Under pressure: how widespread vs severe competitor unethical practices shape responsible artificial intelligence deployment

Samuel N. Kirshner and Jessica Lawson
School of Information Systems and Technology Management, UNSW Sydney, Sydney, Australia

Received 15 January 2025
Revised 14 June 2025
16 August 2025
4 October 2025
4 November 2025
8 December 2025
Accepted 12 December 2025

Abstract

Purpose – This study explores how competitive pressure and organizational goals influence responsible AI (RAI) decisions when introducing AI-based digital services. We examine how two types of unethical competition, horizontal (numerous competitors using similar unethical AI tactics) and vertical (a competitor using highly unethical AI), interact with regulatory-focus-driven objectives to shape RAI deployment.

Design/methodology/approach – We design three experimental studies featuring scenarios involving the launch of an unethical AI service to assess how competitive pressures and organizational goals affect RAI decisions. Each experiment manipulates horizontal and vertical unethical competition. Participants' regulatory focus is measured in Study 1 ($N = 249$) and is manipulated through organizational goals in Study 2 ($N = 304$) to assess their interactions with competitive pressure. Study 3 ($N = 159$) tests moral disengagement theory as the underlying mechanisms.

Findings – The results show that horizontal unethical competition increases the launch of unethical AI, regardless of regulatory focus. We uncover a novel interaction effect between vertical unethical competition and firm objectives. While the severity of a competitor's unethical behavior reduces RAI deployment directly, prevention-focused goals can counteract this effect under vertical unethical competition, promoting more responsible decisions.

Originality/value – This research advances RAI scholarship by introducing two unethical competitive contexts to analyze how competitive pressure and organizational goals shape decisions to launch unethical AI services. By isolating horizontal and vertical competition, we provide new insights into how external competition drives UPB. The findings provide a behavioral framework for understanding ethical trade-offs in AI deployment, linking high-level ethics to practical governance in competitive settings.

Keywords Responsible artificial intelligence, Moral disengagement, Regulatory focus, Competition, Unethical pro-organizational behavior

Paper type Research article

Introduction

The integration of artificial intelligence (AI) into digital products and services has created unprecedented opportunities and emerging challenges around ethics, privacy, security, human rights and fairness that managers must navigate to avoid unintended consequences of AI-integrated systems (Berente *et al*., 2021; Marjanovic *et al*., 2022). Responsible AI (RAI) practices can deliver value by strengthening fairness perceptions and lowering the incidence of discriminatory recommendations (Ebrahimi *et al*., 2025). A recent global McKinsey survey of over 750 leaders found that most plan to invest more than 1 million USD in RAI in 2025, with larger firms investing between 25 and 50 million (Luget *et al*., 2025). Yet, social injustices and harmful outcomes can occur when organizations implement AI systems without adequate governance and processes or prioritize economic gains over user welfare (Berente *et al*., 2021; Mikalef *et al*., 2022).

Internet Research
Vol. 36 No. 3, 2026
pp. 1073-1092
Emerald Publishing Limited
e-ISSN: 2054-5657
p-ISSN: 1066-2243
DOI 10.1108/INTR-01-2025-0114

Although demand-side dynamics driven by consumer expectations push organizations to adopt AI systems, supply-side competitive pressures intensify the need for rapid innovation, often sacrificing ethical considerations in the process (Alsheibani *et al.*, 2020; Enholm *et al.*, 2022). These pressures create launch decisions for AI-based digital services and products, where organizations must navigate trade-offs between achieving competitive advantage and ensuring ethical responsibility (Benbya *et al.*, 2021; Kirshner and Lawson, 2025; Lee *et al.*, 2018; Papagiannidis *et al.*, 2025; Wu *et al.*, 2023). Competitive pressures force organizations into rapid design-to-market strategies that bypass thorough testing, creating AI systems that pass internal validation but fail under real-world conditions (Marabelli *et al.*, 2021; Rhue and Washington, 2020). This testing gap significantly increases risks of privacy violations, bias and manipulative behaviors, while AI's inherent opacity obscures these failures from detection (Wu *et al.*, 2023; Kordzadeh and Ghasemaghaei, 2021; Hinds *et al.*, 2020; Asatiani *et al.*, 2021).

While competitive pressures can increase unethical pro-organizational behavior (UPB), actions that violate ethical norms but benefit the organization (Moore, 2008; Welsh *et al.*, 2015), limited research explores how external competitive environments influence decisions to launch unethical AI. Ethical guidelines provide high-level principles like transparency and fairness (Jobin *et al.*, 2019) but fail to account for competitive trade-offs (Birkstedt *et al.*, 2023), such as prioritizing speed-to-market over thorough testing or accuracy over explainability (Liang *et al.*, 2021; Langer *et al.*, 2021). We seek to understand how competitive pressure influences decisions to launch unethical AI by examining two distinct forms of competition. First, firms may encounter horizontal competitive pressure, where many competitors deploy unethical AI systems, creating industry norms of lower ethical standards. Second, organizations may face vertical competitive pressure, where a key competitor deploys an extremely unethical AI system, establishing a new unethical floor. We posit that these competitive conditions can activate moral disengagement (Bandura, 1991). Horizontal unethical AI pressure from multiple competitors can dilute individual accountability, a process known as diffusion of responsibility (Bandura, 1991; Moore, 2008). Vertical unethical AI pressure can trigger advantageous comparison, where actions are justified by contrasting them with more egregious violations (Bandura, 1991).

As organizational goals can shape responses to external competitive forces, we employ regulatory focus theory (Higgins, 1998). Promotion-focused organizations, driven by growth and achievement, are more prone to trade-offs that compromise RAI (Kirshner and Lawson, 2025) and may use advantageous comparison with competitors' more severe violations to justify launching unethical AI. In contrast, prevention-focused firms, motivated by security and avoiding losses, may succumb to horizontal unethical pressure through diffusion of responsibility when industry practices threaten their competitive position. This suggests that firm goals may moderate how different competitive conditions impact AI launch decisions. Thus, we investigate the following two research questions:

RQ1. How do horizontal and vertical unethical competitive pressure impact decisions to launch unethical AI systems?

RQ2. How do organizational goals impact the relationship between competitive pressure and the decision to launch unethical AI?

We conducted three experiments to examine how competitive pressures influence decisions to launch unethical AI. In all three studies, we manipulated competitive conditions to test the effect of horizontal and vertical pressure. Our results reveal that horizontal competitive pressure consistently increases unethical AI launches. Surprisingly, vertical competitive pressure interacted with prevention-focused goals, decreasing the likelihood of launching unethical AI compared to promotion-focused goals. We discuss how these findings respond to calls for exploring the dynamic nature of ethical AI implementation (Birkstedt *et al.*, 2023; Li and Wang, 2023) and bridge the gap between high-level ethical principles and practical governance mechanisms (Morley *et al.*, 2020), demonstrating how competitive environments and organizational orientations can shape RAI deployment.

Background theory

Moral disengagement theory

To investigate the role of competitive environments on ethical AI decision making, we leverage moral disengagement theory. Moral disengagement is a context-dependent, malleable phenomenon shaped by social influences and organizational environments (Moore, 2008). Moral disengagement theory identifies eight mechanisms for rationalizing immoral behavior (Bandura *et al.*, 1996), including behavioral activities (moral justification, euphemistic labelling and advantageous comparison), agency factors (diffusion of responsibility and displacement of responsibility), outcomes (distortion of consequences) and victim characteristics (attribution of blame and dehumanization). As downplaying harm is a function of moral disengagement, firms can take advantage of moral disengagement to implement unethical processes, pre-empting employee's cognitive dissonance where immoral behaviors may contradict their values (Moore, 2008; Bandura, 1991).

Regulatory focus theory

Individual decision-making is shaped by motivations and context, influencing goal setting and actions, where both failure and success drive motivation (Van-Dijk and Kluger, 2004). Regulatory Focus Theory reconciles these views, distinguishing between promotion-focused individuals, driven by success and prevention-focused individuals, motivated by avoiding failure (Higgins, 1998). Promotion focus is associated with risk-taking, creativity and innovation, often incentivized by rewards such as bonuses (Johnson *et al.*, 2017). In contrast, prevention focus prioritizes stability and security, aiming to minimize losses and maintain the status quo.

