firm sustainability image	CSR	CSI	N	Firm sustainability image	CSR	CSI
Social underperformance	137	2.33 (1.22)	5.68 (2.80)	30.50 (25.00)	132	2.34 (1.23)	5.80 (2.75)	28.07 (22.01)
No social performance feedback	140	4.79 (0.91)	4.93 (2.41)	35.84 (22.54)	137	4.78 (0.91)	4.92 (2.36)	34.61 (21.14)
Social overperformance	138	6.24 (0.77)	4.86 (2.01)	43.45 (21.59)	133	6.25 (0.77)	4.96 (1.93)	41.45 (19.29)

Note. Means are reported in each cell, with standard deviations reported in parentheses.

Table B4. Descriptive Statistics and Correlations of Key Variables in Study 2A: Before and After Excluding Outliers

	Full Sample ($N = 415$)				Trimmed Sample ($N = 402$)			
	M	SD	Firm sustainability image	CSR	M	SD	Firm sustainability image	CSR
Firm sustainability image	4.46 (4.46)	1.88 (1.88)	--	--	4.46	1.88	--	--
CSR	5.16 (5.22)	2.45 (2.40)	-.193***	--	5.22	2.40	-.212***	--
CSI	36.61 (34.72)	23.63 (21.49)	.237***	-.286***	34.72	21.49	.278***	-.252***

Note. M = Mean. SD = Standard deviation. * $p < .1$; ** $p < .05$; *** $p < .01$

Appendix C: Supplementary Analyses for Study 2B

Attention Check Analyses

Ten participants failed at least one of the three attention checks in Study 2B. Before testing the robustness of our effects with and without these failed attention check cases, we examined whether these participants differed significantly in demographic or key study variables compared to those who passed all attention checks. Independent sample t -tests and chi-square tests revealed that those who failed one or more attention checks did not significantly differ from those who passed all attention checks. Specifically, their age, $t(23.99) = .98, p = .339$, race, $\chi^2(9) = 4.42, p = .881$, gender, $\chi^2(1) = 1.76, p = .185$, education, $\chi^2(3) = 6.67, p = .083$, and managerial experience, $t(78) = .14, p = .893$, did not differ significantly. They also did not statistically differ in their responses to key study variables across the three conditions.

To test the robustness of our findings, we conducted the same analyses on a trimmed sample that excluded these 10 participants ($N = 71$). Table C1 contains the means and standard deviations of key variables by condition for both the full and trimmed samples, and Table C2 provides the corresponding descriptive statistics and within-person correlations of key variables. We first conducted a repeated measures ANOVA to examine the effectiveness of our manipulation. Mauchly's test of sphericity indicated that the assumption of sphericity was met, $\chi^2(2) = 1.94, p < .001$, and the ANOVA results revealed a significant effect of condition on the manipulation check, $F(2, 140) = 824.56, p < .001, \eta^2 = .92$. Planned contrasts showed that social performance was rated as higher in the overperformance condition ($M = 6.09, SE = .70$) compared to the control, $M = 3.60, SE = .75, F(1,70) = 614.96, p < .001, \eta^2 = .90$, and the underperformance condition reported lower social performance than the control, $M = 1.99, SE = .78, F(1,70) = 286.39, p < .001, \eta^2 = .80$.

Perceived Firm Sustainability Image. The sphericity assumption was also met for perceived firm sustainability image, $\chi^2(2) = 3.23, p = .199$, and results from a repeated-measures ANOVA revealed a significant effect of condition on perceived firm image, $F(2, 140) = 176.17, p < .001, \eta^2 = .72$. Planned contrasts showed that participants perceived the firm's sustainability image to be higher in the overperformance condition ($M = 5.94, SD = .94$) compared to the control, $M = 4.62, SD = .1.20, F(1,70) =$

58.36, $p < .001$, $\eta^2 = .46$, and the underperformance condition reported lower perceived firm sustainability image than the control condition, $M = 2.74$, $SD = 1.26$, $F(1,70) = 149.03$, $p < .001$, $\eta^2 = .68$.

Serial Mediation Analysis. We tested the proposed serial mediation using the same methods as in the full sample. Results indicated that managers' perceived firm sustainability image positively predicted their own moral self-image ($b = 0.55$, $SE = .06$, $p < .001$), indicating that viewing the firm as more sustainable enhanced managers' sense of personal moral standing. Importantly, the indirect effect of firm underperformance on managers' moral self-image through perceived firm sustainability image was significant (Estimate = -1.03 , $SE = .14$, 95% CI [-1.31 , $-.76$]), as was the indirect effect of overperformance (Estimate = $.71$, $SE = .11$, 95% CI [$.52$, 0.94]). For CSR engagement, moral self-image was negatively related to CSR ($b = -0.24$, $SE = .11$, $p = .017$).

Moreover, the serial indirect effect of underperformance on subsequent CSR through firm sustainability image and managers' moral self-image was positive and significant (Estimate = $.31$, $SE = .13$, 95% CI [$.08$, $.59$]), consistent with the full sample. Likewise, the serial indirect effect of social overperformance also provided support for moral licensing, consistent with the full sample (Estimate = $-.21$, $SE = .09$, 95% CI [$-.41$, $-.06$]). For CSI engagement, moral self-image was not significantly related to CSI ($b = -0.20$, $SE = 1.25$, $p = .44$). As a result, the proposed serial indirect effects on CSI engagement were not significant for either underperformance or overperformance, resembling the results found with the full sample.

Furthermore, we identified CSI outliers using the same method as in Study 2A. Removing these outliers from the sample did not materially affect our results. Moral self-image continued to have no significant effect on CSI ($b = -0.49$, $SE = 1.07$, $p = .33$), and thus the proposed serial indirect effects on CSI engagement remained not significant.

Analyses Ruling Out Alternative Explanations

Stakeholder expectations. To rule out the possibility that strategic concerns about stakeholder expectations—rather than moral self-regulation—were driving the effects, we included a two-item measure of stakeholder expectations at the between-person level: “Maintaining CSR at a moderate level is

the safest strategic choice for the firm” and “If CSR performance deviates too far from the organization’s goal, stakeholders are more likely to intervene.” Results controlling for stakeholder expectations remained the same in that managers’ perceived firm sustainability image continued to positively predict their own moral self-image ($b = 0.62$, $SE = .06$, $p < .001$), and the indirect effect of firm underperformance on managers’ moral self-image was significant (Estimate = -1.14 , $SE = .14$, 95% CI [-1.43 , $-.88$]), as was the indirect effect of overperformance (Estimate = $.78$, $SE = .12$, 95% CI [$.57$, 1.02]).

