ETHICAL LEADERSHIP IN AI-AUGMENTED ORGANIZATIONS: A STRUCTURAL MODEL OF TRUST AND ACCOUNTABILITY IN ENHANCING DECISION QUALITY AND EMPLOYEE BEHAVIORAL OUTCOMES

SONIA N. KHAWAND

Lebanese International University
E-mail: Sonia.khawand@liu.edu.lb

Abstract - The rapid proliferation of artificial intelligence (AI) in organizational decision-making has introduced profound challenges for leadership, trust, and accountability. Despite growing scholarly attention to AI adoption and its organizational consequences, the mechanisms through which leadership shapes the relational and cognitive dimensions of human–AI interaction remain theoretically underdeveloped. Drawing on Social Learning Theory and the Technology Acceptance Model, this paper advances a structural theoretical model in which ethical leadership moderates the relationship between AI use and employee trust — encompassing both trust in AI systems and trust in leadership — with downstream effects on decision quality and employee behavioral outcomes. We propose seven formal propositions and develop a conceptual model that situates ethical leadership as a critical boundary condition governing whether AI augmentation strengthens or undermines employee trust within organizations. The model contributes to the literatures on ethical leadership, organizational trust, human–AI collaboration, and responsible AI governance. Implications for theory, management practice, and future empirical research are discussed.

Keywords - Ethical Leadership, Artificial Intelligence, Employee Trust, Trust in AI, Decision quality, behavioral Outcomes, Social Learning Theory, Technology Acceptance Model, AI Governance

I. INTRODUCTION

The decision-making, process of organizing the work, and the distribution of accountability are changing at a fundamental level in organizations from all sectors. AI technologies are integrated into organization processes — from algorithmic hiring and performance management to AI support for strategic planning and predictive analytics — bringing a new type of actor into organizational life: AI capable, opaque, and consequential [1], [2], [3]. A key question management scholars and practitioners must grapple with is when does the human underpinnings of effective organizational functioning get enhanced by AI augmentation, and when does it get undermined?

Trust is among these bases. Trust — in colleagues, in management, and in organizational systems — is a prerequisite for cooperation, psychological safety, and adaptive performance, as established in the organizational trust literature [4], [5]. The integration of AI into organizational decision-making disrupts current trust dynamics in unique ways. Concerns that AI recommendations might not be reliable [6], [7], [8], that they reduce the autonomy and relevance of individual employees, or that they appear to come from unsavory intent behind introducing AI into the institution — these fears impact how employees interact with AI tools, the quality of decisions they make when working with AI, and the downstream effects on their organizational attitudes.

Leadership becomes a key ingredient in this context. Ethical leaders are uniquely placed to influence the meaning and reaction of employees to AI-enhanced work settings [9], [10]. Ethical leaders can serve as interpretive anchors to navigate the psychological implications of AI on employee trust by setting clear expectations, explaining the goals of AI application, and ensuring a consistent approach to employee wellbeing and fairness. However, AI can foster mistrust, poorer decision-making, and negative employee behaviors in settings characterized by ethical confusion or indifferent corporate leadership.

Although this intersection is theoretically rich, no research has yet identified a single model that specifies how ethical leadership moderates the AI–trust relationship, how trust dimensions mediate downstream effects, and how these relationships flow to behavioral outcomes. The present paper advances a theoretical model from two complementary theories: Social Learning Theory (SLT) [11], [12] and the Technology Acceptance Model (TAM) [13], [14]. SLT offers a microfoundational argument for the mechanisms through which ethical leaders influence employee cognition and behavior through observation, modeling, and reinforcement. TAM dimensions perceived usefulness and ease of use constitute the cognitive architecture of employees' evaluation of AI systems, co-created by the ethical leadership context.

The paper is organized as follows: Section II reviews the relevant literatures over the five core constructs. Section III develops the theoretical framework and

presents seven formal propositions around a conceptual model. Section IV discusses theoretical contributions and implications. Section V recognizes limitations and suggests future research directions. Section VI concludes.