Regulatory focus can affect decision-making through its relationship with risk (Crowe and Higgins, 1997; Gino and Margolis, 2011). Promotion-focused individuals pursue ambitious goals, fostering exploration, growth and innovation, but also by engaging in riskier behaviors (Higgins, 2002). Prevention-focused individuals adopt cautious strategies, reducing risk and prioritizing ethical behavior through vigilance (Higgins, 2002; Gino and Margolis, 2011). At the firm level, regulatory focus affects employee behaviors and ethical outcomes, as individual tendencies, such as extraversion and neuroticism, can translate into professional orientations (Lanaj *et al.*, 2012). A match between employee focus and organizational environment enhances engagement and moral reactions (Achar and Lee, 2022). Industry competition, leadership and organizational culture also shape focus, with promotion focus driving innovation and citizenship behaviors, while prevention focus supports safety and compliance (Lanaj *et al.*, 2012). Recently, Regulatory Focus has found extensive application in IS, offering valuable insights into user behavior and digital platform engagement (e.g. Lwin *et al.*, 2016; Lee *et al.*, 2019; Wu *et al.*, 2019; Zhou *et al.*, 2024). Most recently, Kirshner and Lawson (2025) found support that promotion-focused behaviors can lead to increased AI-based UPB, by influencing decision-maker's RAI values. The current work extends their findings by integrating regulatory focus with moral disengagement to examine how competitive pressures interact with organizational goals to impact AI-based UPB decisions.

Hypothesis development

Organizations deploying AI systems face competitive pressures that can affect ethical decision-making. We examine how two ethical competitive configurations activate distinct moral disengagement mechanisms to influence decisions to launch potentially unethical AI systems. First, when many competitors adopt similar unethical AI practices, we posit that this widespread adoption activates diffusion of responsibility. Decision-makers may experience reduced personal obligation to maintain ethical standards as the behavior becomes normalized

(Bandura, 1991). Second, when a focal competitor implements extremely unethical AI practices, we posit that this scenario will activate advantageous comparison (Bandura, 1991), where individuals justify deploying moderately unethical AI by contrasting it against more severe misconduct.

We propose that competition activating these moral disengagement mechanisms predicts how external environments influence AI deployment decisions. Under horizontal pressure, decision-makers may rationalize deployment, since "everyone else is doing it" and delaying the launch while competitors gain market advantages would be strategically detrimental. Similarly, under vertical unethical pressure, decision-makers may launch a moderately unethical AI as a "lesser evil" while maintaining competitive positioning and appearing more ethical than their focal competitor. Based on these mechanisms, we hypothesize that different competitive environments will systematically affect AI deployment decisions as follows:

H1a. [Horizontal]: A higher number of competitors with unethical AI offerings increases the propensity for launching potentially unethical AI.

H1b. [Vertical]: A higher severity of unethical AI from a focal competitor increases the propensity for launching potentially unethical AI.

Kirshner and Lawson (2025) explored the relevance of promotion and prevention-based organizational goals for AI-based UPB, and found that promotion-focused goals lead to lower levels of AI-based UPB. We expand on Kirshner and Lawson (2025) by positing a relationship between the competitive environments and regulatory focus on AI deployment decisions. Prevention-focused goals and competitive threats: Prevention-focused individuals tend to be risk-avoiders, focusing on minimizing losses through security and stability, and are conservative in their efforts to avoid adverse outcomes (Higgins, 2002). When many competitors deploy unethical AI systems, this creates a direct competitive threat to the organization's market position. Prevention-focused decision-makers may perceive this widespread adoption as creating a competitive disadvantage that threatens organizational survival and their own job security, given their motivation from resilience and the fear of failure, prioritizing the security of employment and maintaining the status quo (Johnson *et al*., 2017; Van-Dijk and Kluger, 2004). This fear of falling behind competitively may activate diffusion of responsibility, allowing them to justify deploying similar unethical AI by reasoning that "everyone else is doing it" and that maintaining competitive parity is necessary to avoid organizational losses. The industry-wide practice becomes normalized, reducing individual accountability for the ethical implications.

H2a. A higher degree of prevention-focused goals compared to promotion-focused goals increases the impact of a higher number of unethical competitors on the propensity for launching potentially unethical AI.

Promotion-focused individuals approach tasks eagerly, centering on achievements and positive outcomes (Higgins *et al*., 2003), and have greater risk tendencies, focusing on maximizing success by innovative and creative means, leading to unethical behavior (Gino and Margolis, 2011). Accordingly, we propose that promotion-focused goals provide an opportunity for individuals to take advantage of a primary competitor implementing an AI system with worse consequences to engage in advantageous comparison and justify an unethical AI action. Put differently, individuals may increase the likelihood of implementing moderately unethical AI when competitors implement even more unethical AI. Therefore, a scenario where an organization can implement a new AI and competitors are implementing worse AI may create a natural regulatory fit for a promotion-focused individual to leverage advantageous comparison, leading to AI-based UPB.

H2b. A higher degree of promotion-focused goals compared to prevention-focused goals increases the impact of the severity of unethicalness of competitors on the propensity for launching potentially unethical AI.

The complete research model and hypotheses are presented in Figure 1.

Study 1

Overview

To explore how external competitive pressures impact AI-based UPB, Study 1 examined the relationship between individuals' regulatory focus orientation and behaviors in a competitive scenario characterized either by a high number of unethical competitors (Horizontal) or by a severely unethical competitor (Vertical). Utilizing a between-subjects design with three treatments (Control vs Horizontal vs Vertical), we manipulated the potential for moral disengagement to test H1a and H1b and measured participants underlying regulatory focus to explore potential moderating effects of regulatory focus, testing H2a and H2b.

Participants

We recruited American MTurkers, compensating them with 2 USD for completing the study. We screened [1] participants and used attention checks to ensure quality engagement from participants. The sample size was 240 participants, which is above GPower's recommendation (d = 0.5, alpha = 0.05, and 1-beta = 0.8) due to measuring regulatory focus. In total, 249 participants (females: 111; other: 2; age: M = 44.04, SD = 11.03) completed the experiment.

Procedure

Participants indicated how strongly they agreed with eight promotion and prevention-focused statements adapted from Lockwood *et al*. (2002) to measure regulatory focus. We discussed that firms use AI to undertake targeted marketing and personalized pricing by tracking consumer activity and spending habits from previous interactions and third-party data. Participants were asked to imagine they worked for a firm that sells tourism services and reported directly to the Chief Marketing Officer (CMO), who planned to introduce AI into the organization. This scenario and attention check questions were taken from Kirshner and Lawson (2025) [2]. Participants answered two comprehension-check questions, and if they passed, they were asked to decide on a proposed AI implementation, addressing the question: "Should the firm use AI to target marketing and alter pricing of services based on data collected from their consumers based on previous interactions and data about their consumers purchased from third parties?".

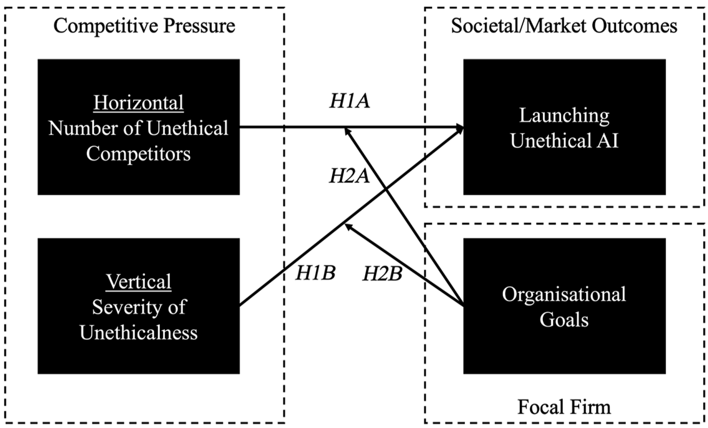

Figure 1. Research model for the impact of competition and goals on launching unethical AI. Source: Authors' own work

Participants were then allocated to one of the three conditions: Control, Horizontal or Vertical. All participants read the Control condition, reiterating that AI can track consumer activity through cookies and that the AI analyzes the data from previous interactions and data purchased from third parties for target marketing to improve profits, but it may have potential ethical issues. Participants in the Horizontal treatment were told (in addition to the Control group information) that 95% of their firm's competitors use AI to target marketing and alter pricing using data from previous interactions and purchases from third parties. We noted that this was the same strategy as the CMO's proposal (see Table A1). Participants in the Vertical treatment were told that the firm's main competitor used AI for target marketing and altering pricing using data from previous interactions and data purchased from third parties, which is the same strategy as the CMO's proposal. (This is the same as in the Horizontal treatment, but with one firm.) In addition, the competitor accessed microphones on consumer devices to collect data and sold this data without consent from consumers, which we emphasized is highly unethical.