Moral self-image also remained negatively related to CSR ($b = -.24$, $SE = .11$, $p = .018$), and the serial indirect effect of underperformance on CSR through firm sustainability image and managers’ moral self-image stayed positive and significant (Estimate = $.27$, $SE = .13$, 95% CI [$.02$, $.54$]). Likewise, the serial indirect effect of social overperformance also showed support for moral licensing consistent with the full sample (Estimate = $-.18$, $SE = .09$, 95% CI [$-.38$, $-.01$]). For CSI engagement, moral self-image was not significantly related to CSI ($b = -0.25$, $SE = 1.16$, $p = .42$), as were the proposed serial indirect effects on CSI, mirroring the results found with the full sample. Thus, controlling for stakeholder expectations did not materially alter our results.

Felt responsibility. We also examined felt responsibility as a potential boundary condition, proposing that managers who feel more responsible for the social performance feedback they receive are more likely to internalize that feedback and, consequently, exhibit stronger moral-cleansing and moral-licensing responses. We manipulated felt responsibility by randomly assigning participants to assume high responsibility for performance feedback or low responsibility across each of the three within-person conditions. Participants in the high condition reported feeling slightly more responsible for the performance feedback ($M = 5.81$, $SD = .93$) than participants in the low condition ($M = 5.30$, $SD = 1.20$, $t(79) = 2.15$, $p = .034$). However, including the felt responsibility manipulation check as a within-person control variable in the serial mediation analysis did not alter the results. Managers’ perceived firm sustainability image continued to positively predict their own moral self-image ($b = 0.62$, $SE = .06$, $p < .001$), and the indirect effect of firm underperformance on managers’ moral self-image was significant (Estimate = -1.13 , $SE = .14$, 95% CI [-1.43 , $-.87$]), as was the indirect effect of overperformance

(Estimate = .78, SE = .11, 95% CI [.57, 1.01]).

Moral self-image also remained negatively related to CSR ($b = -.24$, $SE = .11$, $p = .017$), and the serial indirect effect of underperformance on CSR through firm sustainability image and managers' moral self-image stayed positive and significant (Estimate = .26, SE = .13, 95% CI [.02, .54]). Likewise, the serial indirect effect of social overperformance also provided support for moral licensing, consistent with the full sample (Estimate = $-.18$, SE = .09, 95% CI [$-.38$, $-.01$]). For CSI engagement, moral self-image was not significantly related to CSI ($b = -0.20$, $SE = 1.16$, $p = .43$), as were the proposed serial indirect effects on CSI, mirroring the results found with the full sample. Thus, felt responsibility did not meaningfully affect the results.

Table C1. Means and Standard Deviations for Key Variables by Condition in Study 2B: Before and After Excluding Attention Check Failures

	Full Sample ($N = 243$, 81 participants X 3 conditions)					Trimmed Sample ($N = 213$, 71 participants X 3 conditions)				
	N	Firm sustainability image	Moral self-image	CSR	CSI	N	Firm sustainability image	Moral self-image	CSR	CSI
Social underperformance	81	2.80 (1.28)	3.27 (1.33)	6.36 (2.43)	25.02 (24.87)	71	2.74 (1.26)	3.13 (1.27)	6.33 (2.52)	25.31 (24.58)
Control group	81	4.63 (1.19)	5.35 (1.52)	5.82 (1.84)	27.86 (23.52)	71	4.62 (1.19)	5.34 (1.45)	5.80 (1.92)	27.65 (23.97)
Social overperformance	81	5.90 (1.10)	6.65 (1.19)	5.43 (1.71)	22.57 (23.91)	71	5.94 (.94)	6.69 (.89)	5.40 (1.78)	21.65 (22.47)

Note. Means are reported in each cell, with standard deviations reported in parentheses.

Table C2. Within-Person Correlations of Key Variables in Study 2B: Before and After Excluding Attention Check Failures

	Full Sample ($N = 243$)			Trimmed Sample ($N = 213$)		
	Firm sustainability image	Moral self-image	CSR	Firm sustainability image	Moral self-image	CSR
Moral self-image	.826***	--	--	.832***	--	
CSR	-.321***	-.380***	--	-.330***	-.414***	
CSI	-.079	-.084	.094	-.103	-.104	.112

Note. All variables are person-centered. * $p < .1$; ** $p < .05$; *** $p < .01$

Appendix D: Constructive Replication of Study 2B with Revised Decision Tasks

In this supplementary experiment, we conducted a constructive replication of Study 2B to address a potential alternative interpretation of the CSR allocation measure. In Study 2B, participants decided how much of the firm’s discretionary budget to allocate to CSR initiatives after receiving social performance feedback. However, the original wording stated that CSR allocations “compete with other strategic priorities,” such as product development, business expansion, and operational efficiency. Although this framing was intended to reflect the reality that CSR decisions often involve tradeoffs with other organizational investments, it also introduced a potential ambiguity. Participants may have interpreted social overperformance or underperformance not only as feedback about the firm’s social performance, but also as indirect information about whether the firm had over- or under-invested in other strategic priorities. Consequently, changes in CSR allocation could have reflected a general strategic rebalancing logic rather than the moral rebalancing process proposed by our theory.¹

To address this concern, we conducted a supplementary experiment with participants recruited through Prolific. The current study retained the core design of Study 2B while revising the decision task to provide a more conservative test of our theorizing.² Specifically, we replaced the original language emphasizing CSR as competing with “other strategic priorities” with broader language describing the CEO’s responsibility for allocating discretionary funds across multiple organizational activities, including product development, business growth, employee support, and societal contributions. We also clarified that funds allocated to CSR initiatives could not be used for other organizational purposes, without implying that the firm had under- or over-invested in any particular non-CSR activity. Thus, the revised measure preserves the fundamental resource-allocation tradeoff inherent in CSR decision making while reducing the likelihood that participants would interpret the feedback as a signal about neglected strategic priorities outside the social performance domain. Additionally, we also revised the CSI decision measure

¹ We thank one of our reviewers for raising this important point.

² The social performance feedback manipulation check was omitted because it had already been validated in Study 2B.

to use more general language around the lower-cost supplier. All materials, data, and syntax files for this constructive replication of Study 2B are available on our Open Science Framework (OSF) repository.³

This constructive replication serves two purposes. First, it examines whether the findings of Study 2B are robust to an alternative operationalization of the CSR and CSI decision tasks. Second, and more importantly, it strengthens the construct validity of the CSR allocation measure by testing whether social performance feedback continues to influence subsequent CSR decisions when the decision context is less likely to invite a strategic portfolio-balancing interpretation.

