Quantifying Calibration: Bridging Trust and Reliance in Automation Across Dispositional Factors

Evelyn Goroza[1,*], Gavin McCarthy-Bui[1], Anne Zhao[1], Elsa R. Ostenson[1], and Dave B. Miller[1]

[1] *Tufts University, 200 College Avenue, Medford MA 02155, USA*

Abstract

As more sophisticated automation increases in prevalence across domains and users, it becomes more important for the Human-Computer Interaction community to better understand how to identify and foster calibrated trust and appropriate reliance. Analyzing the relationships between the variables trust, reliance, and system capability using ratio-scale measures provide a new way to quantify these factors. Using a cooperative human-robot task in a gameplay scenario, we aim to empirically investigate how well we can quantify the relationships between the three variables: i.e., trust to system capability, reliance to system capability, and trust to reliance. Because automation use does not occur in a vacuum, our model includes understudied yet salient measures of dispositional trust alongside a cooperative game task, to explore the effects of personality and cultural factors on human-automation trust and reliance across different levels of system capability. Understanding trust and reliance calibration in this way will offer insights valuable to designers of especially novel systems and the field of human-computer interaction study.

Keywords

Trust, Reliance, Automation, Dispositional Factors, Culture, Personality, Calibrated Trust, Appropriate Reliance, Human-Automation Interaction

1. Introduction

Interaction with sophisticated automated systems is increasingly part of everyday life, and for such systems to be properly used, they need to be designed to encourage human users to trust them appropriately. If a user's mental model is not properly calibrated, they can exhibit mistrust, which can lead to misuse of the system [1], [2]. Conversely, distrust can lead to disuse, thus forgoing a potential advantage of the benefits of automation [3].

In addition to system performance, individual traits—including both individual and cultural factors—influence trust and reliance [4]. To design automated systems that will elicit appropriate use, designers of such systems need to understand the correspondence between the actual system capabilities, operator trust (an attitude [2]) and operator reliance (a behavior referring to the user's engagement of the system [1], [2]). We therefore introduce an exploratory empirical study quantifying the alignment of trust, reliance, and system capability. Alongside studying this three-variable alignment, individual traits including both cultural factors and personality factors [5] are likely influences on trust and reliance, which our research includes as covariates in our model. To validate our conceptual model, we are in the process of conducting an empirical study on trust and reliance, using a cooperative task in a novel sorting task game utilizing an adaptable automation system.

2. Review of the Literature

2.1. Factors Influencing Calibrated Trust and Reliance on Automation

Lee and See [2] define calibrated trust as "[the level of] operator trust that matches system capability, leading to appropriate use" [2]. Experts such as pilots or industrial system operators are trained to know the capabilities and limits of the systems they use—but members of the public often do not

AutomationXP25: Hybrid Automation Experiences, April 27, 2025, Yokohama, Japan. In conjunction with ACM CHI'25.

* Corresponding author.
† These authors contributed equally.
evelyn.goroza@tufts.edu (E. Goroza); dave.miller@tufts.edu (D. B. Miller)

know the true limits of the systems the use, sometimes finding those limits with deadly consequences [6]. Hoff and Bashir's three-layer trust model [4] includes the layer Dispositional Trust, the defining characteristic being that is that it is a relatively stable trait over time. Derived from an empirical review of the literature, this layer identifies four primary sources of dispositional trust variability: culture, age, gender, and personality [4]. Razin and Feigh's work [7] explicates the factors from the dispositional layer as external antecedents to trust constructs, but still consider personality and cultural factors as "antecedents" to trust. Chiou and Lee's 2023 model [8] builds on these earlier works, factoring in the collaborative nature of more current technologies compared to those available twenty years before.

Alarcon et al. [9] discuss the impact of personality factors as antecedents of trust to illustrate the increasing relevance of distinguishing between human-human and human-machine interactions. These authors highlight an ongoing, pertinent debate over which of two models best characterizes human-automation interaction: the Computers are Social Actors paradigm [10], or the unique-agent hypothesis [11].

Measures of trust include self-report scales, such as those developed by Jian, Bisantz, and Drury [12], Lee and Moray [13], Merritt and Ilgen [14], and Merritt [15]. Considering trust as an antecedent of reliance [16] makes it possible to use behavioral reliance as a relevant measure as well, as was done by Miller et al. [17] and Fu et al. [18] in empirical studies where the driver of a partially automated vehicle needed to demonstrate appropriate trust: either allowing the vehicle full control, or taking control of the vehicle when necessary.