II. LITERATURE REVIEW

A. Ethical Leadership

Ethical leadership is the modeling of normatively appropriate conduct through personal actions and interpersonal relations, and the facilitation of such behavior to followers through two-way communication, reinforcement, and decision-making [9, p. 120]. This definition emphasizes two complementary dimensions: moral person (integrity, concern for others, ethical decision-making) and moral manager (role modeling, ethical communications, and use of rewards and discipline to ensure follower accountability) [15], [10]. Both dimensions are essential and carry distinct meanings for organizational behavior.

Evidence from research has consistently found connections between ethical leadership and positive organizational outcomes — followers' trust [9], [16], organizational citizenship behaviors (OCB) [17], [18], voice behavior [19], and psychological safety [20]. Such effects are hypothesized to occur primarily via social exchange processes: workers who perceive their leaders as ethical respond in kind by increasing trust, discretionary effort, and organizational identification [21], [22].

However, the ethical leadership literature has been largely developed in human-to-human leadership contexts. AI's impact on organizational decision-making is a new development that current ethical leadership models have yet to fully incorporate. These constitute the basic theoretical reasons for the model presented in this paper.

B. Artificial Intelligence in Organizations

AI is an umbrella term referring to computing-based systems that perform tasks ordinarily considered to require human intelligence, including reasoning, perception, and learning [23]. In organizations, AI is used across functions ranging from simple automation to advanced prediction, classification, and recommendation tasks [1], [24]. In complex and uncertain domains, AI augmentation — supporting rather than replacing human judgment — is generally superior to unaided human judgment and fully automated AI processes [25], [6]. However, employees must trust the recommendations AI can offer rather than reject or over-depend on them for this benefit to be realized.

AI also poses ethical dilemmas that directly impact HR professionals and organizational leaders: algorithmic bias and fairness [26], [27], opacity and explainability [28], accountability for AI-mediated decisions [2], and employee surveillance and autonomy concerns [29]. Without effective leadership on these factors, employees may be skeptical of AI adoption in their workflows, leading to lower productivity and resistance.

C. Employee Trust: A Two-Dimensional Conceptualization

The term trust has been defined at interpersonal, organizational, and institutional levels [4], [5], [30]. In this model, we differentiate two analytically distinct yet empirically related types of trust: trust in AI and trust in leadership.

Trust in AI is defined as employees' acceptance and reliance on AI-recommended decisions in critical work contexts [8], [31]. Following Mayer et al. [4], trust in AI can be decomposed into perceived ability (competence of the AI system), benevolence (the AI system designed to serve employee interests), and integrity (the AI system operating according to acceptable principles). Not only are technical characteristics of the AI system important, but also the organizational context and leadership credibility regarding AI's purpose and protections [32], [33].

Trust in leadership, by contrast, is confidence in the reliability, integrity, and beneficence of direct managers and organizational leaders [9], [34]. Trust in leadership is especially critical in AI-augmented environments because leaders are the primary human agents with decision-making authority over AI implementation, its necessity, its limitations, and the opportunity to integrate AI recommendations into organizational culture and interests.

D. Decision Quality

Decision quality is defined as the extent to which organizational decisions are well-reasoned and complete, consistent with organizational goals, and error-robust [35], [36]. In AI-enhanced environments, decision quality results from the synthesis of human judgment and AI-generated information. Employees who trust AI recommendations are more inclined to interact with AI outputs constructively, integrate them with relevant contextual information, and make higher-quality decisions than those who either reject or blindly accept AI outputs [6], [37], [25].

Trust in leadership is also directly linked to decision quality through its effects on information sharing, psychological safety, and organizational learning [38], [39]. In environments where employees feel confident in their leaders, they are more inclined to articulate concerns about AI suggestions, surface contextual information that may conflict with AI

inputs, and participate in thoughtful deliberation necessary for high-quality group decision-making.

E. Employee Behavioral Outcomes

Employee behavioral outcomes encompass the range of discretionary and prescribed behaviors expressing employees' orientations toward the organization, their work, and their colleagues. In this model, we focus on two behavioral dimensions particularly relevant to AI-augmented organizational contexts: organizational citizenship behavior (OCB) and adaptive performance.

OCB refers to discretionary employee behaviors that support organizational effectiveness without being formally required or directly rewarded [40], [41]. OCBs are sensitive to employee perceptions of fairness, leader integrity, and organizational support — constructs directly implicated in the ethical leadership and trust dynamics modeled here [17]. Adaptive performance refers to employees' capacity to modify their behavioral strategies in response to changing work demands — including those created by AI-augmented workflows [42], [43].