Participants and Design

We recruited 90 participants through Prolific to complete an online within-person experiment. Prior to analysis, we excluded two participants who did not meet our preregistered selection criteria (<https://aspredicted.org/ek6jn3.pdf>), resulting in a final sample of 88 participants ($M_{\text{age}} = 38.73$; $SD = 9.63$; 50% Male; 77.3% White). Participants reported an average of 8.56 years of managerial experience ($SD = 6.82$) across a variety of industries. Following the procedures of Study 2B, participants imagined themselves as the CEO of three different companies and completed the same set of questions after each scenario, resulting in 264 within-person observations. After reviewing each company's CSR performance report, participants reported their perceptions of the firm's sustainability image, completed a measure of moral self-image, and responded to the revised CSR and CSI decision tasks.

Results

Table D1 presents the means and standard deviations of the key variables across each condition, and Table D2 shows a within-person correlation matrix for the study's key variables. We ran all analyses with and without the eight participants who failed one or more attention checks, and the results are not substantively different. Thus, we report the results using all 88 participants, consistent with our prior studies.

Main Effect Analyses. A repeated measures ANOVA showed that CSR, $\chi^2(2) = 16.10$, $p < .001$,

³ https://osf.io/fc78h/overview?view_only=f01b6eae074d493dbace7c762c08e3cf

and CSI, $\chi^2(2) = 8.14, p = .017$, failed to meet the sphericity assumption. Thus, Greenhouse-Geisser estimates were used to test the within-person effects. Results revealed a significant within-person effect on CSR, $F(1.71, 148.62) = 4.50, p = .017, \eta^2 = .05$, with participants showing greater CSR after receiving social *underperformance* feedback, $M = 6.37, SD = 2.47$, compared to the control, $M = 5.76, SD = 1.48$, $F(1,87) = 7.36, p = .008, \eta^2 = .08$, supporting the moral cleansing hypothesis. However, CSR did not differ significantly following social *overperformance* feedback ($M = 5.69, SD = 1.89$) relative to the control condition, $F(1,87) = 0.09, p = .766, \eta^2 = .00$. In addition, CSI showed no significant within-person effect, $F(1.83, 159.59) = 1.65, p = .198, \eta^2 = .019$.

These results are mostly consistent with the original findings in Study 2B, with one exception: the overall contrast between social overperformance and the control condition was not significant for CSR. Given the previously noted concern that CSR allocation decisions may be interpreted as part of a broader resource-balancing process, we conducted exploratory analyses to examine whether the order in which the dependent variables were presented influenced the results.⁴ This analysis is theoretically important because participants made two morally relevant resource-allocation decisions in sequence. When the CSR decision was presented first, it represented participants' initial behavioral response to the social performance feedback. In contrast, when the CSI decision was presented first, participants may have already engaged in moral or strategic balancing through the supplier-selection decision before making the CSR allocation decision.

A repeated-measures ANOVA revealed a significant main effect of performance feedback on CSR, $F(2, 87.23) = 4.08, p = .020$, but no main effect of dependent-variable order on CSR, $F(1, 196.08) = 0.61, p = .438$. Interestingly, the interaction between performance-feedback condition and dependent-variable order was significant, $F(2, 171.46) = 4.49, p = .013$, indicating that the effect of social performance feedback on CSR depended on whether the CSR or CSI decision was presented first. Table D3 reports the means and standard errors of the dependent variables across the three performance-

⁴ This exploratory analysis was not preregistered and should therefore be interpreted as an examination of a potential boundary condition rather than a confirmatory test.

feedback conditions by decision order (CSR first vs. CSI first). Figure D1 plots this interaction.

To further probe this interaction, we conducted mixed-model analyses separately by dependent-variable order. When CSR was presented first, the effect of the performance feedback scenario on CSR was significant, $F(2,41.40) = 7.23, p = .002$. Consistent with moral cleansing, participants allocated more to CSR in the social underperformance condition, $M = 6.93, SE = 0.32$, than in the control, $M = 5.98, SE = 0.20, p = .001$. Consistent with moral licensing, participants allocated less to CSR in the social overperformance condition, $M = 5.34, SE = 0.27$, than in the control condition, $p = .034$. Thus, when CSR was participants' first behavioral response to the feedback, the results supported both the moral cleansing and moral licensing predictions.

When the CSI decision was presented first, however, the predicted CSR pattern was attenuated. This finding suggests that participants' CSR allocations may have been influenced by the preceding supplier decision, consistent with the possibility that participants engaged in balancing across the two morally relevant choices. We interpret these order effects cautiously, as they emerged from exploratory analyses and the study was not designed to isolate sequential decision-making processes. Nevertheless, the pattern is informative in light of concerns that resource-allocation outcomes may reflect broader balancing considerations. Notably, the clearest test of the CSR decision—when it was presented first and therefore not preceded by another morally relevant choice—replicated the predicted patterns of moral cleansing and moral licensing. Future research could more directly investigate how decision order influences moral self-regulation across multiple CSR- and CSI-related decisions.

Perceived Firm Sustainability Image. For perceived firm sustainability image, the sphericity assumption was not met, $\chi^2(2) = 31.26, p < .001$, and the Greenhouse-Geisser estimates from a repeated-measures ANOVA revealed a significant within-person effect of social performance feedback on perceived firm image, $F(1.53,133.36) = 307.532, p < .001, \eta^2 = .78$, consistent with the original findings. Planned contrasts showed that the *underperformance* condition, $M = 2.79, SD = 1.28$, elicited lower perceptions of firm sustainability image than the control condition, $M = 4.67, SD = 1.05, F(1, 87) = 206.42, p < .001, \eta^2 = .70$, and the *overperformance* condition, $M = 6.18, SD = .85$, led to higher firm

sustainability image perceptions compared to the control, $F(1, 87) = 214.74, p < .001, \eta^2 = .71$, consistent with the original Study 2B findings.

Serial Mediation Analysis. We used the same serial mediation model using Bayesian estimation in MPlus (version 9), as in Study 2B, to examine the extent to which firm social performance feedback influences CSR and CSI sequentially through managers' perceived firm sustainability image and their own moral self-image (for summary of results, see Figure D2).

Consistent with Study 2B, results show that managers' perceived firm sustainability image positively predicted their moral self-image, $b = 0.73$, 95% CI [0.63, 0.83]. In addition, the indirect effect of social underperformance, relative to the control condition, on managers' moral self-image through perceived firm sustainability image was negative ($Estimate = -1.35$, 95% CI [-1.61, -1.11]), whereas the indirect effect of social overperformance was positive ($Estimate = 1.13$, 95% CI [0.92, 1.36]).