Kohn et al. [19] note nine trust behaviors, comprising: combined team performance, outcomes, compliance/agreement rate, decision time, delegation, stakes invested, intervention, reliance, response time, and verification. Notably, they postulate that the trust behavior known as delegation—assigning a task to the automation when the task could be performed by a human operator—is a relatively novel measure but is a strong indication of trust in that automation. This behavior is characterized by the participant ceding control to the agent, rather than taking it away as in the case of human intervention. By splitting a task with an automated system or robot, it is then possible to directly measure reliance on a continuous ratio scale—and this can be combined with survey measures of trust in a mixed-methods approach.

2.2. Dispositional Factors: Culture and Personality Traits

While much prior work has focused on system attributes, less attention has been given to how personality traits and cultural factors shape trust in and reliance on automation. While culture and personality are noted in the Hoff and Bashir model as dispositional factors of trust [4] and are noted in empirical research [5], [12], [14], [20], the specifics of how these factors influence trust and reliance necessitate further research.

Leung and Cohen [21] developed the Culture x Person x Situation (CuPS) approach to offer an integrated, balanced account of within- and between-culture variation, consisting of three distinct cultural logics: Dignity, Face, and Honor—in which the ideal types were developed from two underlying thematic principles considered salient in all societies: social order and self-valuation. In Dignity Cultures (e.g., Western Europe, North America), self-worth is internally derived and evaluated by personal standards. In contrast, Face Cultures (e.g., East Asia, Taiwan) derive self-worth externally, based on maintaining harmony and stable social hierarchies. These cultures emphasize collectivism, high power distance, and conformity to institutional norms that govern behavior. Finally, Honor Cultures (e.g., Middle East, Latin America, Mediterranean countries) derive self-worth from personal interactions, reputation, and the need to defend honor.

The CuPS approach circumvents issues of overly reductionist approaches to typifying individuals, such as incorrectly placing the sole focus on prototypical individuals of a single culture or incorrectly placing the sole focus on individual differences—both of these approaches ignore any emergent interactions between personality and culture, which has the potential to account for differences in behavior. Considering the cultural factors that influence personality traits of human interactants with automation, drawing from the CuPS model provides a platform on which to build an understanding of the cultural background component of human-automation trust and reliance, and calibration between these measures and system capability.

A study of operator trust in automation applying the cultural logics, conducted by Chien et al. [20] is one of the few studies examining culture as a dispositional trust factor. These logics relate to interaction with automated systems in terms of where the locus of control is placed, and how the reliability of the system working in collaboration (or at cross-purposes) with the human will affect total human-system performance.

Awad et al.'s survey research investigating the ideal moral orientation of autonomous vehicles [22] also found three clusters of cultural mores, resembling the ascribed "Ideal Types" of Dignity, Face, and Honor cultures. This finding further reinforces the importance of understanding the significance of culture and personality in the study of human-automation interactions, especially regarding trust and reliance.

3. Proposed Research Design

3.1. Research Questions

1. Can we quantify the alignment of trust, reliance, and system capability to classify calibration states continuously throughout human-automation interactions?
2. How do dispositional factors (e.g., propensity to trust, cultural values) influence trust and reliance calibration with true automated system capability?

3.2. Study Overview

We developed a mixed-methods, repeated-measures experimental design, with planned analyses conducted both within and across participants. The study will be administered via Prolific to a representative sample of US participants, and includes the following components:

1. Survey measures to assess dispositional traits.
2. The Calibratio Game designed for the purposes of the current study.

3.3. Participants

We plan to recruit a representative sample of approximately 100 American participants through the Prolific platform, based on powering the study to detect a small-medium effect with linear multiple regression modeling. As feasible, we will seek a balanced demographic representation (e.g., age, gender, cultural background) to better explore the influence of dispositional factors on trust and reliance.

3.4. Survey Measures

We chose to focus our survey measure battery on investigating the largely understudied dispositional factors of culture and personality. These measures are detailed in **Table** 1.