III. THEORETICAL FRAMEWORK AND PROPOSITIONS

A. Theoretical Grounding

The conceptual model is grounded in two complementary theoretical traditions. Social Learning Theory (SLT; Bandura, 1977, 1986) [11], [12] provides the logic for the mechanisms through which ethical leaders affect employee trust cognitions. SLT suggests that people learn their attitudes, values, and behavioral scripts from direct experience and from salient role models — those possessing power, status, and credibility. In organizational settings, leaders serve as dominant social models whose behavior conveys norms of acceptable work practice, including the use of AI systems [12], [9].

SLT is complemented by the Technology Acceptance Model (TAM) [13], [14], which explains the cognitive architecture employees use to assess AI systems. TAM posits that the primary determinants of technology adoption are perceived usefulness (PU) — the extent to which a technology is thought to improve job performance — and perceived ease of use (PEOU) — the extent to which using a technology will not require much mental effort. We propose that ethical leadership systematically influences both PU and PEOU perceptions by providing employees with contextual information, psychological safety, and relational confidence, enabling them to assess and interact effectively with AI tools.

B. The Conceptual Model

The model is structured around a central moderation hypothesis: ethical leadership moderates the relationship between AI use and employee trust, such that the positive association between AI use and trust is stronger under conditions of high ethical leadership than under conditions of low ethical leadership. This moderation operates through two trust dimensions — trust in AI and trust in leadership — which together mediate the effects of AI use on decision quality and, through decision quality, on employee behavioral outcomes. Figure 1 presents a schematic representation of the full conceptual model.

The model advances three related theoretical claims. First, ethical leadership is a necessary but insufficient condition for AI use to generate positive trust outcomes. Second, trust in AI and trust in leadership are complementary but functionally distinct — each carries unique variance in decision quality outcomes. Third, the cascade from trust to decision quality to behavioral outcomes is conditioned by the quality of the leadership context in which AI use occurs.

C. Propositions

1) AI Use and Employee Trust

The relationship between AI use and employee trust is theoretically bidirectional: AI adoption can increase trust by demonstrating organizational investment in decision support and analytical capability, or decrease trust by generating uncertainty about job security, autonomy, and the fairness of AI-mediated decisions [6], [8], [7]. Drawing on TAM, we argue that in the absence of contextual enablers, employees are likely to evaluate AI use with ambivalence — aware of its potential usefulness, but uncertain about its implications for their agency and organizational treatment.

Proposition 1a: The use of AI in organizational decision-making is positively associated with employee trust in AI systems, such that greater AI use exposure increases employees' perceived ability, benevolence, and integrity attributions toward AI.

Proposition 1b: The use of AI in organizational decision-making is positively associated with employees' trust in organizational leadership, insofar as AI adoption is perceived as a legitimate investment in organizational capability and employee decision support.

2) Ethical Leadership as Moderator

SLT provides the primary theoretical basis for our core moderation hypothesis. Ethical leaders model accountable behavior, communicate transparently about the purposes and limitations of AI use, and establish normative expectations that AI will be deployed in ways consistent with employee interests and organizational values [9], [10]. These behaviors function as contextual signals that resolve the trust

ambivalence associated with AI use, enabling employees to develop calibrated confidence in both AI systems and their leadership. Conversely, in low ethical leadership contexts, the normative signals are absent or contradictory, and even technically capable AI systems are unlikely to generate sustained employee trust.

Proposition 2a: Ethical leadership positively moderates the relationship between AI use and trust in AI, such that the positive association between AI use and trust in AI is stronger when ethical leadership is high than when it is low.
Proposition 2b: Ethical leadership positively moderates the relationship between AI use and trust in leadership, such that the positive association between AI use and trust in leadership is stronger when ethical leadership is high than when it is low.