We also found that moral self-image credibly predicted CSR, $b = -0.16$, 95% CI [-0.29, -0.02]. Moreover, the serial indirect effect of *underperformance* on subsequent CSR through firm sustainability image and managers' moral self-image was positive ($Estimate = 0.21$, 95% CI [0.03, 0.41]), supporting H3a. Similarly, the serial indirect effect of *overperformance* on CSR through firm sustainability image and managers' moral self-image was negative ($Estimate = -0.18$, 95% CI [-0.34, -0.03]), supporting H3b.

Moral self-image also credibly predicted CSI, $b = -1.66$, 95% CI [-3.23, 0.11], although in the opposite direction from that hypothesized. Moreover, the serial indirect effect of *underperformance* on CSI was negative ($Estimate = 2.22$, 95% CI [0.15, 4.49]), while the serial indirect effect of *overperformance* on CSI was negative ($Estimate = -1.85$, 95% CI [-3.76, -0.13]), which suggests moral consistency rather than moral cleansing or licensing. Consistent with Study 2B, therefore, we found no evidence that CSI served as a mediating pathway for moral cleansing or moral licensing.

Discussion. Overall, the findings from this supplementary experiment provide additional support for the robustness of the proposed moral self-regulation mechanism for CSR, while suggesting that the mechanism may operate differently for CSI. This constructive replication was designed to address the possibility that the original wording of Study 2B's CSR allocation task encouraged a strategic-balancing

interpretation by explicitly describing CSR investments as competing with specific alternative strategic priorities. In the revised study, we preserved the realistic resource-allocation context while removing language that could imply that social performance feedback signaled prior over- or under-investment in particular non-CSR activities.

Across the revised procedures, social performance feedback again influenced managers' perceptions of firm sustainability image and their own moral self-image, and the indirect effects through the CSR pathway were substantively consistent with the moral self-regulation account. Specifically, social underperformance increased CSR allocations relative to the control condition, consistent with moral cleansing. Although the overall direct effect of social overperformance on CSR did not fully replicate the moral licensing effect observed in Study 2B, exploratory analyses of decision order revealed that the predicted licensing pattern emerged when CSR was presented as participants' first behavioral response to the feedback. Under this condition, participants allocated more resources to CSR following social underperformance and fewer resources to CSR following social overperformance, relative to the control condition, providing support for both the moral cleansing and moral licensing predictions.

Consistent with the findings of Study 2B, the hypothesized effects on CSI did not emerge. Whereas the original Study 2B yielded no significant effects on CSI, this replication revealed a pattern of moral consistency rather than compensatory behavior. These findings suggest that, in the present decision context, participants were more likely to respond to social performance feedback by adjusting their willingness to invest in CSR than by compensating through changes in socially questionable supplier practices.

Taken together, these findings alleviate concerns that the wording of "strategic priorities" drove the results observed in Study 2B. The revised study provides additional evidence that the proposed moral self-regulation process is robust across alternative operationalizations of CSR decisions. At the same time, the observed order effects suggest that future research may benefit from more directly examining how moral self-regulation unfolds across sequences of related organizational decisions.

Table D1. Means and Standard Deviations of Key Variables Across Conditions in Constructive Replication of Study 2B

	Firm sustainability image	Moral self-image	CSR	CSI
Underperformance	2.79 (1.28)	3.08 (1.45)	6.37 (2.47)	35.12 (30.14)
Neutral	4.64 (1.05)	4.86 (1.21)	5.76 (1.48)	33.02 (24.34)
Overperformance	6.18 (0.85)	6.22 (0.85)	5.69 (1.89)	29.87 (26.94)

Note. $N = 264$.

Table D2. Within-Person Correlations for Key Variables in Constructive Replication of Study 2B

	Firm sustainability image	Moral self-image	CSR
Moral self-image	.89***	--	
CSR	-.10*	-.10	--
CSI	-.06	-.00	-.04

Note. $N = 264$. * $p < .1$; ** $p < .05$; *** $p < .01$

Table D3. Means and Standard Deviations for Dependent Variables by Dependent Variable Order in Constructive Replication of Study 2B

	CSR First		CSI First	
	CSR	CSI	CSR	CSI
Underperformance	6.93 (0.32)	37.59 (29.50)	6.24 (0.82)	32.30 (4.33)
Neutral	5.98 (0.20)	31.97 (29.38)	5.48 (0.75)	33.76 (3.30)
Overperformance	5.34 (0.27)	31.76 (29.36)	6.17 (0.77)	30.65 (3.99)

Note. $N = 264$.

