sciendo

Scientific Bulletin
Vol. XXVIII, No. 2(56), 2023

ALGORITHMIC ACCURACY IN CONTEXT: ASSESSING THE IMPACT OF FINANCIALLY MOTIVATED DISHONESTY

Jennifer BATES
jbates@marymount.edu

Amel BEN ABDESSLEM
abenabde@marymount.edu

SCHOOL OF BUSINESS, MARYMOUNT UNIVERSITY, ARLINGTON, VIRGINIA

ABSTRACT:

Algorithms are expected to radically improve the way businesses operate. To fully realize these benefits, customers have to accept the algorithmic implementations. Accuracy is a critical component of algorithmic acceptance models. However, thus far, research into accuracy has been limited to user ratings of the accuracy of the algorithmic output despite strong evidence for customer (user) acts of embellishment and other moral hazards. This article aims to deepen the understanding of accuracy ratings by assessing the impact of variation in user input accuracy on the accuracy ratings of the algorithm's ultimate response. Drawing on honesty, a survey was presented to 500 individuals on the Cloud Research platform. The quantitative analysis showed that the more inaccurately a user felt and behaved, the more accurately they rated the algorithmic response of 'decline', with contextual accuracy factors explaining up to 35% of the variation in ratings of the algorithm's accuracy. This finding raises important implications for researchers and practitioners who want to improve algorithmic acceptance but may have limited their improvement focus to explainability or model accuracy without considering the user behavior. Algorithmic accuracy ratings and acceptance should be evaluated in the context of the user and their likelihood to provide inaccurate inputs.

KEYWORDS:

Accuracy, algorithmic acceptance, honesty, perception

1. Introduction

Algorithms are being incorporated into a broad range of business applications such as healthcare, finance, and insurance (Willson, 2017), but adoption rates have been affected by user aversion toward algorithmic decision-making (Mahmud et al., 2022). This has propelled an interest in understanding the factors that may increase adoption rates. In scholarly literature, qualitative measurement of algorithmic accuracy has been accomplished via questions such as "How much do you trust that the algorithm made a good-quality decision?" (Lee, 2018). Yet, the nature of these questions focuses on user perceptions of accuracy, particularly when there is no objectively correct answer to supplement with.

While receiving less focus in algorithmic literature, the accuracy of inputs has been a necessary concern for insurance and fintech businesses. Fraud in business dealings is generally categorized into first, second, and third-party fraud,

DOI: 10.2478/bsaft-2023-0012
© 2023 Jennifer Bates et al. This work is licensed under the Creative Commons Attribution-Non Commercial-No Derivatives 3.0 License.

which are all purposeful and planned fraud activities. However, there is general acceptance that the fourth category, known alternatingly as opportunistic fraud, consumer opportunism, or moral hazard, is of significant importance due to its prevalence and impact, with as many as 20% of loan applications showing inflated incomes, contributing to losses of up to 5% of business revenue (Cohn, 2020; Garmaise, 2015). Despite this, user accuracy is not addressed by algorithmic research, which focuses solely on the accuracy of the algorithmic outputs.

This research intends to close that gap in through investigating how variation in the accuracy of user inputs affects their ratings of algorithmic outputs. In order to investigate this research question, a deception scenario was developed. Its purpose was to recreate a realistic financial application in order to encourage realistic opportunistic fraud behaviors. The accuracy of the survey respondents' answers was measured via both direct and indirect methods, and existing scales were used to measure their ratings of algorithmic accuracy.

The results of this study will benefit researchers and practitioners in several ways. When algorithms are deployed in the real world, it is guaranteed that inputs provided to them will have varying levels of accuracy. It is essential that the degree to which a user may provide inaccurate information be accounted for when designing and optimizing algorithms. To encourage algorithmic acceptance, consideration and effort need to be divided, with a focus not just on improvements to transparency and explainability but also on encouraging the user to provide accurate information.