3) Trust and Decision Quality
We propose that both trust dimensions are positive antecedents of decision quality, but through distinct mechanisms. Trust in AI promotes decision quality by increasing employees' willingness to engage with and appropriately weight AI-generated recommendations, reducing both algorithmic aversion and automation bias [6], [44]. Calibrated trust — neither excessive nor deficient — enables the productive integration of human and AI judgment. Trust in leadership promotes decision quality by fostering psychological safety, information sharing, and constructive deliberation [38], [39].
Proposition 3a: Employee trust in AI is positively associated with decision quality, such that calibrated trust — neither excessive nor deficient — maximizes the productive integration of human and AI judgment.
Proposition 3b: Employee trust in leadership is positively associated with decision quality, through its facilitation of psychological safety, information sharing, and constructive deliberation in AI-augmented decision contexts.

4) Decision Quality and Employee Behavioral Outcomes
High-quality decision processes generate positive employee behavioral outcomes through multiple pathways. When employees observe that AI-augmented decisions are well-reasoned, fair, and consistent with organizational values, they experience procedural justice — a powerful predictor of OCB [45], [46]. Decision quality also creates a positive feedback loop in which employees' confidence in organizational competence increases their willingness to invest discretionary effort. In AI-augmented contexts specifically, high decision quality demonstrates the viability of human–AI collaboration, reducing employee anxiety about AI displacement and supporting adaptive performance.
Proposition 4: Decision quality is positively associated with employee behavioral outcomes — specifically, organizational citizenship behavior and adaptive performance — such that higher decision quality increases both discretionary organizational contribution and behavioral flexibility in AI-augmented work contexts.

5) The Conditional Indirect Effect
The foregoing propositions collectively imply a conditional indirect effect: the positive influence of AI use on employee behavioral outcomes, transmitted through trust and decision quality, is contingent on the moderating role of ethical leadership. This conditional indirect effect — moderated mediation [47], [48] — specifies that ethical leadership is not merely additive to the AI–trust–behavior chain: it is a constitutive boundary condition without which the positive effects of AI use on organizational outcomes cannot be reliably expected.

Proposition 5: The indirect effect of AI use on employee behavioral outcomes (via trust in AI, trust in leadership, and decision quality) is moderated by ethical leadership, such that this indirect effect is significantly positive when ethical leadership is high but attenuated or non-significant when ethical leadership is low.

Figure 1. Conceptual Model: Ethical Leadership as Moderator in the AI–Trust–Decision Quality–Behavioral Outcomes Chain.

IV. DISCUSSION

A. Theoretical Contributions
This paper makes several distinct contributions to the theoretical literatures on ethical leadership, organizational trust, human–AI collaboration, and responsible AI governance.

First, we extend ethical leadership theory into the domain of AI-augmented organizations. By specifying ethical leadership as a moderator of the AI–trust relationship, the model identifies a new and theoretically important boundary condition for the effects of AI adoption on organizational outcomes. This contribution extends Brown et al.'s [9] foundational model by introducing AI use as a novel contextual variable.

Second, we advance a more refined conceptualization of trust in AI-augmented organizational contexts. By distinguishing between trust in AI and trust in leadership as complementary but functionally distinct constructs, the model clarifies the multiple pathways through which ethical leadership shapes trust and through which trust shapes decision quality. This two-dimensional conceptualization addresses a limitation in prior work, which has tended to treat trust as a unitary construct [8], [4].

Third, by integrating SLT and TAM within a single theoretical framework, the paper advances a more comprehensive explanation of the psychological processes through which ethical leadership shapes AI-related employee outcomes than either theory provides alone. SLT accounts for the social–normative mechanisms; TAM accounts for the cognitive-evaluative processes. Importantly, ethical leadership systematically shapes both the normative context and the evaluative criteria through which employees engage with AI systems.

Fourth, the model contributes to the emerging literature on responsible AI governance [2], [49], [50] by theorizing leadership as a governance mechanism. While existing work tends to focus on formal institutional structures, the present model identifies the everyday leadership behaviors of organizational actors as a critical, and potentially more proximal, governance resource.

B. Practical Implications

For organizational leaders, the model underscores that AI deployment decisions carry significant relational and behavioral consequences not reducible to technical considerations. Ethical leadership communication about AI — including transparency about algorithmic logic, fairness safeguards, and human accountability for AI-mediated decisions — is a strategic leadership capability that organizations should actively develop.