Figure D1. Interaction Plot from Constructive Replication of Study 2B

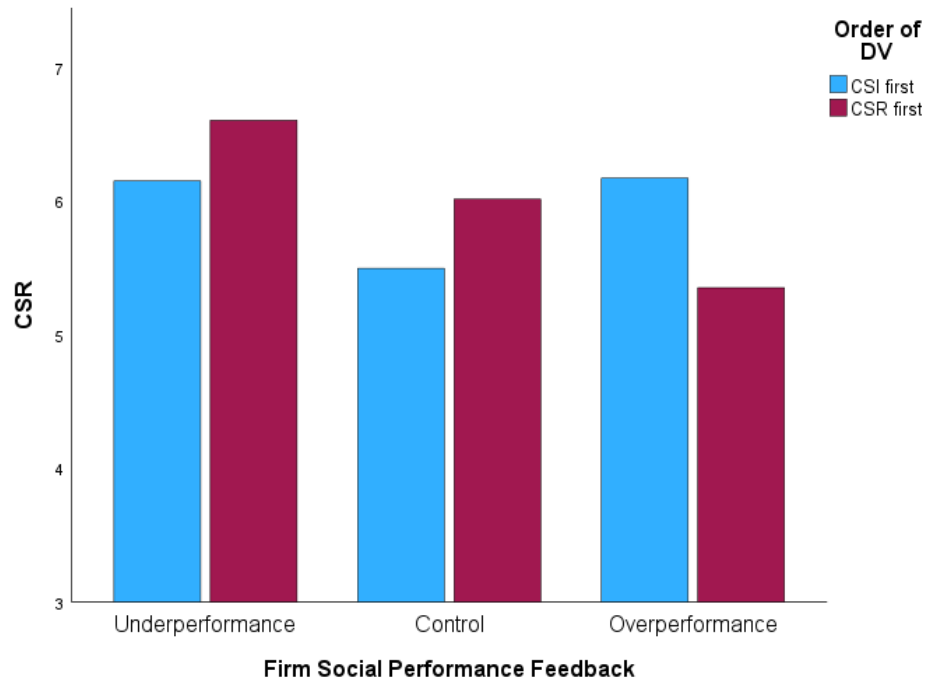
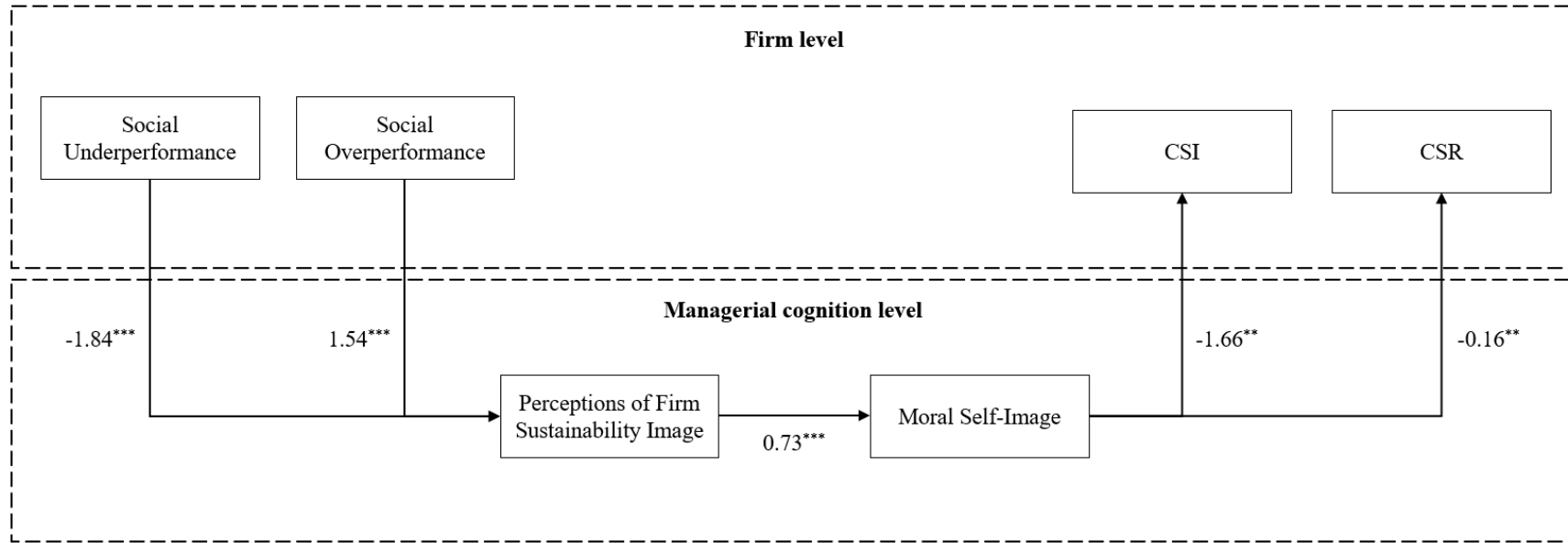


Figure D2. Summary of Full Model Results in Constructive Replication of Study 2B



Note. $N = 264$. * $p < .1$; ** $p < .05$; *** $p < .01$

Appendix E: Supplementary Analyses for Study 2C

Attention Check Analyses

We included two attention check questions in Study 2C. Six participants failed both attention checks, and 12 participants failed one. We first examined whether these 18 participants differed significantly on demographic or key study variables compared to those who passed all attention checks. Independent sample *t*-tests and chi-square tests revealed that those who failed one or more attention checks did not significantly differ from those who passed all attention checks. Specifically, the two groups did not differ in age ($p = .938$), race ($p = .250$), education ($p = .122$), managerial experience ($p = .851$), or their responses to key study variables, including the manipulation check ($p = .752$), CSR ($p = .285$), and CSI ($p = .202$). A chi-square analysis further confirmed that the failed attention checks were not unequally distributed across experimental conditions (9.3% of control, 9.0% of overperformance, and 4.9% of underperformance, $\chi^2 = 1.39$, $p = .500$). There were differences in gender among those who failed versus those who passed attention checks, $\chi^2 = 7.83$, $p = .005$, with 88.9% of participants who failed attention checks being male.

To test the robustness of our findings, we conducted the same analyses on a trimmed sample that excluded these 18 participants ($N = 217$). Table E1 contains the means and standard deviations of key variables by condition for both the full and trimmed samples, and Table E2 provides the corresponding descriptive statistics and correlations of key variables. Consistent with the full sample, a Welch's ANOVA revealed a significant effect of the condition on the manipulation check, $F(2, 137.06) = 53.68$, $p < .001$, $\eta^2 = .35$. Planned contrasts indicated that participants in the high moral self-image condition reported higher scores on the manipulation check items ($M = 6.07$, $SD = .89$) than those in the control ($M = 5.17$, $SD = 1.08$; $p < .001$) and low moral self-image ($M = 3.68$, $SD = 1.88$; $p < .001$) conditions. The difference between the low and control conditions was also statistically significant ($p < .001$).

Consistent with the full sample analysis and our theoretical prediction that higher (lower) moral self-image decreases (increases) CSR, moral self-image was negatively correlated with CSR ($r = -.34$, $p < .01$). Moral self-image's effect on CSR also remained, $F(2, 141.75) = 5.38$, $p = .006$. Planned contrasts

revealed that participants in the high moral self-image condition reported lower CSR ($M = 4.71$, $SD = 1.69$) than those in the low moral self-image condition ($M = 5.72$, $SD = 2.10$, $t(144.88) = -3.21$, $p = .002$). Consistent with the full sample analysis, this pattern was driven primarily by moral cleansing response: CSR was higher in the low moral self-image condition than the control condition, $M = 4.97$, $SD = 1.41$, $t(135.40) = -2.54$, $p = .012$, whereas CSR in the high moral self-image condition did not significantly differ from the control, $t(134.40) = .97$, $p = .335$.

Additionally, consistent with our theoretical expectation that higher (lower) moral self-image increases (decreases) CSI, moral self-image was positively correlated with CSI ($r = .21$, $p < .01$). Planned contrasts revealed a marginal difference in CSI between the high moral self-image ($M = 46.80$, $SD = 28.50$) and low moral self-image conditions, $M = 38.65$, $SD = 28.50$, $t(214) = 1.83$, $p = .069$. Although neither the low moral self-image condition, $t(214) = .45$, $p = .656$, nor the high moral self-image, $t(198) = -1.33$, $p = .184$, differed significantly from the control condition ($M = 40.66$, $SD = 23.97$), the pattern of group means was directionally consistent with predictions. Thus, the results of the trimmed sample are consistent with those of the full sample.

Additional Analyses

Following our preregistered analytic plan, we examined whether agentic and communal values moderated the relationship between moral self-image and either CSR or CSI. Results showed that agentic values were not significantly related to CSR ($r = .02$, $p = .792$) but were positively associated with CSI ($r = .14$, $p = .035$). Conversely, communal values were positively associated with CSR ($r = .27$, $p < .001$) and negatively related to CSI ($r = -.15$, $p = .029$). Despite these theoretically meaningful patterns, neither agentic nor communal values significantly moderated the effects of moral self-image on CSR or CSI. Moreover, including agentic and communal values as covariates in separate univariate models yielded results that were substantively identical to those from models without these predictors.