The remainder of this article is structured as follows. Section 2 reviews the current understanding of the importance of accuracy to algorithmic acceptance as well as the understanding of consumer honesty in business scenarios. Section 3 outlines the theoretical background and hypothesis development. Next, Section 4 outlines the research methodology, survey design, and data analysis techniques. Section 5 presents the results and Section 6 discusses their implications. Finally, Section 7 discusses limitations and suggestions for future research.

2. Literature review

2.1. Algorithmic acceptance model

As users become increasingly aware of the existence of algorithms, pressure is placed on ensuring users also understand and accept the algorithms. The Algorithmic Acceptance Model (AAM) was developed on the backbone of UX research, including influences from Expectation Confirmation Theory (ECT) and Adaptive Decision-Making Theory (ADM) (Shin, 2020a; Payne et al., 1993). Taken together and applied to Algorithmic Experiences (AX), these theories imply that user acceptance of algorithms will be evaluated subjectively based on the criteria of the specific scenario, and this initial evaluation will strongly affect their perceptions of the algorithmic performance (confirmation bias) which will, in turn, affect their intention to use the algorithmic process further.

Models for evaluating algorithmic acceptance have noted the importance of the perceptions of satisfaction, fairness, and trust from the onset. Shin & Park (2019) formed the initial AAM model, evaluating algorithms on fairness (of which accuracy is a subset), accountability, transparency, and trust. Empirical studies have evaluated the AAM, and have shown it to have strong correlations to user heuristics, though explainability was added in 2020 to better quantify trust's effects as a moderator (König et al., 2022; Shin, 2021; Shin & Park, 2019). While each factor is consistently mentioned across AX research, in some instances, the factors of explainability and accuracy are bucketed within transparency and fairness, respectively. Despite including contextual factors in ECT and ADM, they are not currently included in the AAM.

2.2. Consumer honesty

Currently, there is no comprehensive list of motivations for consumers to lie. Fudge Factor Theory (FFT) posits a non-zero limit for human beings to be dishonest without feeling as though they are immoral or dishonest themselves (Andrighetto et al., 2016). The General Theory of Deception (GTD), though focused mainly on the creation of lies, does outline "BPRED" factors (benefit, punishment, risk, execution, and dissonance), which represent an individual's subjective assessment of a scenario and whether to be truthful (Srour & Py, 2022). These academic theories line up with industry research showing consumer honesty levels are impacted by both financial factors, such as access to better loan terms or higher insurance payouts, as well as non-economic factors, such as social good or feelings of rejection (Tseng, 2019; Van Der Zee et al., 2016). Evidence has shown that corporations can employ BPRED-related techniques such as a reduction of anonymity, task instructions focused on truthfulness, and encouraging feelings of loyalty towards the corporation to encourage more truthful behavior (Hydock, 2018; Leal et al., 2016).

3. Theoretical background

3.1. Role of cognition

A review of FFT and GDT highlights that dishonest acts can be categorized not merely as honest/dishonest but also on factors of awareness, intent, and size of resulting action (Srour & Py, 2022). Self-concept maintenance theory, when applied to honesty, posits that individuals desire to maintain a self-image that they are honest (Mazar et al., 2008). The desire to maintain a self-image of honesty can be so strong as to induce those who lie to forget they told the lie altogether, particularly when they are highly involved with producing the inaccurate information (Y. Li & Liu, 2021). Taking into consideration the role of cognition in dishonest behavior, particularly in generating more significant dishonest behaviors, the first hypothesis is proposed:

H1: *Individuals who engage in small lies are less likely to report dishonest behavior or consider themselves dishonest.*

3.2. Role of reciprocity

Studies have indicated that reciprocity creates an expectation that honest and accurate interactions will be returned with honesty and accuracy from others (K. Li & Xiao, 2022). With the rise of algorithmic interest, research has posited that reciprocity will also be a significant component of human-algorithm interactions (Upadhyaya & Galizzi, 2023; Sandoval et al., 2016). Some experimental studies have shown that individuals will similarly interact with algorithms as with other humans (Sandoval et al., 2016), while others found that humans were more likely to have reduced reciprocity, taking advantage of the algorithm's expected behavior for their gain (Upadhyaya & Galizzi, 2023). Together, this points to individual actors' assessment of algorithms to be influenced by their behavior in given interactions.