For human resource and talent management professionals, the model highlights decision quality and behavioral adaptation as key outcomes to monitor during AI adoption. Training programs for managers in AI-augmented environments should incorporate ethical leadership development components focused specifically on AI accountability, transparency, and employee-centered communication.

For AI governance practitioners and policymakers, the model implies that frameworks focusing exclusively on technical standards — algorithmic fairness, explainability, privacy — are necessary but insufficient. Organizational culture, leadership quality, and trust climate are endogenous governance factors that powerfully shape whether formal governance mechanisms produce trustworthy AI outcomes.

V. LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

Like all theoretical models, the present one has limitations that should guide future empirical work. Four main constraints are recognized.

First, the model is theoretical and has yet to be tested empirically. The propositions offered here are theoretical predictions, not findings. Future studies should operationalize the constructs and test the hypothesized relationships with appropriate quantitative methods — multivariate regression, structural equation modeling, and conditional process analysis [48] — across different organizational settings.

Second, ethical leadership is treated as a relatively stable moderating variable. In practice, ethical leadership in AI-enhanced organizations is likely to be fluid and variable. Future research should examine the conditions under which ethical leadership can be sustained in AI-intensive organizations.

Third, the model is employee-level and does not explicitly develop organizational- or industry-level consequences of the proposed relationships. Multi-level studies ranging from the individual to the team to the organizational level of analysis would greatly enrich the model.

Fourth, while SLT and TAM provide robust explanatory tools, other theoretical approaches — institutional theory, agency theory, and moral foundations theory — could make complementary or competing predictions. Future theoretical work should evaluate the boundary conditions and scope conditions of the proposed model.

Directions for future research include: (a) investigating cross-cultural variation in the ethical leadership–AI trust relationship; (b) examining individual difference variables — AI literacy, technology anxiety, and propensity to trust — as moderators; and (c) examining longitudinal dynamics and the possibility of trust erosion following high-profile AI failures.

VI. CONCLUSION

The integration of AI into organizational decision-making represents one of the most consequential transformations in the contemporary landscape of management and organizing. Yet technological capability alone does not determine organizational outcomes: the relational, ethical, and leadership context in which AI is deployed is at least as important as the sophistication of the AI system itself. This paper has argued that ethical leadership is a critical boundary condition governing whether AI augmentation strengthens or undermines the trust structures that underpin effective organizational functioning.

Through a structural theoretical model grounded in Social Learning Theory and the Technology Acceptance Model, we have proposed seven formal propositions specifying how ethical leadership moderates the AI–trust relationship, how trust in AI

and trust in leadership jointly shape decision quality, and how decision quality cascades into employee behavioral outcomes. The model advances a unified theoretical account of the leadership conditions under which AI-augmented organizations can realize the promise of human–AI collaboration while preserving the relational integrity that sustains organizational health.

The implications of this model extend beyond academic theory. As organizations invest increasingly in AI capabilities, the development of ethical leadership that can credibly navigate the accountability, transparency, and fairness challenges of AI-augmented work environments becomes a critical organizational capability. Scholars, practitioners, and policymakers alike would benefit from attending not only to the machines that augment organizational intelligence, but to the human leaders who shape the conditions under which that intelligence is — or is not — put in service of human flourishing.

REFERENCES

[1] E. Brynjolfsson and A. McAfee, The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. New York: W. W. Norton & Company, 2014.

[2] A. Dafoe, `AI governance: A research agenda,` Future of Humanity Institute, Univ. of Oxford, 2018.

[3] M. H. Jarrahi, "Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making," Business Horizons, vol. 61, no. 4, pp. 577–586, 2018.

[4] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," Acad. Manage. Rev., vol. 20, no. 3, pp. 709–734, 1995.

[5] D. M. Rousseau, S. B. Sitkin, R. S. Burt, and C. Camerer, "Not so different after all: A cross-discipline view of trust," Acad. Manage. Rev., vol. 23, no. 3, pp. 393–404, 1998.

[6] B. J. Dietvorst, J. M. Logg, and J. Logg, "Algorithm aversion: People erroneously avoid algorithms after seeing them err," J. Exp. Psychol.: Gen., vol. 144, no. 1, pp. 114–126, 2015.