We operationalize the factor culture with the scale developed by Yao et al. [23]. In accordance with the CuPS approach [21], the Yao scale operationalizes the cultural logics—Dignity, Honor, and Face—by measuring perceived cultural norms. Their scale is grounded in an approach that focuses on measuring norms ("what is appropriate") rather than individual values ("what is important"). This is justified by the rationale that individuals use cultural norms to interpret context and guide actions in social interactions [23]—a necessity to study behavior which is lacking in the value approach, found in more commonly used cross-cultural scales, such as the seminal Hofstede's Cultural Dimensions [24] and Triandis' cultural syndromes [25]. Our use of the Yao scale will consider each of the logics corresponding to Dignity, Honor, and Face cultures as a separate factor as found in the original study [23].

Our personality dimensions are largely based on traits which describe an operator's dispositional trust in technology. We initiated this process by investigating the most relevant trust constructs to our study, drawing on information from recent reviews of trust measurement [7][19]. The model derived in Razin and Feigh's meta-review [7] ultimately guided our selection of trust constructs. Of these identified constructs, we narrowed our focus to the three constructs we deemed most relevant to our study design: capability-based trust, general trust, and faith in technology. Next, we operationalized the suggested scales which captured each of these constructs.

For general trust, we use Frazier et al.'s [26] Propensity to Trust scale (coded as PTT). We also selected McKnight's [27] Trusting Stance—General Technology scale, and renamed this to Propensity to Trust in Machines scale (coded as PM) for the purpose of highlighting a the comparison with Frazier's interpersonally-oriented PTT scale.

For Faith in Technology, we use McKnight's [27] Faith in General Technology (coded as FIGT).

To assess capability-based trust, we use McKnight's [27] Trusting Belief-Specific Technology (combining subscales for Reliability and Functionality), adapted for the automated agent in our study. We operationalize this adaptation by presenting the scale alongside a vignette briefly describing the task—the Calibratio Game—and our automated agent, Otto (further detailed in the following Subsection 3.5, Calibratio Game). Accordingly, we refer to this scale as Capability-Based Trust in Otto (coded as TIO).

Table 1
Selected Dispositional Factors

Dispositional Factor	Construct from Razin & Feigh [7]	Scale	Code
Culture	Culture	A Measurement Model for Dignity, Face, and Honor Cultural Norms [23]	DFH
Personality	Faith in Technology	Faith in General Technology [27]	FIGT
Personality	General Trust	Propensity to Trust in Machines [27]	PM
Personality	General Trust	Propensity to Trust (Interpersonal) [26]	PTT
Personality	Capability-Based Trust	Trust in a Specific Technology—Functionality and Reliability subscales [27]	TIO

We designed the game, Calibratio, for the present study to assess repeated measures of trust and reliance on the game's automated agent Otto. A sample of the game interface is depicted in **Figure 1**. The game is a collaborative task which simulates interaction modeled with adaptable automation. This is an especially useful model of control facilitated in this setting to induce the participant to modulate their desired allocation of reliance over a continuous period of time.

3.5. Calibratio Game

To measure trust and reliance (operationalized as delegation), participants will engage in a shared puzzle-piece sorting task with the automated agent, Otto. Participants sort three different puzzle pieces as they travel down a conveyor belt, which accumulates a combined player + Otto score in order to reflect a shared goal with the agent. Participants earn 1 point for every shaped sorted correctly when it reaches the sorting zone, 0 points for every shape sorted incorrectly, and 0 points for every shape missed (i.e., the piece passes the sorting zone). The workload is initially split with the participant given 100% of the workload and Otto given 0%. Participants are able to change the delegation of workload continuously throughout the round, as well as during pauses every 20 pieces where the conveyor stops and the participant is queried about their trust in Otto, and their trust in their own capability.

First, participants engage in a preliminary Baseline Round without Otto, in order to determine an individually calibrated spawn rate that will require the delegation of part of the sorting task to Otto. The spawn rate is dynamically adjusted to converges to the rate where the participant's performance lets them sort 80% of pieces working alone, in order to control for preexisting differences in skill level across operators. This rate is carried over to the following three experimental rounds. After the Baseline Round and before Round 1, Otto is introduced to the participants as the robot sorting agent who will help them sort the shapes if they delegate workload using the up and down keyboard keys. The participants are informed that Otto earns points in the same point system, points being combined in order to reflect a shared goal with the agent.