H2: *Contextual factors will affect ratings of algorithmic accuracy.*

H2a: *Individuals who engage in minor inaccuracies or are entirely accurate will rate algorithmic responses as more accurate.*

H2b: *Individuals who are meaningfully inaccurate or dishonest will rate algorithmic responses as less accurate.*

4. Methodology

4.1. Deception scenario

Consider the question, "What is the approximate value of your car", a survey respondent in a neutral environment would likely provide an honest estimate, in contrast, if asked as part of an insurance claim, the respondent encounters a moral hazard, encouraging inflated responses (Champiri et al., 2020; Tseng, 2019). This led to a two-part survey design in which respondents completed a single study task

broken up into two sections. The first section was a deception scenario designed to capture answers while operating in a context with a moral hazard, creating a possible and realistic incentive for answering inaccurately. The respondent was told they were providing inputs to an application for a financial product they strongly desired. Before the second section, the survey respondent was notified that the algorithm declined them for a desired financial product when, in reality, no algorithm existed, and everyone was declined. The respondent then answered questions about their perception of the algorithm's result, as well as a question set designed to understand the accuracy level of the answers they provided, their self-perceived honesty, and basic demographic information.

4.2. Honesty assessments

The literature review discussed the distinction between planned fraud, delineated into first-, second-, and third-party frauds, and opportunistic or unplanned fraud, and the distinctions those might cause in the self-awareness of dishonest behavior. With that framework in mind, the survey design leveraged two honesty assessments for the independent variables. Respondents were first assessed using self-reported Likert-scale responses on generalized honesty scales (Ramokapane et al., 2021; Henle et al., 2019). Respondents who scored at least 31 out of 32 on the honesty scale were labeled as 'Honest' (N = 325), while respondents with self-reported scores of less than 31 were labeled 'Claimed Dishonest' (N = 171). The second honesty measure was created by counting the differences in answers to the financial application questions when motivated by the mortgage scenario compared to a neutral survey environment. Respondents whose answers to the Likert-scaled questions varied by one or less were labeled as 'Honest' (N = 306), while respondents with larger variation were labeled as 'Acted Dishonest' (N = 190). Respondents who both claimed to be dishonest and acted dishonest (N = 91) were grouped into 'Claimed and Acted Dishonest' for the analysis.

4.3. Algorithmic scales

Ratings of the algorithm's fairness, accuracy, and explainability were dependent continuous variables. The survey questions asked participants for a level of agreement using a standard five-point or seven-point Likert-scale. The dependent variable was represented by totaling the scores on five questions, with three subtypes of fairness, accuracy, and explainability (Shin, 2020b; Lee, 2018). Algorithmic awareness was calculated by totaling the number of correct responses to a series of questions about algorithms, including what they are and where they are used (Dogruel et al., 2021). Respondents were also asked for Likert-scale agreement to three questions as possible moderators: one assessing whether they thought the algorithm would know if inaccurate answers were provided, a second assessing whether they believed the algorithm used other information, and a third to determine if they believed their inputs were important to the algorithm.

4.4. Descriptives

Data was collected via a Qualtrics survey from 500 individuals recruited via Cloud Research. Four participants were removed due to not finishing the survey. For data analyses, 496 responses were analyzed. The Cloud Research tool enforced an even split for male and female respondents and required that users be located in the United States when taking the survey. The age of participants ranged from 18 to 84, with average age being in the range from 35 - 44 years old. Participants ranged from having not completed high school to possessing a doctorate. Regarding employment, 60% of respondents were employed full-time, 11.3% were self-employed, and 3.4% were unemployed. 74.4% of participants identified as White, 10.1% as Black, 7.8% as Hispanic, 6.6% as Asian, and the remainder as Other.