[7] D. B. Shank, C. Graves, A. Gott, P. Gamez, and S. Rodriguez, "Feeling our way to machine minds: People's emotions when considering artificial intelligence minds," Computers Hum. Behav., vol. 98, pp. 256–266, 2019.

[8] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," Hum. Factors, vol. 46, no. 1, pp. 50–80, 2004.

[9] M. E. Brown, L. K. Treviño, and D. A. Harrison, "Ethical leadership: A social learning perspective for construct development and testing," Organ. Behav. Hum. Decis. Process., vol. 97, no. 2, pp. 117–134, 2005.

[10] L. K. Treviño, M. Brown, and L. P. Hartman, "A qualitative investigation of perceived executive ethical leadership," Hum. Relations, vol. 56, no. 1, pp. 5–37, 2003.

[11] A. Bandura, Social Learning Theory. Englewood Cliffs, NJ: Prentice-Hall, 1977.

[12] A. Bandura, Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[13] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," MIS Q., vol. 13, no. 3, pp. 319–340, 1989.

[14] V. Venkatesh and F. D. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies," Manage. Sci., vol. 46, no. 2, pp. 186–204, 2000.

[15] L. K. Treviño, L. P. Hartman, and M. Brown, "Moral person and moral manager: How executives develop a reputation for ethical leadership," Calif. Manage. Rev., vol. 42, no. 4, pp. 128–142, 2000.

[16] K. Kalshoven, D. N. Den Hartog, and A. H. De Hoogh, "Ethical leadership at work questionnaire (ELW): Development and validation of a multidimensional measure," Leadersh. Q., vol. 22, no. 1, pp. 51–69, 2011.

[17] D. M. Mayer, M. Kuenzi, R. Greenbaum, M. Bardes, and R. B. Salvador, "How low does ethical leadership flow? Test of a trickle-down model," Organ. Behav. Hum. Decis. Process., vol. 108, no. 1, pp. 1–13, 2009.

[18] R. F. Piccolo, R. Greenbaum, D. N. Den Hartog, and R. Folger, "The relationship between ethical leadership and core job characteristics," J. Organ. Behav., vol. 31, no. 2–3, pp. 259–278, 2010.

[19] F. O. Walumbwa and J. Schaubroeck, "Leader personality traits and employee voice behavior: Mediating roles of ethical leadership and work group psychological safety," J. Appl. Psychol., vol. 94, no. 5, pp. 1275–1286, 2009.

[20] J. B. Avey, M. E. Palanski, and F. O. Walumbwa, "When leadership goes unnoticed: The moderating role of follower self-esteem on the relationship between ethical leadership and follower behavior," J. Bus. Ethics, vol. 98, no. 4, pp. 573–582, 2011.

[21] P. M. Blau, Exchange and Power in Social Life. New York: Wiley, 1964.

[22] R. Cropanzano and M. S. Mitchell, "Social exchange theory: An interdisciplinary review," J. Manage., vol. 31, no. 6, pp. 874–900, 2005.

[23] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed. Upper Saddle River, NJ: Pearson, 2016.

[24] T. H. Davenport and R. Ronanki, "Artificial intelligence for the real world," Harvard Bus. Rev., vol. 96, no. 1, pp. 108–116, 2018.

[25] E. Kamar, "Directions in hybrid intelligence: Complementing AI systems with human intelligence," in Proc. 25th Int. Joint Conf. Artif. Intell., 2016, pp. 4070–4073.

[26] C. O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.

[27] S. Barocas and M. Hardt, "Fairness in machine learning," NeurIPS Tutorial, 2017.

[28] S. Wachter, B. Mittelstadt, and C. Russell, "Counterfactual explanations without opening the black box," Harvard J. Law & Technol., vol. 31, no. 2, pp. 841–887, 2017.

[29] S. Zuboff, The Age of Surveillance Capitalism. New York: PublicAffairs, 2019.

[30] D. H. McKnight and N. L. Chervany, "Trust and distrust definitions: One bite at a time," in Trust in Cyber-societies. Berlin: Springer, 2001, pp. 27–54.

[31] M. Heerink, B. Kröse, V. Evers, and B. Wielinga, "Assessing acceptance of assistive social agent technology by older adults," Int. J. Soc. Robot., vol. 2, no. 4, pp. 361–375, 2010.