After exposure to the treatment statements, participants indicated whether to implement the AI on a 7-point Likert scale (1 = "Strongly disagree" to 7 = "Strongly agree"). This was in accordance with the scale as used by Lee *et al*. (2018) to avoid participant fatigue. Participants in the moral disengagement treatments were asked a manipulation check question. We asked what percentage of the firm's competitors disregard consumer privacy to undertake targeted marketing and personalized pricing in the Horizontal group, which was measured on a 5-point scale from "Much lower than average" to "Much higher than average." The Vertical group was asked how they would rate the ethicalness of the main competitor compared to the CMO's proposed action on a 5-point scale from "Much lower" to "Much higher." Participants then reported age and gender.

Results

Participants in the Horizontal group perceived the number of the firm's competitors using the same AI as higher than average ($M = 4.60$, $SD = 0.92$, $t(80) = 15.74$, 95% CI: [4.40, 4.81], $p < 0.001$), and participants in the Vertical group perceived the ethicalness of the main competitor compared to the CMO's proposal as being lower ($M = 1.94$, $SD = 1.05$, $t(83) = -9.29$, 95% CI: [1.71, 2.17], $p < 0.001$). Therefore, both competition manipulations were successful in creating moral disengagement scenarios. The responses were internally consistent across the four promotion-focused measures (Cronbach's alpha = 0.84) and the four prevention-focused measures (Cronbach alpha = 0.85). Overall, participants were slightly more promotion-focused ($M = 5.68$, $SD = 0.99$, t(248) = 26.75, 95% CI: [5.56, 5.81], $p < 0.001$) than prevention-focused ($M = 3.79$, $SD = 1.39$, $t(248) = -2.44$, 95% CI: [3.61, 3.96], $p = 0.016$). We deducted participants' prevention-focus scores from their promotion-focus scores to evaluate regulatory focus (Lockwood *et al*., 2002).

To examine how participants' competitive context (Vertical and Horizontal) and regulatory focus influenced RAI decision-making, we used Hayes' PROCESS Model 1 with indicator coding for the multi-categorical treatment variable (Hayes and Montoya, 2017). The dependent variable was participants' decision on whether the firm should implement the AI system, reverse-coded so that higher values indicated the more ethical choice. The moderator regulatory focus was mean-centered and conditional effects were estimated at low (-1 SD; 16th percentile), medium (0 SD, 50th percentile) and high (+1 SD; 84th percentile) values of regulatory focus.

Table 1 presents the estimates of the overall model, which was significant ($F(5, 243) = 3.97$, $p = 0.002$). At the mean regulatory focus level, participants in the Horizontal group were significantly less likely to make ethical decisions than those in the Control group ($b = -0.70$, $SE = 0.31$, $p = 0.023$), supporting H1a. The difference between the Vertical and Control conditions at the mean level of regulatory focus was not significant ($b = 0.42$, $SE = 0.30$, $p = 0.165$), providing no support for H1b. The analysis above addresses RQ1.

Table 1. PROCESS model 1 regression output for study 1

	b	SE	*t*	*p*	95% CI
Intercept	3.51	0.21	16.42	<0.001	[3.09, 3.93]
Vertical	0.42	0.30	1.39	0.165	[−0.17, 1.02]
Horizontal	−0.70	0.31	−2.28	0.023	[−1.30, −0.10]
Regulatory Focus	−0.02	0.11	−0.15	0.882	[−0.23, 0.20]
Vertical × Reg Focus	−0.30	0.17	−1.82	0.070	[−0.63, 0.03]
Horizontal × Reg Focus	0.04	0.15	0.29	0.772	[−0.25, 0.34]

Source(s): Authors' own work

We next tested whether regulatory focus moderated the effects of treatment condition. The interaction between regulatory focus and Horizontal was not significant ($b = 0.04$, SE $= 0.15$, $p = 0.772$), providing no support for H2a. However, the interaction between regulatory focus and Vertical was marginally significant ($b = -0.30$, SE $= 0.17$, $p = 0.070$). Probing this interaction revealed that participants low in regulatory focus (-1 SD) were significantly more likely to make ethical decisions in the Vertical condition compared to the Control group ($b = 1.00$, SE $= 0.44$, $p = 0.024$). At high levels of regulatory focus ($+1$ SD), there was no significant difference between Vertical and Control ($b = -0.22$, SE $= 0.46$, $p = 0.640$). These results support H2b and suggest that participants lower in promotion focus were more sensitive to a competitor's severely unethical behavior and responded by making more responsible AI recommendations. The analysis above addresses RQ2.

We then examined if regulatory focus moderated the effects of competitive context on launching the AI. The interaction between regulatory focus and Horizontal was not significant ($b = 0.04$, $SE = 0.15$, $p = 0.772$), providing no support for H2a. However, the interaction between regulatory focus and Vertical was marginally significant ($b = -0.30$, $SE = 0.17$, $p = 0.070$). Contrary to the direction predicted by H2b, probing this interaction revealed that participants low in regulatory focus were significantly more likely to make ethical decisions in the Vertical condition compared to the Control group ($b = 1.00$, $SE = 0.44$, $p = 0.024$, 95% CI [0.13, 1.86]; see Table 2). This effect was not significant at the mean level of regulatory focus ($b = 0.32$, $SE = 0.31$, $p = 0.307$), and reversed but remained nonsignificant at high levels of regulatory focus ($b = -0.22$, $SE = 0.46$, $p = 0.640$). These findings suggest that participants lower in promotion focus (i.e. more prevention-focused) responded more strongly to the presence of a severely unethical competitor by making more responsible AI decisions.

These observations are supported by Figure 2, which displays the predicted ethical decision scores across the range of regulatory focus levels for each treatment group. Participants in the Vertical condition made the most ethical decisions when their regulatory focus was low, but this effect diminished as regulatory focus increased. In contrast, participants in the Horizontal condition consistently made less ethical decisions regardless of regulatory orientation.

Table 2. Conditional effects of treatment by regulatory focus levels

Regulatory focus	Contrast with control	*b*	SE	t	*p*	95% CI
−1.90	Vertical	1.00	0.44	2.27	0.024	[0.13, 1.86]
	Horizontal	−0.78	0.41	−1.88	0.061	[−1.59, 0.04]
0.35	Vertical	0.32	0.31	1.02	0.307	[−0.29, 0.92]
	Horizontal	−0.68	0.31	−2.20	0.029	[−1.29, −0.07]
2.10	Vertical	−0.22	0.46	−0.47	0.640	[−1.13, 0.69]
	Horizontal	−0.61	0.44	−1.37	0.171	[−1.47, 0.26]

Source(s): Authors' own work

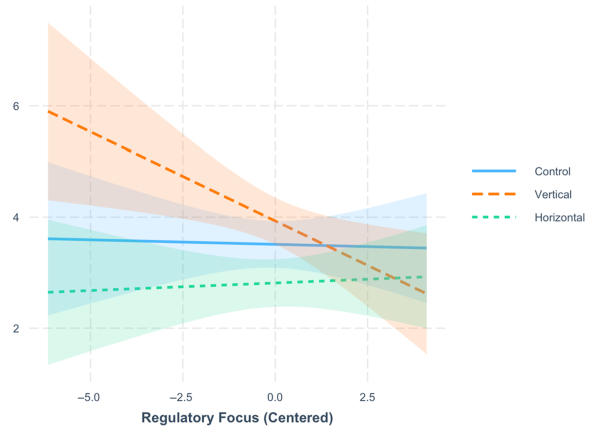

Figure 2. Predicted AI-based UPB scores (*y*-axis) as a function of regulatory focus (*x*-axis) and treatment condition (group) in Study 1. Source: Authors' own work

Overall, these results provide support for H1a, showing that widespread unethical competition undermines ethical decision-making, while the pattern of interaction observed for H2b runs counter to expectations, indicating that prevention-focused individuals were more responsibly responsive to the ethical severity of a focal competitor.

Study 2

Overview

Study 2 tested the robustness of our findings in Study 1. Moreover, we explicitly tested the interaction effects by manipulating the regulatory focus goal of an imagined employee and organization. Thus, the study is a 2 (regulatory focus: promotion vs prevention-focused goals) x 3 (competitive environment: Control vs Horizontal vs Vertical) between-subjects design to test the effects of competition on RAI launch decisions.