5. Results

5.1. Hypothesis 1 – Honesty types

Hypothesis 1 was assessed based on a comparison of the results of the honesty measures. An Independent Samples One Way Analysis of Variance (ANOVA) was performed to evaluate within-group differences of those assessed as having Acted Dishonestly. A one-way between-groups ANOVA found that, of those who acted dishonestly, the degree of dishonesty varied significantly between the two groups who did/did not self-report (claim) dishonesty, ($F(1,188) = 27.5$, $p < .001$, $\eta^2 = 0.127$). Unadjusted post-hoc comparisons revealed significant differences in the degree of dishonesty between those who both Claimed and Acted Dishonestly ($M = 5.38$, $SD = 3.55$) and those who claimed to be honest ($M = 3.28$, $SD = 1.74$, $p < .001$, $d = .761$, 95% CI [0.464, 1.06]). This supports H1, indicating that respondents committing small acts of dishonesty may still view themselves as honest individuals.

5.2. Hypothesis 2 – Contextual factors

5.2.1. - Contextual influence of acting dishonest

A multiple linear regression analysis tested how ratings of the level of influence one's inputs have on outcome, whether the algorithm would know of lies, the amount the algorithm used outside data, and the degree of dishonesty affected ratings of the algorithm's accuracy. The combination of predictors significantly predicted the algorithmic accuracy, explaining 35% of the variance ($F(4,491) = 65.4$, $p < .001$, adjusted $R^2 = .348$). An algorithm's accuracy rating was significantly impacted by how much an algorithm's output was affected by the user's inputs ($\beta = .45$, $t(97) = 11.33$, $SE = .06$, $p < .001$), the ratings of whether the algorithm used other inputs ($\beta = -.24$, $t(97) = -6.07$, $SE = .05$, $p < .001$), and the perception that the algorithm would know if someone lied, ($\beta = .10$, $t(97) = 2.57$, $SE = .05$, $p < .01$). Whether the respondent acted dishonestly was not significant as an independent factor. The results for each predictor within the regression model are shown in Table no. 1. Increased agreement with the respondent's inputs affecting the algorithm's outputs significantly correlated with increased algorithmic accuracy ratings, which support H2, indicating that an individual's actions while interacting with the algorithm will have a significant effect on their ratings of its accuracy.

Table no. 1

Standard Multiple Regression Results for Honest Actions and Impact Variables

		95% Confidence Interval for *β*			
	β	**LL**	**UL**	***B***	***p***
Inputs Affect Outcome	0.45	0.37	0.52	0.69	< .001
Algorithm Knows Lies	0.10	0.02	0.17	0.14	0.010
Use of Other Info	-0.24	-0.32	-0.16	-0.33	< .001
Acted Dishonest	0.07	0.00	0.14	0.25	0.059

5.2.2. - Contextual influence of claimed dishonest

An additional multiple linear regression tested how ratings of whether the algorithm would know of lies and the degree to which individuals view themselves as dishonest affected ratings of the algorithm's accuracy. The combination of predictors significantly predicted the algorithmic accuracy, explaining 30% of the variance ($F(2,493) = 102$, $p < .001$, adjusted $R^2 = .290$). An algorithm's

accuracy rating was significantly affected by both whether the algorithm would know of lies (β = -.08, t(97) = -2.12, SE = .01, $p < .05$) and the degree to which an individual views themselves as honest (β = .53, t(97) = 14.05, SE = .06, $p < .001$).

The results for each predictor within the regression model are shown in Table no. 2. Increased agreement with the respondent's inputs affecting the algorithm's outputs significantly correlated with increased algorithmic accuracy ratings. In contrast, respondents who believed themselves to be honest rated the algorithm as less accurate. These findings further support H2.