[32] K. E. Schaefer, J. Y. Chen, J. L. Szalma, and P. A. Hancock, "A meta-analysis of factors influencing the development of trust in automation," Hum. Factors, vol. 58, no. 3, pp. 377–400, 2016.

[33] P. A. Hancock et al., "A meta-analysis of factors affecting trust in human-robot interaction," Hum. Factors, vol. 53, no. 5, pp. 517–527, 2011.

[34] K. T. Dirks and D. L. Ferrin, "Trust in leadership: Meta-analytic findings and implications for research and practice," J. Appl. Psychol., vol. 87, no. 4, pp. 611–628, 2002.

[35] P. C. Nutt, Why Decisions Fail. San Francisco: Berrett-Koehler, 2002.

[36] J. W. Dean and M. P. Sharfman, "Does decision process matter? A study of strategic decision-making effectiveness," Acad. Manage. J., vol. 39, no. 2, pp. 368–396, 1996.

[37] M. L. Cummings, "Man versus machine or man + machine?" IEEE Intell. Syst., vol. 29, no. 5, pp. 62–69, 2014.

[38] A. Edmondson, "Psychological safety and learning behavior in work teams," Admin. Sci. Q., vol. 44, no. 2, pp. 350–383, 1999.

[39] T. L. Simons and R. S. Peterson, "Task conflict and relationship conflict in top management teams," J. Appl. Psychol., vol. 85, no. 1, pp. 102–111, 2000.
[40] D. W. Organ, Organizational Citizenship Behavior: The Good Soldier Syndrome. Lexington, MA: Lexington Books, 1988.
[41] P. M. Podsakoff, S. B. MacKenzie, J. B. Paine, and D. G. Bachrach, "Organizational citizenship behaviors: A critical review," J. Manage., vol. 26, no. 3, pp. 513–563, 2000.
[42] E. D. Pulakos, S. Arad, M. A. Donovan, and K. E. Plamondon, "Adaptability in the workplace: Development of a taxonomy of adaptive performance," J. Appl. Psychol., vol. 85, no. 4, pp. 612–624, 2000.
[43] M. A. Griffin, A. Neal, and S. K. Parker, "A new model of work role performance," Acad. Manage. J., vol. 50, no. 2, pp. 327–347, 2007.
[44] R. Parasuraman and D. H. Manzey, "Complacency and bias in human use of automation," Hum. Factors, vol. 52, no. 3, pp. 381–410, 2010.
[45] R. Folger and M. A. Konovsky, "Effects of procedural and distributive justice on reactions to pay raise decisions," Acad. Manage. J., vol. 32, no. 1, pp. 115–130, 1989.
[46] R. H. Moorman, "Relationship between organizational justice and organizational citizenship behaviors," J. Appl. Psychol., vol. 76, no. 6, pp. 845–855, 1991.
[47] K. J. Preacher, D. D. Rucker, and A. F. Hayes, "Addressing moderated mediation hypotheses," Multivariate Behav. Res., vol. 42, no. 1, pp. 185–227, 2007.
[48] A. F. Hayes, Introduction to Mediation, Moderation, and Conditional Process Analysis. New York: Guilford Press, 2013.
[49] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," Nat. Mach. Intell., vol. 1, no. 9, pp. 389–399, 2019.
[50] J. Fjeld, N. Achten, H. Hilligoss, A. Nagy, and M. Srikumar, "Principled artificial intelligence," Berkman Klein Center Research Publication, 2020.
[51] D. H. Autor, F. Levy, and R. J. Murnane, "The skill content of recent technological change," Q. J. Econ., vol. 118, no. 4, pp. 1279–1333, 2003.
[52] M. E. Brown and M. S. Mitchell, "Ethical and unethical leadership," Bus. Ethics Q., vol. 20, no. 4, pp. 583–616, 2010.
[53] A. Newman, K. Kiazad, Q. Miao, and B. Cooper, "Examining the cognitive and affective trust-based mechanisms underlying the relationship between ethical leadership and organisational citizenship," J. Bus. Ethics, vol. 123, no. 1, pp. 113–123, 2014.

★★★