Across the three experimental rounds, there are three capability level values for Otto's performance, the order of capability levels being randomized to control for order effects. The capability level reflects the proportion of shapes which Otto is able to sort in time; Otto's sorting decision is programmed to always be correct. The three capability values are 20%, 50%, and 80%, these being chosen to test trust and reliance on an agent with low, medium, and high capability levels.

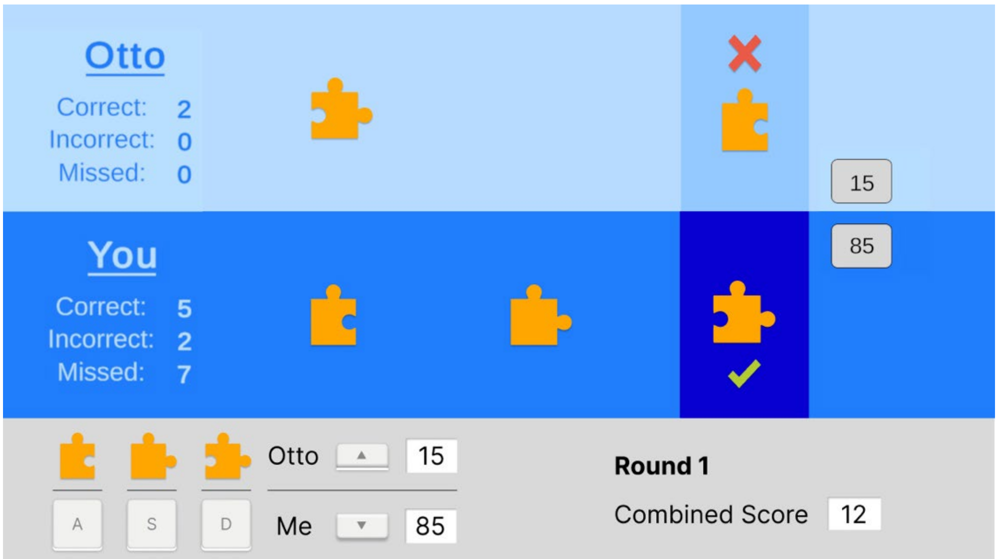

Figure 1: A sample prototypical interface of the Calibratio Game.

4. Data Analysis

Regression analysis will be used to determine the impact of self-reported culture and personality measures on trust and reliance. Below is a sample of our planned analysis for trust regression across participants for a given system capability level.

Equation 1 describes the variation in trust for each level of system capability (20, 50, or 80, as described in Section 3. Methods Subsection 3.5 Calibratio Game) accounted for by participants' self-reported dispositional factors in **Table** 1. Each of these selected dispositional factors will be numbered as a sequential trait, as denoted in the model.

Equation 1. Trust Regression Model for a Given System Capability Level.

$$T_i = \mu + X_{1,i}\beta_1 + \cdots + X_{n,i}\beta_n + \epsilon_i \quad (1)$$

where...

T_i is the trust level for participant i

μ is the average trust across participants

$x_{1,i}$ is trait 1 for person i

$x_{n,i}$ is trait n for person i

Regression analysis for reliance will resemble the trust regression model in **Equation 1**, but with the additional incorporation of reliance as a dynamic measure across each round of system capability.

We are currently investigating two potential courses for analysis: first, one which assumes that reliance values converge to a certain value by the end of the round, as exemplified in some preliminary pilot trials; and second, an alternate case in which reliance values are aggregated or characterized to capture all values over the course of the round.

5. Expected Contributions

5.1. Theoretical and Methodological Advancements

By empirically studying the relationship between trust and reliance using a ratio-scale measure, integrating dispositional antecedent factors into the trust calibration model, our work bridges a critical gap in the human-automation interaction literature. While others have related personality factors [28] and cultural influences to trust in automation [20], relating these traits to trust and reliance in a gameplay scenario is a novel contribution. Where extant models [5], [28], [29], [30] posit the existence of relationships between trust and reliance, automation capability and trust, and automation capability and reliance, and recent models propose the quantification of calibrated trust [5], [29], [30], [31], we examine these relationships empirically, on ratio scales, which to our knowledge has not yet been done.