Table no. 2

Standard Multiple Regression Results for Self-Rated Honesty and Impact Variables

		95% Confidence Interval for *β*			
	β	**LL**	**UL**	***B***	***p***
Inputs Affect Outcome	0.53	0.46	0.61	0.82	< .001
Claimed Dishonest	-0.08	-0.16	-0.01	-0.03	0.034

5.2.3. Contextual influence of honesty type

An ANOVA analysis was used to calculate if the differences in dishonesty type created any difference in responses to algorithmic accuracy. A one-way between-groups ANOVA found that ratings of algorithmic accuracy differed significantly across honesty types (F(3,492) =3.15, $p < .05$, $\eta^2 = 0.019$). Unadjusted post-hoc comparisons revealed significant differences in accuracy ratings, with honest individuals providing significantly lower ratings (M = 3.84, SD = 1.84) than both those who both Claimed and Acted Dishonest (M = 4.43, SD = 1.48, $p < .01$, d = .343, 95% CI [0.10, 0.59]) and those who only Claimed Dishonest (M = 4.3, SD = 1.56, $p < .05$, d = .268, 95% CI [0.01, 0.52]). The estimated marginal means are presented in Figure no. 1. Remembering that the algorithmic outcome was 'Decline' for all participants, these findings support H2a and H2b. Specifically, respondents with the least honest and accurate behavior gave the highest algorithmic accuracy ratings.



Figure no. 1: Estimated Marginal Means by Honesty Type
(Note: Error bars represent 95% confidence intervals)

6. Discussion and implications

This research aimed to determine whether the contextual factors present in ECT and ADM should be included in future studies focused on algorithmic acceptance. The study findings strongly suggest that contextual factors, particularly those of a user's honesty, affect their perception of the algorithm and its accuracy. H2a and H2b were supported, showing that the accuracy of a negative outcome was rated more highly by those who were dishonest. This finding is in line with more generalized honesty and reciprocity work, which indicates that individuals create an expectation that their actions will be reflected by non-human actors (Sandoval et al., 2016).

6.1. Theoretical implications

This article contributes several important theoretical considerations to the literature on algorithmic acceptance, specifically concerning accuracy. Though a large body of work has examined methods to improve algorithmic acceptance, it often hinges on improving transparency and explainability (König et al., 2022), with accuracy often discussed as a trade-off of explainability (Herm et al., 2023). This study advances our understanding to show that it may be possible to improve accuracy without trade-offs via improvements to contextual factors that influence the user's behaviors.

This study supports the inclusion of the contextual factors present in ECT and ADM into the AAM. Rather than focus exclusively on how users rate accuracy, it should be assessed as a function of the user and their environment. In particular, the use of generalized honesty scales can be leveraged in future research to better explain variances in algorithmic accuracy assessments, particularly in research environments where the user's actual accuracy may be difficult to determine due to preferred anonymity and privacy concerns.

6.2. Practical implications

Whether via data mining, selling, or direct user entry, misrepresented data is likely to wind up in an algorithm used by businesses. Different business stakeholder groups may apply this knowledge in different ways. *Model developers* should consider the context in which training data was supplied or collected to interpret accuracy ratings better. *UX designers* should consider using honesty prompting to collect better information and improve user adoption of algorithms (Srour & Py, 2023; Leal et al., 2016). *Organizational leaders* considering algorithmic implementations should consider factors such as their industry and the context of the model's usage as a requirement for making deployment decisions. Additionally, there are scenarios where there are regulatory or practical limits to the BPRED factors a business can implement to improve user accuracy despite known concerns or instances of moral hazard. This may create an upper limit to user accuracy ratings of any algorithms deployed in this space and should be considered before deploying developer resources toward accuracy improvements.

6.3. Limitations and avenues for future research

This research has several possible limitations as it is based on a single scenario given to survey respondents in the United States. The behaviors and perceptions of individuals concerning both honesty (Andrighetto et al., 2016) as well as algorithmic perceptions (Grzymek et al., 2019; Nitto et al., 2017) have been shown to vary across countries and cultures, indicating the research should be repeated with non-US based respondents. Furthermore, two limitations were introduced due to using a survey hosting platform. The first is that all respondents were given the same algorithmic 'outcome,' a mortgage decline. An interesting avenue of future research would be to vary the

categorical result or use a numerical result that varies in size/degree. The second limitation imposed by a survey platform and goals of data security is that the respondents were anonymous. It is both possible and likely that some users may report being honest and give the same answers to both question sets, causing them to be labeled as acting honestly despite being dishonest in one or both scenarios. In real-world scenarios, there are external data sources, such as W-2s and FICO scores, which could have been used in a mortgage scenario to verify individuals who acted honestly more accurately.