Participants

We recruited American MTurkers, incentivizing them with 1.85 USD for completing the study. As in the previous studies, we screened participants using Qualtrics' fraud direction services and incorporated attention-check questions. We aimed to recruit 50 participants per treatment as the sample size based on GPower's recommendation for an effect size of $d = 0.5$, alpha $= 0.05$, and 1-beta $= 0.8$. In total, 304 participants (females: 118; other: 3; age: $M = 39.13$, $SD = 11.16$) completed the study.

Procedure

The procedure resembled Study 1 except that we introduced a character named "Alex", who works for a company called MarkTech and reports to the CMO [3]. The survey began with an explanation of black box AI, like Study 1, with the same checks. Participants were randomly assigned to a promotion or prevention-focused manipulation that gave descriptions of Alex's goal orientation as being either a promotion-focused or prevention-focused employee (see Table A2). We asked a manipulation check for whether the participant believed "Alex is someone who is primarily striving to become their ought or ideal self?" Participants were then given a description of MarkTech as either a promotion-focused or prevention-focused organization and were asked a similar manipulation check question (i.e. "Do you think MarkTech primarily strives on achieving gains or preventing losses?").

We utilized the competitive environment manipulations from Study 1 for the Control, Horizontal, and Vertical groups. The same AI-based UPB question was asked to participants, but regarding what Alex should recommend: "Should Alex recommend that MarkTech use black box AI for targeted marketing and personalized pricing of products and services based on data collected from their consumers from previous interactions and data about their consumers purchased from third parties to increase profits and build technological capabilities?". Finally, we asked participants for demographic information.

Results

We first tested the success of the regulatory focus manipulations. Participants rated Alex as more focused on achieving ideal goals ($M = 4.45$, $SD = 2.04$) in the promotion-focused group, which was significantly higher than the scale midpoint of 4, $t(158) = 2.80$, $p = 0.006$, 95% CI [4.13, 4.77]. In contrast, participants in the prevention-focused condition rated Alex as more focused on avoiding undesired outcomes ($M = 3.03$, $SD = 1.93$), significantly below the midpoint, $t(144) = -6.02$, $p < 0.001$, 95% CI [2.72, 3.35]. Participants also rated MarkTech as more promotion-focused in the promotion condition ($M = 6.02$, $SD = 0.88$), $t(158) = 28.86$, $p < 0.001$, 95% CI [5.88, 6.16], and more prevention-focused in the prevention condition ($M = 2.61$, $SD = 1.88$), $t(144) = -8.94$, $p < 0.001$, 95% CI [2.30, 2.91]. These results confirm the success of the regulatory focus manipulations.

We first conducted a 2×3 factorial ANOVA to examine the main and interaction effects. The dependent variable was participants' recommendation for whether Alex should advise MarkTech to implement the AI system, reverse-coded so that higher values indicated more ethical (i.e. more cautious and responsible) choices. The ANOVA revealed a significant main effect of regulatory focus, $F(1, 298) = 33.47$, $p < 0.001$, a significant main effect of competitive condition, $F(2, 298) = 6.42$, $p = 0.002$ and a significant interaction between regulatory focus and competitive condition, $F(2, 298) = 4.05$, $p = 0.018$.

To further examine the interaction and test our specific hypotheses, we used Hayes' PROCESS Model 1 (see Table 3). The moderator regulatory focus was coded as promotion-focused (0) versus prevention-focused (1), with conditional effects estimated for both regulatory focus orientations. The overall PROCESS model was significant, $F(5, 298) = 10.88$, $p < 0.001$. At the promotion-focused level of regulatory focus, participants in the Vertical condition were directionally, but not significantly less likely to make ethical decisions than those in the Control group ($b = -0.35$, $SE = 0.34$, $p = 0.305$), providing little support for H1b. However, the difference between Horizontal and Control at the prevention-focused level was moderately significant ($b = -0.61$, $SE = 0.32$, $p = 0.057$), providing support [4] for H1a.

We next tested whether regulatory focus moderated the effects of the treatment. Table 3 shows that the interaction between regulatory focus and the Horizontal group was not significant ($b = 0.06$, $SE = 0.47$, $p = 0.894$), while the interaction between regulatory focus and Vertical was ($b = 1.20$, $SE = 0.47$, $p = 0.012$). Probing this interaction revealed that

Table 3. PROCESS model 1 regression output for study 2

Variable	*b*	SE	*t*	*p*	95% CI
Intercept	3.13	0.23	13.62	<0.001	[2.68, 3.58]
Vertical	−0.35	0.34	−1.03	0.305	[−1.01, 0.32]
Horizontal	−0.61	0.32	−1.92	0.057	[−1.23, 0.02]
Prevention Focus	−0.64	0.33	−1.92	0.057	[−0.02, 1.30]
Vertical × Prevention	1.20	0.47	−2.53	0.012	[0.27, 2.13]
Horizontal × Prevention	0.06	0.47	0.13	0.894	[−0.87, 0.99]

Source(s): Authors' own work

participants in the prevention-focused frame were significantly more likely to make ethical decisions in the Vertical condition compared to the Control group (b = 0.85, SE = 0.33, p = 0.011, 95% CI [0.20, 1.51]; see Table 4), but the effect was not significant for promotion-focused participants (b = −0.34, SE = 0.34, p = 0.305).

Thus, overall, the pattern of results in Study 2 largely replicates and extends the findings from Study 1. Prevention-focused participants demonstrated greater ethical sensitivity when faced with a severely unethical competitor, making significantly more responsible AI decisions in the Vertical condition compared to the Control. In contrast, promotion-focused participants showed no significant difference in ethical decision-making across conditions. These findings provide additional support that regulatory focus can moderate responses to competitive ethical contexts. These results are presented in Figure 3.

Study 3

Overview

Study 3 tested whether the Horizontal and Vertical competition treatments used in Studies 1 and 2 activated moral disengagement. Specifically, we assessed whether exposure to a high number of unethical competitors and whether exposure to a severely unethical competitor increased perceptions of unethicalness. This study helps validate the theoretical interpretation that these treatments encourage unethical decisions by enabling moral disengagement. Thus, this study tests the following hypothesis:

Table 4. Conditional effects of treatment at levels of regulatory focus for study 2

Regulatory Focus	Contrast with control	*b*	SE	*t*	*p*	95% CI
Prevention	Vertical	0.85	0.33	2.55	0.011	[0.20, 1.51]
	Horizontal	−0.54	0.35	−1.56	0.121	[−1.23, 0.14]
Promotion	Vertical	−0.34	0.34	−1.03	0.305	[−1.01, 0.32]
	Horizontal	−0.61	0.32	−1.91	0.057	[−1.23, 0.02]

Source(s): Authors' own work

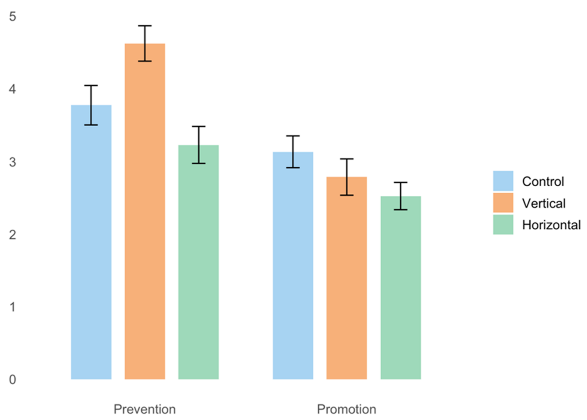

Figure 3. AI-based UPB (y-axis) as a function of regulatory focus (x-axis) and competitive context (group) in Study 2. Source: Authors' own work

H3. Competitive pressure reduces perceived moral accountability, mediated by moral disengagement, with horizontal unethical competition activating diffusion of responsibility and vertical unethical competition activating advantageous comparison.

Participants

We recruited 160 British participants through Prolific and paid each participant 1.50 GBP, based on a target sample size of approximately 50 participants per treatment [5]. Participants were 18 years or older and fluent in English. Attention checks were embedded throughout the survey to ensure data quality and one participant who completed the study was dropped because Qualtrics flagged them as a duplicate, leaving a sample of 159 (females: 69; other: 1; age: $M = 43.29$, $SD = 13.58$).