5.2. Practical Implications

Insights from our research relating trait factors and the alignment of these variables—trust, reliance, and system capability—will inform the design of adaptable automation systems enabled to facilitate adjustable allocation of reliance to suit individual users' needs. This adaptation can draw on both personal factors to set initial automation parameters, which can subsequently adjust based on demonstrated reliance. As what may be appropriate may vary substantially between individuals and groups, our work provides new insights for understanding human-automation trust and reliance relationships by integrating individual cultural and relevant personality factors.

The applications of this work may inform the surmounting research in ethical design of adaptive and adaptable human-AI applications, including the assessment of cognitive health [32], educational practice [33], design of decision-support systems [34], and digital agriculture [35]. For instance, a recent article defines Artificial-Intelligence-Chatbots (AICs) Induced Cognitive Atrophy (AICICA), which refers to the potential deterioration of essential cognitive abilities resulting from an overreliance on AICs [32]. The authors call for research to investigate the effect of AICs across individual differences, as the human-like conversational nature, and immediacy of active and/or personalized information (as compared to static information, such as search engine results) might foster a deep sense of trust and reliance in some users, which can induce changes in brain circuitry—such as decision-making processes, learning, and emotional responses. They call for studies meticulously controlling for diverse populations and contexts to gain insights into engagement with AICs to assess overreliance and implications on cognitive functioning [32]. This study paves the way to research which focuses on engagement across cultural and dispositional factors of trust, which can aid in the work towards designing responsible technologies.

As prophesized by Bainbridge's seminal work on the "Ironies of Automation" [36], there exists the pitfall of system designs which expand, rather than eliminate, more problems for the human operator [36]. Operationally, this aligns with the onus on designers and implementors of extant and novel technologies to design automation technologies which foster Lee and See's cornerstone pillars of calibrated trust and appropriate reliance [2].

5.3. Workshop Engagement

Preliminary results, insights from building the game, and the development of the methodology will be shared with AutomationXP 2025 workshop participants working in the areas of human-agent interaction. Workshop attendees can test the gameplay experience and learn from our experience developing this mixed-methods approach to combining trait and behavioral measures to study trust and reliance.

6. Discussion and Future Work

As an exploratory study combining a number of measures, this study embarks on multimethod research to investigate the relationships between individual operator traits with trust and reliance. Future studies would benefit from more diverse user populations, such as a worldwide sample. Conducting research on trust and reliance conducted in more naturalistic setting rather than using an online game experience can also provide other insights into how cultural and personality factors influence the calibration of trust and reliance. Longitudinal studies may also provide more accurate insights of the evolution of these dynamic variables, as trust models are not static and will almost surely change over longer and repeated interactions between humans and automated agents. Further exploration of how multiple dispositional factors interact to influence trust calibration is also warranted.

7. Conclusion

Understanding the interplay between individual traits, trust, and reliance is crucial for advancing automation design, especially as the users of sophisticated automation become more diverse. By introducing the Calibration Ratio as a way to quantify the relationship between trust and reliance, and by incorporating dispositional factors into our trust and reliance model, our work offers both theoretical and practical contributions to the field of human-automation interaction. We look forward to engaging with the CHI community to further refine these ideas and explore their broader implications.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] R. Parasuraman and V. Riley, "Humans and Automation: Use, Misuse, Disuse, Abuse," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 39, no. 2, pp. 230–253, Jun. 1997, doi: 10.1518/001872097778543886.

[2] J. D. Lee and K. A. See, "Trust in Automation: Designing for Appropriate Reliance," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 46, no. 1, pp. 50–80, Mar. 2004, doi: 10.1518/hfes.46.1.50_30392.

[3] B. D. Sawyer, D. B. Miller, M. Canham, and W. Karwowski, "Human Factors and Ergonomics in Design of A3: Automation, Autonomy, and Artificial Intelligence," in *Handbook of Human Factors and Ergonomics*, 5th ed., John Wiley & Sons, Ltd, 2021, pp. 1385–1416. doi: 10.1002/9781119636113.ch52.

[4] K. A. Hoff and M. Bashir, "Trust in Automation Integrating Empirical Evidence on Factors That Influence Trust," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 57, no. 3, pp. 407–434, May 2015, doi: 10.1177/0018720814547570.