7. Conclusion

With the growing business desire to implement algorithms in various industries and contexts, it's increasingly important to understand how those contexts may influence algorithmic adoptions. This article set out to close the gap in understanding how these contexts may influence algorithmic adoption, beginning with accuracy. It showed that up to 30% of the variance in user accuracy ratings may be attributed to feelings or actions of inaccuracy. As illustrated in this study, future research should be wary of only asking users about their perceptions of the algorithmic accuracy without inquiring about the other factors at play.

CRediT authorship contribution statement

Jennifer Bates: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Visualization. **Amel Ben Abdesslem:** Resources, Writing - Review & Editing, Supervision.

REFERENCES

Andrighetto, G., Zhang, N., Ottone, S., Ponzano, F., D'Attoma, J., & Steinmo, S. (2016). Are Some Countries More Honest than Others? Evidence from a Tax Compliance Experiment in Sweden and Italy. *Frontiers in Psychology, Vol. 7*. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00472.

Champiri, Z.D., Fisher, B., Kiong, L.C., & Danaee, M. (2020). How contextual data influences user experience with scholarly recommender systems: An empirical framework. In *HCI International 2020 – Late Breaking Papers: User Experience Design and Case Studies*, 635-661. Springer International Publishing.

Cohn, M. (2020, April 17). Organizations lose 5 percent of revenue to fraud every year. *Accounting Today,* available at: https://www.accountingtoday.com/news/organizations-lose-5-of-revenue-to-fraud-every-year.

Dogruel, L., Masur, P., & Joeckel, S. (2021). Development and validation of an algorithm literacy scale for internet users. *Communication Methods and Measures, Vol. 16, Issue 2,* 115-133. Available at: https://doi.org/10.1080/19312458.2021.1968361.

Garmaise, M.J. (2015). Borrower misreporting and loan performance. *The Journal of Finance, Vol. 70, Issue 1*, 449-484. Available at: https://doi.org/10.1111/jofi.12156.

Grzymek, V., Puntschuh, M., & Bertelsmann Stiftung. (2019). What Europe knows and thinks about algorithms: Results of a representative survey. *Discussion Paper Ethics of Algorithms,* available at: https://doi.org/10.11586/2019008.

Henle, C.A., Dineen, B.R., & Duffy, M.K. (2019). Assessing intentional resume deception: Development and nomological network of a resume fraud measure. *Journal of Business and Psychology, Vol. 34, Issue 1*, 87-106. Available at: https://doi.org/10.1007/s10869-017-9527-4.

Herm, L.-V., Heinrich, K., Wanner, J., & Janiesch, C. (2023). Stop ordering machine learning algorithms by their explainability! A user-centered investigation of performance and explainability. *International Journal of Information Management, Vol. 69*, 102538. Available at: https://doi.org/10.1016/j.ijinfomgt.2022.102538.

Hydock, C. (2018). Assessing and overcoming participant dishonesty in online data collection. *Behavior Research Methods, Vol. 50, Issue 4*, 1563-1567. Available at: https://doi.org/10.3758/s13428-017-0984-5.

König, P.D., Wurster, S., & Siewert, M.B. (2022). Consumers are willing to pay a price for explainable, but not for green AI. Evidence from a choice-based conjoint analysis. Big *Data & Society, Vol. 9, Issue 1*. Available at: https://doi.org/10.1177/ 20539517211069632.

Leal, S., Vrij, A., Nahari, G., & Mann, S. (2016). Please be honest and provide Evidence: Deterrents of deception in an online insurance fraud context. *Applied Cognitive Psychology, Vol. 30, Issue 5,* 768-774. Available at: https://doi.org/10.1002/acp.3252.

Lee, M.K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society, Vol. 5, Issue 1*. Available at: https://doi.org/10.1177/2053951718756684.