Procedure

Participants read the same AI implementation scenario from Study 2, in which MarkTech was considering adopting a black box AI system for targeted marketing and personalized pricing. The system was described as potentially increasing profits, but raising ethical concerns due to its use of consumer data. As in Studies 1 and 2, participants were randomly assigned to one of three treatment conditions. In the Control condition, no information was provided about competitors. In the Horizontal condition, participants read that 90% of competing firms had already implemented similar AI systems. In the Vertical condition, participants read that the focal competitor had implemented the AI and additionally accessed consumer device microphones to collect sensitive data. After reading the materials, participants were told that Alex had recommended implementing the AI. They then evaluated this recommendation using three items on 7-point scales assessing ethicality, justice and moral acceptability (following Reidenbach and Robin, 1990).

Participants then completed nine accountability items designed to assess whether the treatment conditions influenced how responsible Alex was perceived to be for potential ethical consequences. Each participant responded to three items per condition (Control, Horizontal, Vertical), presented in a fixed order. The Control items assessed baseline moral accountability (e.g. “If no other companies were using this type of AI, then Alex should feel accountable for any ethical issues that arise”). The Horizontal items reflected diffusion of responsibility (e.g. “If many competitors were already using this AI, Alex should feel personally responsible for any unfair outcomes”), and the Vertical items reflected advantageous comparison (e.g. “If other firms are doing worse things than using this AI, then Alex should feel accountable for any ethical issues that arise”). Responses were made on 7-point Likert scales (1 = “Strongly disagree” to 7 = “Strongly agree”). Lower agreement with these accountability statements in the Horizontal and Vertical conditions indicated reduced perceived responsibility, consistent with disengagement reasoning. While this study included both ethical evaluation and accountability measures, its primary aim was to test whether the competition manipulations reduced perceived ethical responsibility through the mechanisms theorized to underpin moral disengagement.

Results

We first examined the internal consistency of the three ethical evaluation items (ethicality, justice and moral acceptability). Responses were highly consistent (Cronbach alpha = 0.93), and we averaged them to form a composite ethical decision score. We then conducted an ANOVA analysis with the treatment conditions (Control, Horizontal, Vertical) as the independent variable. The effect of treatment was significant, $F(2, 156) = 4.68$, $p = 0.011$, $n^2 = 0.06$. Tukey’s HSD revealed that participants in the Horizontal condition evaluated Alex’s recommendation as significantly less ethical than participants in the Control condition (Difference = 0.86, 95% CI [0.19, 1.52], $p = 0.007$). The difference between the Vertical and

Control condition was not significant (Difference = 0.45, 95% CI [−0.21, 1.11], $p = 0.241$), nor was the difference between the Vertical and Horizontal conditions (Difference = −0.40, 95% CI [−1.06, 0.26], $p = 0.323$).

We then assessed whether the Horizontal and Vertical manipulations activated moral disengagement mechanisms. Participants completed nine moral disengagement items (three per condition), where reliability was high for the Control (Cronbach alpha = 0.85), Horizontal (Cronbach alpha = 0.81) and Vertical items (Cronbach alpha = 0.79). Paired-samples *t*-tests revealed that participants expressed less moral accountability in the Horizontal condition than in the Control condition, $t(158) = -9.55$, $p < 0.001$, 95% CI [−1.00, −0.66], Difference = −0.83. They also expressed less accountability in the Vertical condition than in the Control, $t(158) = -10.24$, $p < 0.001$, 95% CI [−1.15, −0.78], Difference = −0.96. Figure 4 shows the results for both the between-subject comparisons on ethical evaluations and the within-subject responses to accountability statements reflecting the moral disengagement logic. Although participants in the Vertical condition did not judge Alex's recommendation as significantly more or less ethical than in the Control condition, they nevertheless held Alex significantly less personally accountable for potential ethical consequences. This is consistent with the logic of advantageous comparison, where unethical actions are less problematic when compared to worse alternatives. Similarly, participants in the Horizontal condition both rated the decision as significantly less ethical and reported lower personal accountability, consistent with diffusion of responsibility.

To formally evaluate H3, we tested if the effect of treatment condition on ethicality was mediated by participants' perceived responsibility. Using PROCESS Model 4 with 5,000 bootstrap samples (see Table 5), the treatment was a multi-categorical independent variable, with the mediator matched to the moral accountability associated with the treatment condition, and the ethicality of the decision being the dependent variable. Relative to the control condition, the Horizontal treatment significantly reduced perceived responsibility ($b = -1.01$, $SE = 0.26$, $p < 0.001$, 95% CI [−1.51, −0.50]), as did the Vertical treatment ($b = -0.80$, $SE = 0.26$, $p = 0.002$, 95% CI [−1.30, −0.30]). In turn, perceived responsibility negatively predicted ethicality ($b = -0.32$, $SE = 0.08$, $p < 0.001$, 95% CI [−0.48, −0.15]). The indirect effect of the Horizontal treatment on ethicality via responsibility was significant ($b = 0.32$, $SE = 0.13$, 95% CI [0.10, 0.59]), as was the indirect effect for the Vertical treatment ($b = 0.25$, $SE = 0.12$, 95% CI [0.06, 0.52]). Neither the direct effect of the Horizontal treatment on

Figure 4. Ethical evaluations score (*y*-axis) of decisions and moral accountability for each competitive context (group) in Study 3

Table 5. PROCESS model 4 mediation analysis for study 3

Effect	Path	*b*	SE	*t*	*p*	95% CI
Path a (X → M)						
Control → Moral Disengagement	(cont.)	4.99	0.18	27.66	<0.001	[4.63, 5.34]
Vertical → Moral Disengagement	a_1	−0.80	0.26	−3.13	0.002	[−1.30, −0.30]
Horizontal → Moral Disengagement	a_2	−1.01	0.26	−3.95	<0.001	[−1.51, −0.50]
Path b (M → Y)						
Moral Disengagement → Ethicality	b	−0.32	0.08	−3.76	<0.001	[−0.48, −0.15]
Path c' (X → Y, controlling for M)						
Vertical → Ethicality	C'_1	−0.20	0.28	0.72	0.472	[−0.35, 0.75]
Horizontal → Ethicality	C'_2	0.54	0.28	1.90	0.059	[−0.02, 1.09]
Total Effect (c)						
Vertical → Ethicality	c_1	0.45	0.28	1.62	0.108	[−0.10, 1.01]
Horizontal → Ethical	c_2	0.45	0.28	3.06	0.003	[0.30, 1.41]
Indirect Effects (ab)						
Vertical → Moral Disengagement → Ethicality	a_1b	0.25	0.12			[0.06, 0.52]
Horizontal → Moral Disengagement → Ethicality	a_2b	0.32	0.13			[0.100.0.59]

Note(s): The indirect effects' standard errors and confidence intervals are generated based on 5,000 bootstrap samples

Source(s): Authors' own work

ethicality (controlling for the mediator) ($b = 0.54$, $SE = 0.28$, $p = 0.059$, 95% CI [−0.02, 1.09]) nor the direct effect of the Vertical treatment were significant ($b = -0.20$, $SE = 0.28$, $p = 0.472$, 95% CI [−0.35, 0.75]). This pattern provides strong support for H3, demonstrating that competitive pressure operates through the theorized moral disengagement mechanisms. These findings validate that the Horizontal and Vertical manipulations activate distinct moral disengagement mechanisms underlying the behavioral effects observed in Studies 1 and 2.

Discussion

Theoretical implications

While RAI literature has made strides in examining internal organizational factors related to accountability, transparency, governance structures and leadership policies (e.g. Papagiannidis *et al.*, 2025), it has overlooked the role of competition in shaping RAI decision-making. We address calls for more precise conceptualization of how external pressures influence ethical decision-making in AI contexts (Birkstedt *et al.*, 2023) by offering the first systematic study of how competitive environments influence organizational choices related to RAI.

Drawing parallels to economic product differentiation [6], we distinguish between horizontal unethical competition, which arises from pressure to match direct competitors engaging in equally unethical AI practices, and vertical unethical competition, which stems from a competitor engaging in severely unethical AI practices. By theorizing these forms of unethical competition, we advance the understanding of how external market forces condition the implementation and sustainability of RAI initiatives. Thus, existing RAI decision-making models (e.g. Papagiannidis *et al.*, 2025) should integrate external competitive factors to better reflect firm challenges. By incorporating the dynamics of horizontal and vertical competition with regulatory focus, our study shows how external pressures can undermine or reshape internal commitments to ethical AI.