Importantly, these null interactions provide initial support for our theorizing that agentic and communal values may not directly shape how moral self-image translates into CSR or CSI responses

(Step 3 in Figure 2B: *moral balancing*). Instead, these values appear to operate earlier in the process by influencing the extent to which individuals internalize or identify with the firm's sustainability image (Step 2 in Figure 2B: *image internalization*). We acknowledge, however, that this interpretation remains tentative, and further research is needed to examine this possibility more directly.

Table E1. Means and Standard Deviations for Key Variables by Condition in Study 2C: Before and After Excluding Attention Check Failures

	Full Sample ($N = 235$)			Trimmed Sample ($N = 217$)		
	N	CSR	CSI	N	CSR	CSI
High moral self-image	78	4.78 (1.69)	41.90 (28.31)	71	4.71 (1.69)	46.80 (28.50)
Neutral moral self-image	75	4.90 (1.46)	41.90 (24.63)	68	4.97 (1.41)	40.66 (23.97)
Low moral self-image	82	5.64 (2.13)	38.17 (28.14)	78	5.72 (2.10)	38.65 (28.50)

Note. Means are reported in each cell, with standard deviations reported in parentheses.

Table E2. Descriptive Statistics and Correlations of Key Variables in Study 2C: Before and After Excluding Attention Check Failures

	Full Sample ($N = 235$)				Trimmed Sample ($N = 217$)			
	M	SD	CSR	CSI	M	SD	CSR	CSI
CSR	5.12	1.83	--	--	5.15	1.82	--	--
CSI	42.60	27.32	-.223**	--	41.95	27.26	-.250***	--
Moral self-image	4.92	1.69	-.322***	.171***	4.93	1.70	-.340***	.209***

Note. M = Mean. SD = Standard deviation. * $p < .1$; ** $p < .05$; *** $p < .01$

Appendix F: Supplementary Undergraduate Sample for Study 2C

To assess the robustness of Study 2C's findings, we recruited a substantially larger undergraduate sample, increasing statistical power and enabling a more reliable test of the hypothesized effects.

Participants and Design

We recruited 689 undergraduate business students from multiple U. S. universities ($M_{\text{age}} = 21.12$; $SD = 2.33$; 50.9% Male; 66.8% White; 27 unreported demographics). Although 42 participants (6.1%) did not meet our preregistered exclusion criteria (<https://aspredicted.org/f8h79y.pdf>), we report the results for the full sample because they are consistent before and after exclusions. Nearly half of the students were not employed (54.6%), while many others held part-time or full-time jobs in roles mostly at the individual contributor level, indicating limited managerial experience. Accordingly, this student sample serves as a robustness check and a conservative test of our hypotheses. Participants followed the same experimental design and procedures as those used in the main managerial sample in Experiment 2C.

Results

Supporting the successful manipulation of moral self-image and consistent with the results from the managerial sample, a Welch's one-way ANOVA revealed a significant effect of the condition, $F(2, 450.66) = 203.24, p < .001, \eta^2 = 0.40$. Planned contrasts indicated that participants in the high moral self-image condition reported higher scores on the manipulation check items ($M = 6.03, SD = 1.12$) than those in the control, $M = 4.90, SD = 1.02; t(450.71) = -11.31, p < .001$, and participants in the low moral self-image ($M = 3.61, SD = 1.46$) reported lower scores than the control, $t(415.88) = 10.98, p < .001$.

Planned contrasts within an ANOVA framework, $F(2, 685) = 4.12, p = .017, \eta^2 = .012$, revealed that CSR in the high moral self-image condition ($M = 5.85, SD = 1.56$) was not significantly different than CSR in the low moral self-image condition, $M = 6.00, SD = 1.51, t(685) = -1.09, p = .277$. However, we found evidence of moral cleansing: CSR was higher in the low moral self-image condition compared to the control condition, $M = 5.61, SD = 1.33, t(685) = -2.85, p = .005$, whereas the high moral self-image condition did not significantly differ in CSR compared to the control, $t(685) = -1.75, p = .080$. This pattern of results is broadly consistent with the findings from the managerial sample, providing support

for moral cleansing through CSR. However, unlike the managerial sample, ANOVA results indicated no significant effect of moral self-image on CSI, $F(2, 685) = .32, p = .728$, and planned contrasts showed no significant difference in CSI across the three conditions. These results indicate that the moral self-regulatory process, particularly with respect to CSI, may be more pronounced among individuals with managerial experience. We note, however, that this interpretation is tentative, and future research is needed to examine this possibility more directly.

Attention Check Analyses. Forty-two students failed at least one attention check (37 failed only one, and five failed both checks). We examined whether these 42 participants differed significantly on demographic or key study variables compared to those who passed all attention checks. Independent sample *t*-tests and chi-square tests revealed that those who failed one or more attention checks did not significantly differ from those who passed all attention checks. Specifically, their age ($p = .202$), gender ($p = .179$), and managerial experience ($p = .523$) did not differ significantly. Those who failed attention checks were predominantly White (52.8%; $\chi^2 = 14.09, p = .029$). A chi-square analysis further confirmed that the failed attention checks were not unequally distributed across experimental conditions (6.2% of control, 4.4% of high moral self-image, and 7.7% of low moral self-image, $\chi^2 = 2.28, p = .320$). While they did not statistically differ in their responses to the manipulation check ($p = .137$), they showed different means for CSI ($p = .024$) and CSR ($p = .048$).

To test the robustness of our findings, we conducted the same analyses on a trimmed sample that excluded these 42 participants ($N = 647$). Table F1 contains the means and standard deviations of key variables by condition for the full and trimmed samples, and Table F2 provides descriptive statistics and correlations of key variables. A Welch's one-way ANOVA revealed continued support for a successful manipulation, $F(2, 420.51) = 219.55, p < .001, \eta^2 = .44$, with participants in the high moral self-image condition reporting higher scores on the manipulation check items ($M = 6.08, SD = 1.06$) than those in the control, $M = 4.95, SD = .99, t(429.31) = -11.40, p < .001$, and participants in the low moral self-image reporting lower scores than the control, $M = 3.54, SD = 1.45, t(378.93) = 11.77, p < .001$.