[5] G. M. Lucas, B. Becerik-Gerber, and S. C. Roll, "Calibrating workers' trust in intelligent automated systems," *PATTER*, vol. 5, no. 9, Sep. 2024, doi: 10.1016/j.patter.2024.101045.

[6] A. Davies, "Tesla's Latest Autopilot Death Looks Just Like a Prior Crash," *Wired*, May 16, 2019. Accessed: Sep. 09, 2019. [Online]. Available: https://www.wired.com/story/teslas-latest-autopilot-death-looks-like-prior-crash/

[7] Y. S. Razin and K. M. Feigh, "Converging Measures and an Emergent Model: A Meta-Analysis of Human-Machine Trust Questionnaires," *J. Hum.-Robot Interact.*, vol. 13, no. 4, p. 58:1-58:41, Nov. 2024, doi: 10.1145/3677614.

[8] E. K. Chiou and J. D. Lee, "Trusting Automation: Designing for Responsivity and Resilience," *Hum Factors*, vol. 65, no. 1, pp. 137–165, Feb. 2023, doi: 10.1177/00187208211009995.

[9] G. M. Alarcon *et al.*, "The effect of propensity to trust and perceptions of trustworthiness on trust behaviors in dyads.," *Behavior Research Methods*, vol. 50, no. 5, pp. 1906–1921, Oct. 2018, doi: 10.3758/s13428-017-0959-6.

[10] C. Nass and Y. Moon, "Machines and Mindlessness: Social Responses to Computers," *Journal of Social Issues*, vol. 56, no. 1, pp. 81–103, 2000, doi: 10.1111/0022-4537.00153.

[11] E. J. de Visser *et al.*, "Almost human: Anthropomorphism increases trust resilience in cognitive agents," *J Exp Psychol Appl*, vol. 22, no. 3, pp. 331–349, Sep. 2016, doi: 10.1037/xap0000092.

[12] J.-Y. Jian, A. M. Bisantz, and C. G. Drury, "Foundations for an empirically determined scale of trust in automated systems," *International journal of cognitive ergonomics*, vol. 4, no. 1, pp. 53–71, 2000, doi: 10.1207/S15327566IJCE0401_04.

[13] J. D. Lee and N. Moray, "Trust, self-confidence, and operators' adaptation to automation," *International Journal of Human-Computer Studies*, vol. 40, no. 1, pp. 153–184, Jan. 1994, doi: 10.1006/ijhc.1994.1007.

[14] S. M. Merritt and D. R. Ilgen, "Not All Trust Is Created Equal: Dispositional and History-Based Trust in Human-Automation Interactions," *Hum Factors*, vol. 50, no. 2, pp. 194–210, Apr. 2008, doi: 10.1518/001872008X288574.

[15] S. M. Merritt, "Affective Processes in Human–Automation Interactions," *Hum Factors*, vol. 53, no. 4, pp. 356–370, Aug. 2011, doi: 10.1177/0018720811411912.

[16] M. T. Dzindolet, S. A. Peterson, R. A. Pomranky, L. G. Pierce, and H. P. Beck, "The role of trust in automation reliance," *International Journal of Human-Computer Studies*, vol. 58, no. 6, pp. 697–718, Jun. 2003, doi: 10.1016/s1071-5819(03)00038-7.

[17] D. Miller *et al.*, "Behavioral Measurement of Trust in Automation The Trust Fall," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 60, no. 1, pp. 1849–1853, Sep. 2016, doi: 10.1177/1541931213601422.

[18] E. Fu *et al.*, "The Car That Cried Wolf: Driver Responses to Missing, Perfectly Performing, and Oversensitive Collision Avoidance Systems," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2019, pp. 1830–1836. doi: 10.1109/IVS.2019.8814190.

[19] S. C. Kohn, E. J. de Visser, E. Wiese, Y.-C. Lee, and T. H. Shaw, "Measurement of Trust in Automation: A Narrative Review and Reference Guide," *Front. Psychol.*, vol. 12, Oct. 2021, doi: 10.3389/fpsyg.2021.604977.

[20] S.-Y. Chien, M. Lewis, K. Sycara, A. Kumru, and J.-S. Liu, "Influence of Culture, Transparency, Trust, and Degree of Automation on Automation Use," *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 3, pp. 205–214, Jun. 2020, doi: 10.1109/THMS.2019.2931755.