Our empirical findings also reveal distinct and robust effects of competitive pressures on RAI deployment. Horizontal unethical competition undermines ethical behavior, fostering a race to the bottom that normalizes irresponsible AI deployment. Contrasting our initial theorization predicting a uniformly negative effect on the launch of unethical AI, the prevention-focused orientation can dampen or even reverse this effect. Our counterintuitive finding can be explained through the interaction of ethical ambiguity and slippery slope mechanisms. Mittelstadt (2019) argued that RAI principles are inherently abstract and vague, making it difficult to draw clear ethical lines in deployment decisions. This baseline ambiguity allows decision-makers to rationalize launching potentially unethical AI systems, as the ethical boundaries remain unclear. Welsh *et al*. (2015) showed that people become more morally disengaged when potential unethicalness develops gradually rather than abruptly, as small ethical transgressions pave the way for larger future violations. However, ethical risks presented abruptly, e.g. when exposed to severe competitor unethicalness like accessing consumer microphones without consent, can disrupt the moral disengagement process by making the potential endpoint of unethical behavior vividly apparent.

Importantly, Welsh *et al*. (2015) showed that regulatory focus moderates this slippery slope effect, with prevention-focused individuals being particularly motivated to avoid negative ethical trajectories once they become salient. The severe competitor behavior thus serves as an abrupt ethical clarification that breaks through the baseline ambiguity, activating prevention-focused individuals' core motivation to step back from the ethical edge entirely rather than risk sliding down the slippery slope. In contrast, promotion-focused individuals, even when seeing the severe competitor behavior, remain more willing to take calculated risks and may still engage in advantageous comparison reasoning. By demonstrating that competition can activate different moral disengagement pathways and lead to different decision-making depending on goals, we also respond to Achar and Lee's (2022) call for studying disengagement mechanisms individually.

Managerial and policy implications

Our research highlights that firms should analyze competitors' AI use and be cognizant of how moral disengagement tactics can influence decision-makers to justify implementing unethical AI. Comparisons to unethical competitors may provide prevention-focused firms with motivation for implementing RAI. However, as more organizations implement unethical AI, horizontal unethical competition scenarios may increase and RAI could become a competitive advantage as more competitors engage in AI-based UPB. Firms wishing to improve RAI decision-making by emphasizing prevention-focused objectives must also heed slippery slope effects (Welsh *et al.*, 2015) when comparing against unethical competitors. Managers can also respond by establishing competitor intelligence functions that track both technical and ethical choices, running ethics stress-tests of potential launch decisions and integrating RAI criteria into product approval gates. Managers can also reinforce these measures by mandating regular reviews of how competitive pressures are influencing AI deployment decisions.

Technologically advanced firms implementing new and ethically questionable AI products and services may create vertical unethical competition scenarios in their respective industry. Whereas prevention-focused firms may resist engaging in similar behaviors, promotion-focused firms can leverage this as an opportunity to maximize their own success through irresponsible AI deployment. However, a closer analysis of our results indicates that the benefits of prevention-focused goals may only serve to slow the adoption of industry-wide unethical AI. Over time, promotion-focused firms will likely implement unethical practices in advantageous comparison situations. Eventually, the number of firms engaging in these practices may cross a critical threshold, transforming the market from vertical to horizontal unethical competition. As more promotion-focused competitors implement unethical AI systems, horizontal unethical competition may become the dominant method for unethical AI deployment, driving prevention-focused firms to adopt unethical AI practice to protect their competitive position.

To guard against this erosion, managers can take several steps. Firstly, managers can embed RAI checkpoints into launch processes, including independent review of privacy, fairness and explainability before release. Secondly, managers can invest in transparency mechanisms (e.g. public audits and model documentation) to build trust with regulators and customers. Thirdly, joining industry alliances or standard-setting bodies may prevent a race to the bottom by making responsible practices part of the competitive baseline. If firms neglect these steps, market dynamics risk eroding ethical standards across an entire industry. By contrast, early investment in RAI practices can protect reputation, attract ethically minded customers and employees, and reduce the likelihood of costly regulatory intervention. These effects highlight the need for swift organizational, policy and legislative actions. Delaying responses to unethical AI systems allows for more diffusion of responsibility, opportunities to arise, rendering interventions too late.

Limitation and future research

Our exploration centered on two types of competitive scenarios that can activate specific moral disengagement mechanisms. Our results indicate that these competitive scenarios impact RAI decision-making through distinct pathways, with severely unethical competitors unexpectedly motivating more responsible decisions under prevention-focused conditions. We hope future research formally examines the mechanisms underpinning this surprising result. Moreover, this result highlights the value of investigating moral disengagement mechanisms individually since it may expose novel effects across applications and generate interesting interactions with regulatory focus and other relevant behavioral-based theories.

We focused on diffusion of responsibility and advantageous comparison as they are the mechanisms most directly activated by competitive environments. While competitive dynamics are likely to manifest through these mechanisms, the other six moral disengagement mechanisms could also influence AI-based UPB through organizational, cultural or individual-level antecedents, representing avenues for future investigation. Future research should also explore how organizations may systematically leverage different moral disengagement mechanisms to justify AI-based UPB, potentially examining how organizational cultures, leadership practices or industry norms activate specific disengagement pathways beyond competitive dynamics. The studies used vignettes and online samples, creating limitations regarding external validity. While online samples reduce stakes and protect anonymity, therefore are appropriate when investigating ethical dilemmas, future research could enhance external validity by replicating these findings with professionals in relevant fields and contexts.

IRB and data and material availability statements

All studies were approved by the authors' IRB, and all participants provided consent before participating in the studies. The experimental materials, data and analysis code for all studies are available on OSF at https://osf.io/x7emc/?view_only=96b7448084ea4535aa7ad7553958d59a.

Acknowledgments

We thank the Special Issue editors and the anonymous review teams whose feedback substantially strengthened this manuscript. We are grateful to the UNSW Business School and BIZLAB for financial support for the experiments. We also thank seminar participants at University College Cork and the University of Houston for their valuable feedback, with special acknowledgements to Wynne Chin for insightful suggestions that informed the additional study designs conducted after the University of Houston seminar.

Appendix

Table A1. Study 1 horizontal and vertical competition manipulations

Horizontal manipulation	Vertical manipulation
To help make the recommendation, you researched your firm's competitors You found that 95% of your firm's competitors (1) Used target marketing and altered their pricing of tourism services using data collected from their consumers from previous interactions and (2) Used target marketing and altered their pricing of tourism services using data about their consumers purchased from third parties This was the same strategy as your Chief Marketing Officer's proposal	To help make the recommendation, you researched your firm's main competitor Your main competitor (1) Used target marketing and altered their pricing of tourism services using data collected from their consumers from previous interactions and (2) Used target marketing and altered their pricing of tourism services using data about their consumers purchased from third parties This was the same strategy as your Chief Marketing Officer's proposal Beyond this, your main competitor also (1) Accessed the microphone on consumer devices to collect data on their consumers' interests, opinions and recent conversations to further their AI-driven targeted marketing and pricing strategy, and (2) Without explicit consent, sold sensitive data about their consumers to third parties to boost revenue Both of these actions were definitely unethical behaviors

Source(s): Authors' own work

Table A2. Study 2 regulatory focus manipulations

Promotion focus	Prevention focus
Alex Manipulation Alex is someone who undertakes work activities that allow them to get ahead. Alex accomplishes a lot at work, getting work done no matter what, often in a short amount of time	*Alex Manipulation* Alex is someone who follows all rules and regulations. Alex fulfills work obligations by completing work tasks correctly, focusing on the details of their work
MarkTech Manipulation MarkTech strives to be a leader in digital marketing and digital retail As a result, the firm's expectation of employees in positions like Alex's is that they ● They focus on achieving positive outcomes for the organization ● They frequently think about how they can achieve organizational successes MarkTech also wants its workers to strive to become the employees they "want" to be—to achieve their hopes, wishes, and aspirations within the organization. MarkTech firmly believes that instilling this attitude in their workers can help the workers focus on increasing profits and help the organization achieve success	*MarkTech Manipulation* MarkTech strives to maintain its position as a well-regarded firm in digital marketing and digital retail As a result, the firm's expectation of employees in positions like Alex's is that they ● They focus on preventing negative outcomes for the organization ● They frequently think about how they can prevent organizational failures MarkTech also wants their workers to strive to become the employees they "ought" to be—to fulfill their duties, responsibilities, and obligations within the organization. MarkTech believes instilling this attitude in their workers can help the workers focus on limiting profit losses and help the organization prevent failures