Consistent with the full-sample analysis, ANOVA results showed a significant effect of moral

self-image on CSR, $F(2, 644) = 5.32, p = .005$, and CSR in the high moral self-image condition ($M = 5.79, SD = 1.51$) was not significantly different from that in the low moral self-image condition ($M = 6.02, SD = 1.49, t(644) = -1.66, p = .098$). As in the full-sample analysis, CSR was higher in the low moral self-image condition than the control condition, $M = 5.56, SD = 1.32, t(644) = -3.26, p = .001$, supporting a moral cleansing effect for CSR. The high moral self-image condition did not significantly differ in CSR compared to the control, $t(644) = -1.62, p = .106$. An ANOVA also indicated no significant effect of moral self-image on CSI, $F(2, 644) = .50, p = .609$. Thus, removing participants who failed attention checks did not change the pattern of findings observed in the full student sample.

Table F1. Means and Standard Deviations for Key Variables by Condition in Study 2C (Student Sample): Before and After Excluding Attention Check Failures

	Full Sample ($N = 689$)			Trimmed Sample ($N = 647$)		
	N	CSR	CSI	N	CSR	CSI
High moral self-image	229	5.85 (1.56)	44.80 (25.32)	219	5.79 (1.51)	44.14 (25.17)
Neutral moral self-image	226	5.61 (1.33)	46.60 (23.07)	213	5.56 (1.32)	46.42 (23.44)
Low moral self-image	233	6.00 (1.51)	45.65 (23.75)	215	6.02 (1.49)	44.93 (23.66)

Note. Means are reported in each cell, with standard deviations reported in parentheses.

Table F2. Descriptive Statistics and Correlations of Key Variables in Study 2C (Student Sample): Before and After Excluding Attention Check Failures

	Full Sample ($N = 689$)				Trimmed Sample ($N = 647$)			
	M	SD	CSR	CSI	M	SD	CSR	CSI
CSR	5.82	1.48	--	--	5.79	1.45	--	--
CSI	45.68	24.04	-.132***	--	45.15	24.09	-.153***	--
Moral self-image	4.84	1.57	-.041	.010	4.86	1.58	-.073	.016

Note. M = Mean. SD = Standard deviation. * $p < .1$; ** $p < .05$; *** $p < .01$

Appendix G: Constructive Replication of Study 2C with Revised Moral Self-Image Manipulation and Revised Decision Tasks

In this supplementary set of experiments, we examined the robustness and construct validity of Study 2C using revised procedures and measures. First, the original moral self-image manipulation may have incorporated language that extended beyond simply inducing relatively high versus low moral self-image by characterizing participants' moral standing as a "surplus" or a "deficit." Although this wording was intended to reflect the logic of moral self-regulation theory, it may also have conveyed that participants' moral standing was excessively high, insufficiently low, or otherwise in need of correction.⁵ Second, the original Study 2C materials included references to the CEO's competing responsibilities and strategic priorities, which may have encouraged participants to interpret the dependent variables as part of a broader strategic resource-balancing problem rather than as responses primarily driven by moral self-regulation.

To address these concerns, we conducted two supplementary experiments with participants recruited through Prolific. First, we developed and validated a revised moral self-image manipulation that more directly induced relatively high versus low moral self-image without using language such as "excessive," "surplus," or "deficit." Second, we conducted a replication of Study 2C using this revised manipulation along with revised CSR and CSI decision measures that removed language emphasizing competing strategic priorities. Additionally, we removed language indicating a resource trade-off, specifically the phrase "funds allocated to CSR initiatives could not be used for other organizational purposes." We also revised the CSI decision measure to use more general language around the lower-cost supplier. Together, these studies provide a more conservative test of the proposed moral self-regulation mechanism by examining whether variations in moral self-image continue to shape subsequent CSR and CSI decisions when these potential alternative interpretations are minimized.

Alternative Manipulation of Moral Self-Image

⁵ We thank one of our reviewers for raising this important point.

We designed the original moral self-image manipulation based on moral self-regulation theory and prior operationalizations of the construct. Moral self-image is theorized to fluctuate in response to prior behavior, feedback, and social comparison. Thus, prior socially responsible behavior can strengthen people's moral self-image, whereas prior socially irresponsible behavior can weaken people's moral self-image. Consistent with this reasoning, Jordan et al.'s (2015) moral self-image scale asks individuals to compare their current moral self in relation to the person they want to be. Our original manipulation was designed to capture these fluctuations in perceived moral standing.

However, we recognize that terms such as “excessive” and “surplus” may have introduced evaluative meaning beyond the intended manipulation of moral self-image. Specifically, this wording may have led participants to interpret high moral self-image as reflecting “too much” moral standing, or low moral self-image as reflecting a deficiency requiring repair. Consequently, the manipulation may have influenced not only participants' perceived level of moral self-image but also their perception that this level was somehow inappropriate and in need of correction. To examine this possibility, we conducted a supplementary validation study comparing the original manipulation with a revised version that removed this evaluative language and instead manipulated moral self-image more directly. All materials, data, and syntax files for this constructive replication of Study 2C are available on our OSF repository.⁶

Participants and Procedure. We recruited participants through Prolific to complete an online between-subjects experiment. Of the 83 individuals who participated, 67 ($M_{\text{age}} = 37.34$; $SD = 10.01$; 62.7% Male; 43.3% White) met our preregistered selection criteria and were retained for analysis (<https://aspredicted.org/4fz48i.pdf>). On average, participants reported 7.01 years of managerial experience ($SD = 6.31$) across a variety of industries.

Participants were randomly assigned to one of four conditions in a 2 (Moral Self-Image: High vs. Low) $\times$ 2 (Manipulation Version: Original vs. Revised) design. As in our other studies, participants

⁶ https://osf.io/fc78h/overview?view_only=f01b6eae074d493dbaec7c762c08e3cf

assumed the role of CEO and reflected on how they would feel about their prior decisions as CEO. Participants assigned to the original-manipulation conditions received the same moral self-image manipulation used in Study 2C. Participants assigned to the revised-manipulation conditions received a modified version that removed references to an “excessively” high moral self-image, a “moral surplus,” or a “moral deficit.” Instead, the revised manipulation described participants as experiencing either a stronger and more positive sense of their moral self or a weaker and more negative sense of their moral self at that moment. We assessed the effectiveness of both manipulations using the same three-item moral self-image measure employed in Studies 2B and 2C.

Results. Means and standard deviations for each condition are presented in Table G1. Levene’s test for equality of variances was significant, $F(3, 63) = 15.01, p < .001$, indicating that the assumption of homogeneity of variance was violated. Given approximately equal group sizes, the ANOVA is robust to this violation. Results from a 2 (Moral self-image: High vs. Low) $\times$ 2 (Manipulation version: Original vs. Revised) between-subjects ANOVA revealed a significant main effect of moral self-image level on the manipulation check, $F(1, 63) = 54.29, p < .001, \eta^2 = .46$. Participants in the high moral self-image condition ($M = 6.59, SD = .51$) reported significantly higher moral self-image than those in the low moral self-image condition ($M = 4.14, SD = 1.89$).