[21] A. K.-Y. Leung and D. Cohen, "Within- and Between-Culture Variation: Individual Differences and the Cultural Logics of Honor, Face, and Dignity Cultures," *ResearchGate*, Oct. 2024, doi: 10.1037/a0022151.

[22] E. Awad, S. Dsouza, A. Shariff, I. Rahwan, and J.-F. Bonnefon, "Universals and variations in moral decisions made in 42 countries by 70,000 participants," *Proceedings of the National Academy of Sciences*, vol. 117, no. 5, pp. 2332–2337, Feb. 2020, doi: 10.1073/pnas.1911517117.

[23] J. Yao, J. Ramirez-Marin, J. Brett, S. Aslani, and Z. Semnani-Azad, "A Measurement Model for Dignity, Face, and Honor Cultural Norms," *Management and Organization Review*, vol. 13, pp. 1–26, Dec. 2017, doi: 10.1017/mor.2017.49.

[24] G. Hofstede, *Culture's Consequences: International Differences in Work-Related Values*. SAGE, 1984.

[25] H. C. Triandis, "The psychological measurement of cultural syndromes," *American Psychologist*, vol. 51, no. 4, pp. 407–415, 1996, doi: 10.1037/0003-066X.51.4.407.

[26] M. L. Frazier, P. D. Johnson, and S. Fainshmidt, "Development and validation of a propensity to trust scale," *Journal of Trust Research*, vol. 3, no. 2, pp. 76–97, Oct. 2013, doi: 10.1080/21515581.2013.820026.

[27] D. H. McKnight, M. Carter, J. B. Thatcher, and P. F. Clay, "Trust in a specific technology: An investigation of its components and measures," *ACM Trans. Manage. Inf. Syst.*, vol. 2, no. 2, p. 12:1-12:25, Jul. 2011, doi: 10.1145/1985347.1985353.

[28] G. M. Alarcon, A. M. Gibson, S. A. Jessup, and A. Capiola, "Exploring the differential effects of trust violations in human-human and human-robot interactions," *Applied Ergonomics*, vol. 93, p. 103350, May 2021, doi: 10.1016/j.apergo.2020.103350.

[29] G. M. Lucas, B. Becerik-Gerber, and S. C. Roll, “Calibrating workers’ trust in intelligent automated systems,” *Patterns*, vol. 5, no. 9, p. 101045, Sep. 2024, doi: 10.1016/j.patter.2024.101045.

[30] P. L. McDermott and R. N. ten Brink, “Practical Guidance for Evaluating Calibrated Trust,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 63, no. 1, pp. 362–366, Nov. 2019, doi: 10.1177/1071181319631379.

[31] V. L. Pop, A. Shrewsbury, and F. T. Durso, “Individual Differences in the Calibration of Trust in Automation,” *Hum Factors*, vol. 57, no. 4, pp. 545–556, Jun. 2015, doi: 10.1177/0018720814564422.

[32] I. Dergaa *et al.*, “From tools to threats: a reflection on the impact of artificial-intelligence chatbots on cognitive health,” *Front. Psychol.*, vol. 15, Apr. 2024, doi: 10.3389/fpsyg.2024.1259845.

[33] S. Triberti, R. Di Fuccio, C. Scuotto, E. Marsico, and P. Limone, “‘Better than my professor?’ How to develop artificial intelligence tools for higher education,” *Front. Artif. Intell.*, vol. 7, Apr. 2024, doi: 10.3389/frai.2024.1329605.

[34] S. Marocco, A. Talamo, and F. Quintiliani, “From service design thinking to the third generation of activity theory: a new model for designing AI-based decision-support systems,” *Front. Artif. Intell.*, vol. 7, Mar. 2024, doi: 10.3389/frai.2024.1303691.

[35] R. Dara, S. M. Hazrati Fard, and J. Kaur, “Recommendations for ethical and responsible use of artificial intelligence in digital agriculture,” *Front. Artif. Intell.*, vol. 5, Jul. 2022, doi: 10.3389/frai.2022.884192.

[36] L. Bainbridge, “Ironies of automation,” *Automatica*, vol. 19, no. 6, pp. 775–779, Nov. 1983, doi: 10.1016/0005-1098(83)90046-8.