Source(s): Authors' own work

Notes

1. We removed duplicate responses ex-post and pre-screened using embedded variables based on recommendations of Qualtrics, potential subjects with Fraud Scores greater than 30, which means the response is likely fraudulent and with ReCaptcha scores less than 0.8, which is the probability that the responder is human. This exclusion criteria is consistent with data integrity recommendations (Guy *et al*., 2024).
2. The ethical concerns in our scenario stem from several factors. AI systems collect data from previous interactions, diminishing user autonomy by requiring consumers to "surrender" personal information without their explicit consent for these marketing purposes (Wertenbroch *et al*., 2020). The "black box" nature of the algorithm means consumers lack control over their personal data and cannot understand how their information influences pricing decisions, which can lead to discrimination (Saxena *et al*., 2025) and creates feelings of vulnerability (Ackerman *et al*., 2022). Research demonstrates that such practices, particularly involving sensitive third-party data, increase perceived risk and create feelings of exploitation rather than service (Puntoni *et al*., 2021).
3. For this experiment, we adopt a third-person narrative, since participants often conform to social norms and respond in ways consistent with socially desirable. A common method for preventing social desirability bias is to frame the experiment from the perspective of a third person, querying participants on how that person would behave (Chung and Monroe, 2003)
4. We note that the strength of this effect increases when controlling for demographics (e.g. $b = -0.62$, $SE = 0.32$, $p = 0.051$) or when not including the insignificant interaction effect in a regression model (e.g. $b = -0.58$, $SE = 0.23$, $p = 0.014$).
5. The difference in payment for Study 3 is to align with Prolific's payment recommendations.
6. In economics and marketing, horizontal differentiation refers to competing products with similar objective quality but differing in subjective appeal to individual preferences (e.g. color, style, interface). In contrast, vertical differentiation occurs when products vary in objectively ranked quality, such that one product is clearly better or worse than another.

References

Achar, C. and Lee, A.Y. (2022), "Regulatory fit intensifies moral predispositions", *Journal of Personality and Social Psychology*, Vol. 123 No. 3, pp. 481-502, doi: 10.1037/pspa0000306.

Ackermann, K.A., Burkhalter, L., Mildenberger, T., Frey, M. and Bearth, A. (2022), "Willingness to share data: contextual determinants of consumers' decisions to share private data with companies", *Journal of Consumer Behaviour*, Vol. 21 No. 2, pp. 375-386, doi: 10.1002/cb.2012.

Alsheibani, S., Cheung, Y., Messom, C. and Alhosni, M. (2020), "Winning AI strategy: six-steps to create value from artificial intelligence", *Americas Conference on Information Systems*, Vol. 11, pp. 1-10.

Asatiani, A., Malo, P., Nagbøl, P.R., Penttinen, E., Rinta-Kahila, T. and Salovaara, A. (2021), "Sociotechnical envelopment of artificial intelligence: an approach to organizational deployment of inscrutable artificial intelligence systems", *Journal of the Association for Information Systems*, Vol. 22 No. 2, pp. 325-352, doi: 10.17705/1jais.00664.

Bandura, A. (1991), "Social cognitive theory of self-regulation", *Organizational Behavior and Human Decision Processes*, Vol. 50 No. 2, pp. 248-287, doi: 10.1016/0749-5978(91)90022-l.

Bandura, A., Barbaranelli, C., Caprara, G.V. and Pastorelli, C. (1996), "Mechanisms of moral disengagement in the exercise of moral agency", *Journal of Personality and Social Psychology*, Vol. 71 No. 2, pp. 364-374, doi: 10.1037//0022-3514.71.2.364.

Benbya, H., Pachidi, S. and Järvenpää, S. (2021), "Special issue editorial: artificial intelligence in organizations: implications for information systems research", *Journal of the Association for Information Systems*, Vol. 22 No. 2, pp. 281-252, doi: 10.17705/1jais.00662.

Berente, N., Gu, B., Recker, J. and Santhanam, R. (2021), "Managing artificial intelligence", *MIS Quarterly*, Vol. 45 No. 3, pp. 433-1450, doi: 10.25300/misq/2021/16274.

Birkstedt, T., Minkkinen, M., Tandon, A. and Mäntymäki, M. (2023), “AI governance: themes, knowledge gaps and future agendas”, *Internet Research*, Vol. 33 No. 7, pp. 133-167, doi: 10.1108/intr-01-2022-0042.

Chung, J. and Monroe, G.S. (2003), “Exploring social desirability bias”, *Journal of Business Ethics*, Vol. 44 No. 4, pp. 291-302, doi: 10.1023/a:1023648703356.

Crowe, E. and Higgins, E.T. (1997), “Regulatory focus and strategic inclinations: promotion and prevention in decision-making”, *Organizational Behavior and Human Decision Processes*, Vol. 69 No. 2, pp. 117-132, doi: 10.1006/obhd.1996.2675.

Ebrahimi, S., Abdelhalim, E., Hassanein, K. and Head, M. (2025), “Reducing the incidence of biased algorithmic decisions through feature importance transparency: an empirical study”, *European Journal of Information Systems*, Vol. 34 No. 4, pp. 636-664, doi: 10.1080/0960085x.2024.2395531.

Enholm, I.M., Papagiannidis, E., Mikalef, P. and Krogstie, J. (2022), “Artificial intelligence and business value: a literature review”, *Information Systems Frontiers*, Vol. 24 No. 5, pp. 1709-1734, doi: 10.1007/s10796-021-10186-w.

Gino, F. and Margolis, J.D. (2011), “Bringing ethics into focus: how regulatory focus and risk preferences influence (un) ethical behaviour”, *Organizational Behavior and Human Decision Processes*, Vol. 115 No. 2, pp. 145-156, doi: 10.1016/j.obhdp.2011.01.006.

Guy, A.A., Murphy, M.J., Zelaya, D.G., Kahler, C.W. and Sun, S. (2024, In press), “Data integrity in an online world: demonstration of multimodal bot screening tools and considerations for preserving data integrity in two online social and behavioral research studies with marginalized populations”, *Psychological Methods*, doi: 10.1037/met0000696.

Hayes, A.F. and Montoya, A.K. (2017), “A tutorial on testing, visualizing, and probing an interaction involving a multicategorical variable in linear regression analysis”, *Communication Methods and Measures*, Vol. 11 No. 1, pp. 1-30, doi: 10.1080/19312458.2016.1271116.

Higgins, E.T. (1998), “Promotion and prevention: regulatory focus as a motivational principle”, *Advances in Experimental Social Psychology*, Vol. 30, pp. 1-46, doi: 10.1016/S0065-2601(08)60381-0.

Higgins, E.T. (2002), “How self-regulation creates distinct values: the case of promotion and prevention decision making”, *Journal of Consumer Psychology*, Vol. 12 No. 3, pp. 177-191, doi: 10.1207/s15327663jcp1203_01.

Higgins, E.T., Idson, L.C., Freitas, A.L., Spiegel, S. and Molden, D.C. (2003), “Transfer of value from fit”, *Journal of Personality and Social Psychology*, Vol. 84 No. 6, pp. 1140-1153, doi: 10.1037/0022-3514.84.6.1140.

Hinds, J., Williams, E.J. and Joinson, A.N. (2020), “It wouldn’t happen to me: privacy concerns and perspectives following the Cambridge Analytica scandal”, *International Journal of Human-Computer Studies*, Vol. 143, 102498, doi: 10.1016/j.ijhcs.2020.102498.

Jobin, A., Ienca, M. and Vayena, E. (2019), “The global landscape of AI ethics guidelines”, *Nature Machine Intelligence*, Vol. 1 No. 9, pp. 389-399, doi: 10.1038/s42256-019-0088-2.

Johnson, R.E., King, D.D., Lin, S.H.J., Scott, B.A., Walker, E.M.J. and Wang, M. (2017), “Regulatory focus trickle-down: how leader regulatory focus and behaviour shape follower regulatory focus”, *Organizational Behavior and Human Decision Processes*, Vol. 140, pp. 29-45.

Kirshner, S. and Lawson, J. (2025), “Preventing promotion-focused goals: the impact of regulatory focus on responsible AI”, *Computers in Human Behavior: Artificial Humans*, Vol. 3, 100112, doi: 10.1016/j.chbah.2024.100112.

Kordzadeh, N. and Ghasemaghaei, M. (2021), “Algorithmic bias: review, synthesis, and future research directions”, *European Journal of Information Systems*, Vol. 31 No. 3, pp. 388-409, doi: 10.1080/0960085x.2021.1927212.

Lanaj, K., Chang, C.H.D. and Johnson, R.E. (2012), “Regulatory focus and work-related outcomes: a review and meta-analysis”, *Psychological Bulletin*, Vol. 138 No. 5, pp. 998-1034, doi: 10.1037/a0027723.

Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sesing, A. and Baum, K. (2021), "What do we want from explainable artificial intelligence (XAI)? A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research", *Artificial Intelligence*, Vol. 296, 103473, doi: 10.1016/j.artint.2021.103473.

Lee, H.K., Lee, J.S. and Keil, M. (2018), "Using perspective-taking to de-escalate launch date commitment for products with known software defects", *Journal of Management Information Systems*, Vol. 35 No. 4, pp. 1251-1276, doi: 10.1080/07421222.2018.1523604.

Lee, J.Y., Ko, D.W. and Lee, H. (2019), "Loneliness, regulatory focus, inter-personal competence, and online game addiction: a moderated mediation model", *Internet Research*, Vol. 29 No. 2, pp. 381-394, doi: 10.1108/intr-01-2018-0020.

Li, M. and Wan, Y. (2023), "Norms or fun? The influence of ethical concerns and perceived enjoyment on the regulation of deepfake information", *Internet Research*, Vol. 33 No. 5, pp. 1750-1773, doi: 10.1108/intr-07-2022-0561.

Liang, T.P., Robert, L., Sarker, S., Cheung, C.M.K., Matt, C., Trenz, M. and Turel, O. (2021), "Artificial intelligence and robots in individuals' lives: how to align technological possibilities and ethical issues", *Internet Research*, Vol. 31 No. 1, pp. 1-10, doi: 10.1108/intr-11-2020-0668.

Lockwood, P., Jordan, C.H. and Kunda, Z. (2002), "Motivation by positive or negative role models: regulatory focus determines who will best inspire us", *Journal of Personality and Social Psychology*, Vol. 83 No. 4, pp. 854-864, doi: 10.1037/0022-3514.83.4.854.

Luget, A., Asaftei, G.M., Roberts, R., Presten, B. and Ottenbreit, K. (2025), "Insights on responsible AI from the global AI trust maturity survey", McKinsey & Company, available at: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/insights-on-responsible-ai-from-the-global-ai-trust-maturity-survey (accessed 5 October 2025).

Lwin, M.O., Wirtz, J. and Stanaland, A.J. (2016), "The privacy dyad: antecedents of promotion-and prevention-focused online privacy behaviors and the mediating role of trust and privacy concern", *Internet Research*, Vol. 26 No. 4, pp. 919-941, doi: 10.1108/intr-05-2014-0134.

Marabelli, M., Newell, S. and Handunge, V. (2021), "The lifecycle of algorithmic decision-making systems: organizational choices and ethical challenges", *The Journal of Strategic Information Systems*, Vol. 30 No. 3, 101683, doi: 10.1016/j.jsis.2021.101683.

Marjanovic, O., Cecez-Kecmanovic, D. and Vidgen, R. (2022), "Theorising algorithmic justice", *European Journal of Information Systems*, Vol. 31 No. 3, pp. 269-287, doi: 10.1080/0960085x.2021.1934130.

Mikalef, P., Conboy, K., Lundstrom, J.E. and Popovic, A. (2022), "Thinking responsibly about responsible AI and 'the dark side' of AI", *European Journal of Information Systems*, Vol. 31 No. 3, pp. 257-268, doi: 10.1080/0960085x.2022.2026621.

Mittelstadt, B. (2019), "Principles alone cannot guarantee ethical AI", *Nature Machine Intelligence*, Vol. 1 No. 11, pp. 501-507, doi: 10.1038/s42256-019-0114-4.

Moore, C. (2008), "Moral disengagement in processes of organisational corruption", *Journal of Business Ethics*, Vol. 80 No. 1, pp. 129-139, doi: 10.1007/s10551-007-9447-8.

Morley, J., Floridi, L., Kinsey, L. and Elhalal, A. (2020), "From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices", *Science and Engineering Ethics*, Vol. 26 No. 4, pp. 2141-2168, doi: 10.1007/s11948-019-00165-5.

Papagiannidis, E., Mikalef, P. and Conboy, K. (2025), "Responsible artificial intelligence governance: a review and research framework", *The Journal of Strategic Information Systems*, Vol. 34 No. 2, 101885, doi: 10.1016/j.jsis.2024.101885.

Puntoni, S., Reczek, R.W., Giesler, M. and Botti, S. (2021), "Consumers and artificial intelligence: an experiential perspective", *Journal of Marketing*, Vol. 85 No. 1, pp. 131-151, doi: 10.1177/0022242920953847.

Reidenbach, R.E. and Robin, D.P. (1990), "Toward the development of a multidimensional scale for improving evaluations of business ethics", *Journal of Business Ethics*, Vol. 9 No. 8, pp. 639-653, doi: 10.1007/bf00383391.

Rhue, L. and Washington, A.L. (2020), “AI’s wide open: premature artificial intelligence and public policy”, *Boston University Journal of Science and Technology Law*, Vol. 26, p. 353.

Saxena, N.A., Zhang, W. and Shahabi, C. (2025), “Unveiling and mitigating bias in ride-hailing pricing for equitable policy making”, *AI and Ethics*, Vol. 5 No. 2, pp. 1549-1560, doi: 10.1007/s43681-024-00498-3.

Van-Dijk, D. and Kluger, A.N. (2004), “Feedback sign effect on motivation: is it moderated by regulatory focus?”, *Applied Psychology*, Vol. 53 No. 1, pp. 113-135, doi: 10.1111/j.1464-0597.2004.00163.x.

Welsh, D.T., Ordóñez, L.D., Snyder, D.G. and Christian, M.S. (2015), “The slippery slope: how small ethical transgressions pave the way for larger future transgressions”, *Journal of Applied Psychology*, Vol. 100 No. 1, pp. 114-127, doi: 10.1037/a0036950.

Wertenbroch, K., Schrift, R.Y., Alba, J.W., Barasch, A., Bhattacharjee, A., Giesler, M., Knobe, J., Lehmann, D.R., Matz, S., Nave, G., Parker, J.R., Puntoni, S., Zheng, Y. and Zwebner, Y. (2020), “Autonomy in consumer choice”, *Marketing Letters*, Vol. 31 No. 4, pp. 429-439, doi: 10.1007/s11002-020-09521-z.

Wu, J., Huang, L. and Zhao, J.L. (2019), “Operationalizing regulatory focus in the digital age: evidence from an e-commerce context”, *MIS Quarterly*, Vol. 43 No. 3, pp. 745-764, doi: 10.25300/misq/2019/14420.

Wu, X., Zhou, Z. and Chen, S. (2023), “A mixed-methods investigation of the factors affecting the use of facial recognition as a threatening AI application”, *Internet Research*, Vol. 34 No. 5, pp. 1872-1897, doi: 10.1108/intr-11-2022-0894.

Zhou, X., Tang, J. and Wang, T. (2024), “Effect of the fit between situational regulatory focus and feedback focus on customers’ co-design behavior”, *Internet Research*, Vol. 34 No. 5, pp. 1818-1844, doi: 10.1108/intr-11-2022-0861.

Further reading

Akter, S., Dwivedi, Y., Sajib, S., Biswas, K., Bandara, R.J. and Michael, K. (2022), “Algorithmic bias in machine learning-based marketing models”, *Journal of Business Research*, Vol. 144, pp. 201-216, doi: 10.1016/j.jbusres.2022.01.083.

Chen, M., Chen, C.C. and Sheldon, O.J. (2016), “Relaxing moral reasoning to win: how organizational identification relates to unethical pro-organizational behavior”, *Journal of Applied Psychology*, Vol. 101 No. 8, pp. 1082-1096, doi: 10.1037/apl0000111.

Davenport, T.H. and Ronanki, R. (2018), “Artificial intelligence for the real world”, *Harvard Business Review*, Vol. 96 No. 1, pp. 108-116.

Moore, C., Detert, J.R., Klebe Treviño, L., Baker, V.L. and Mayer, D.M. (2012), “Why employees do bad things: moral disengagement and unethical organisational behaviour”, *Personnel Psychology*, Vol. 65 No. 1, pp. 1-48, doi: 10.1111/j.1744-6570.2011.01237.x.

Newman, A., Le, H., North-Samardzic, A. and Cohen, M. (2020), “Moral disengagement at work: a review and research agenda”, *Journal of Business Ethics*, Vol. 167 No. 3, pp. 535-570, doi: 10.1007/s10551-019-04173-0.

Corresponding author
Samuel N. Kirshner can be contacted at: s.kirshner@unsw.edu.au