Results revealed no significant main effect of the manipulation version, $F(1, 63) = 2.01, p = .161, \eta^2 = .031$. Participants exposed to the original manipulation ($M = 5.21, SD = 1.81$) did not differ significantly from those exposed to the new manipulation ($M = 5.57, SD = 1.88$). In addition, the interaction between moral self-image level and the manipulation version was not significant, $F(1, 63) = .951, p = .333, \eta^2 = .015$.

Pairwise comparisons showed no difference in the high condition between the original ($M = 6.52, SE = .32$) and the revised version ($M = 6.67, SE = .34, p = .753$). Likewise, the low condition showed no significant differences between the original ($M = 3.73, SE = .34$) and revised versions ($M = 4.53, SE = .33, p = .098$).

Taken together, these findings indicate that the revised manipulation successfully generated

meaningful differences in moral self-image without relying on language referring to a moral “surplus” or “deficit.” This suggests that the manipulation’s effectiveness does not depend on framing participants’ moral self-image as excessive, deficient, or otherwise in need of correction.

Alternative Dependent Variable Operationalization

In addition to validating the revised moral self-image manipulation, we conducted a constructive replication of Study 2C using the revised manipulation materials and revised dependent-variable measures. This replication was designed to provide a more rigorous test of whether moral self-image influences CSR and CSI decisions when high and low moral self-image are not framed as a moral “surplus” or “deficit,” and when the decision tasks are less likely to encourage a strategic-balancing interpretation.

Participants and Procedure. Using Prolific, we recruited 380 participants to complete the study. Prior to analysis, we excluded 93 participants who did not meet our preregistered selection criteria (<https://aspredicted.org/hz4329.pdf>), resulting in a final sample of 287 participants ($M_{\text{age}} = 40.66$; $SD = 11.11$; 67.6% Male; 75.3% White). Participants reported an average of 10.15 years of managerial experience ($SD = 7.80$) across a variety of industries. Participants were randomly assigned to either a high or low moral self-image condition using the revised manipulation.

As in the other studies, participants assumed the role of CEO of a global technology company and reflected on their prior decisions as CEO. After completing the manipulation check, participants responded to a charitable donation decision task (CSR) and a supplier decision task (CSI), both of which excluded references to balancing competing strategic priorities. We then examined the effects of moral self-image on participants’ CSR and CSI decisions.

Attention Check Analyses. We included one attention-check item and one self-reported effort item. Of the 287 participants, one failed the attention check, and one reported exerting little effort while completing the study. Analyses conducted with and without these participants produced substantively identical results. Consistent with our reporting approach throughout the manuscript, we therefore present results based on the full sample.

Main Analyses. Table G2 reports the means and standard deviations of key variables by condition, and Table G3 provides the corresponding descriptive statistics and correlations of key variables. Levene's test for equality of variances showed that while the homogeneity assumption is not met for the manipulation check, $F(1,285) = 82.39, p < .001$, and CSR, $F(1,285) = 4.33, p = .038$, the assumption was met for CSI, $F(1,285) = 2.65, p = .104$.

The revised manipulation was successful. Participants in the high moral self-image condition reported significantly higher moral self-image ($M = 6.52, SD = 0.59$) than those in the low moral self-image condition ($M = 3.06, SD = 1.59$), $F(1, 285) = 575.89, p < .001$.

Supporting our hypothesis, participants in the high moral self-image condition made less allocation to CSR ($M = 5.60, SD = 1.60$) than participants in the low moral self-image condition ($M = 6.08, SD = 1.99$), $F(1, 285) = 4.97, p = .027$. This pattern is consistent with the proposed moral self-regulation account: when participants experienced a stronger moral self-image, they subsequently allocated less to CSR than when they experienced a weaker moral self-image. Consistent with the results of Study 2C, however, we did not find moral licensing for CSI; instead, participants in the high moral self-image condition reported lower levels of CSI ($M = 32.57, SD = 24.23$) than participants in the low moral self-image condition ($M = 42.65, SD = 26.31$), $F(1, 285) = 11.34, p < .001$, showing moral consistency rather than moral balancing. These findings further highlight the inconsistent operation of moral self-regulation across CSR and CSI decisions.

Discussion. This supplementary set of experiments provides additional evidence consistent with our core findings while directly addressing concerns regarding the original Study 2C materials. First, the validation study demonstrated that the revised manipulation successfully produced meaningful differences in moral self-image even after removing language referring to a moral "surplus" or "deficit." This finding suggests that the manipulation can effectively induce relatively high versus low moral self-image without explicitly framing participants' moral standing as excessive, deficient, or in need of correction. Second, the constructive replication showed that moral self-image continued to influence CSR allocations when participants received the revised manipulation and responded to the revised CSR decision task designed to

reduce the salience of competing strategic priorities.

Taken together, the results were broadly consistent with the CSR findings from Study 2C. Specifically, participants with relatively high moral self-image allocated fewer resources to CSR than participants with relatively low moral self-image, consistent with the proposed moral self-regulation mechanism. In contrast, the CSI results did not provide support for the predicted moral licensing effect. Whereas the original study yielded no evidence of moral self-regulatory effects on CSI, the replication study revealed a pattern of moral consistency rather than moral compensation. We interpret these supplemental findings as reinforcing a broader conclusion of the main manuscript: moral self-regulation is more consistently reflected in CSR decisions than in CSI decisions. These findings suggest that moral self-regulation may not operate symmetrically across all socially relevant organizational choices, highlighting an important boundary condition and a promising avenue for future research.

Table G1. Means and Standard Deviations for MSI Manipulation Check by Condition

Moral Self-Image Condition	Version of Manipulation	<i>M</i>	<i>SD</i>	<i>n</i>
High	Original	6.52	.54	18
High	Revised	6.67	.47	16
Low	Original	3.73	1.57	16
Low	Revised	4.53	2.13	17

Table G2. Means and Standard Deviations for Dependent Variables by Condition

	<i>N</i>	CSR	CSI
High moral self-image	138	5.60 (1.60)	32.57 (24.23)
Low moral self-image	149	6.08 (1.99)	42.65 (26.31)

Note. *N* = 287. Means are reported in each cell, with standard deviations reported in parentheses.

Table G3. Descriptive Statistics and Correlations of Key Variables

	<i>M</i>	<i>SD</i>	CSR	CSI
CSR	5.85	1.83	--	--
CSI	37.80	25.79	-.189***	--
Moral self-image	4.73	2.12	-.158***	-.163***

Note. *N* = 287. *M* = Mean. *SD* = Standard deviation. * $p < .1$; ** $p < .05$; *** $p < .01$