Li, K., & Xiao, W. (2022). Who will help to strive against the "infodemic"? Reciprocity norms enforce the information sharing accuracy of the individuals. *Frontiers in Psychology, Vol. 13*. Available at: https://doi.org/10.3389/fpsyg.2022.919321.

Li, Y., & Liu, Z. (2021). Involvement modulates the effects of deception on memory in daily life. *Frontiers in Psychology, Vol. 12. A*vailable at: https://www.frontiersin.org/ articles/10.3389/fpsyg.2021.756297.

Mahmud, H., Islam, A.K.M.N., Ahmed, S.I., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change, Vol. 175*. Available at: https://doi.org/10.1016/j.techfore.2021.121390.

Mazar, N., Amir, O., & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research (JMR), Vol. 45, Issue 6,* 633-644. Available at: https://doi.org/10.1509/jmkr.45.6.633.

Nitto, H., Taniyama, D., & Inagaki, H. (2017). Social acceptance and impact of robots and artificial intelligence: Findings of survey in Japan, the U.S. and Germany. *NRI Papers, Vol. 211*, 1-17.

Payne, J.W., Bettman, J.R., & Johnson, E.J. (1993). The adaptive decision maker (pp. xiii, 330). Cambridge University Press. Available at: https://doi.org/10.1017/ CBO9781139173933.

Ramokapane, K.M., Misra, G., Such, J., & Preibusch, S. (2021). Truth or dare: Understanding and predicting how users lie and provide untruthful data online. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1-15.* Available at: https://doi.org/10.1145/3411764.3445625.

Sandoval, E.B., Brandstetter, J., Obaid, M., & Bartneck, C. (2016). Reciprocity in human-robot interaction: A quantitative approach through the prisoner's dilemma and the ultimatum game. *International Journal of Social Robotics, Vol. 8, Issue 2,* 303-317. Available at: https://doi.org/10.1007/s12369-015-0323-x.

Shin, D. (2020a). How do users interact with algorithm recommender systems? The interaction of users, algorithms, and performance. *Computers in Human Behavior, 109*. Available at: https://doi.org/10.1016/j.chb.2020.106344.

Shin, D. (2020b). User perceptions of algorithmic decisions in the personalized AI system: Perceptual evaluation of fairness, accountability, transparency, and explainability. *Journal of Broadcasting & Electronic Media, Vol. 64, Issue 4,* 541-565. Available at: https://doi.org/10.1080/08838151.2020.1843357.

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies, Vol. 146*. Available at: https://doi.org/10.1016/j.ijhcs.2020.102551.

Shin, D., & Park, Y.J. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior, Vol. 98*, 277-284. Available at: https://doi.org/10.1016/j.chb.2019.04.019.

Srour, C., & Py, J. (2022). The general theory of deception: A disruptive theory of lie production, prevention, and detection. *Psychological Review*. Available at: https://doi.org/10.1037/rev0000389.

Srour, C., & Py, J. (2023). Making job candidates more honest with a theory-based form: A field study. *Journal of Personnel Psychology, Vol. 22, Issue 2*, 89-98. Available at: https://doi.org/10.1027/1866-5888/a000313.

Tseng, L.-M. (2019). Customer insurance frauds: The influence of fraud type, moral intensity and fairness perception. *Managerial Finance, Vol. 45, Issue 3,* 452-467. Available at: http://dx.doi.org/10.1108/MF-04-2018-0162.

Upadhyaya, N., & Galizzi, M. (2023). In bot we trust? Personality traits and reciprocity in human-bot trust games. *Frontiers in Behavioral Economics, Vol. 2*. Available at: https://doi.org/10.3389/frbhe.2023.1164259.

Van Der Zee, S., Anderson, R., & Poppe, R. (2016). When Lying Feels the Right Thing to Do. *Frontiers in Psychology, Vol. 7*. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00734.

Willson, M. (2017). Algorithms (and the) everyday. *Information, Communication & Society, Vol. 20, Issue 1,* 137-150. Available at: https://doi.org/10.1080/1369118X.2016.1200645.