Keep talking and nobody decides

How can AI augment users' ability to detect misinformation while balancing engagement and workload?

Maged Mortaga[1,*,†], Marvin Sieger[1,†], Lilian Kojan[1], Hendrik Nunner[1], Leonard Stellbrink[1], André Calero Valdez[1] and Tim Schrills[1]

[1] *University of Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany*

Abstract

To detect misinformation, users of social networks potentially utilize AI-based decision support systems (DSS). However, a DSS's ability to augment user behavior depends on how a DSS modifies users' decision-making and interaction experience. We examined how users' performance and experience are affected by the level of automation of a DSS in misinformation detection. In a preregistered within-subjects-experiment with an AI, N=99 participants interacted with two DSS in a simulated environment. The first provided distinct recommendations (higher level of automation), while the second provided solely evaluative support (lower level of automation). We compared their effect on user behavior (here: accuracy, interaction frequency) and experience (here: trust, traceability). Participants showed higher accuracy when receiving recommendations but also interacted less frequently. Trust and perceived traceability did not differ between systems. We discuss whether more intensive processing of the evaluated information could be responsible for the higher number of errors in the evaluative system.

Keywords

Misinformation Detection, Decision Support Systems, Levels of Automation, Recommender System, XAI, HCAI, Trust, Social Media

1. Introduction

Half of all users on social networks use social media as their main source of news [1, 2, 3]. Meanwhile, misinformation on social networks increasingly influences how attitudes and subsequent decisions (e.g., regarding voting) [4] are formed. Previous research demonstrated that misinformation can lead to polarization [5, 6] and echo chambers [7].

Developing technology to inhibit the negative influence of misinformation in social networks constitutes a critical research agenda. One approach has been to reduce the spread of misinformation through automated (AI-based) systems that flag or remove specific posts and accounts from social media platforms (e.g., [8, 9, 10, 11, 12]).

However, automatically flagging content is only a valid strategy when a distinction between truthful and misleading information is possible. As stated, for example, by Hameleers [13] and Vraga [14], misinformation is not one side of a dichotomy contrasted by information on the other: the relation of a piece of information to what we perceive as truthful is strongly dependent on context and time. It evolves, continuously changing until (if possible) a stable point, e.g., a ground truth or its status as a hoax is established [15].

Rather than fully automating the detection and removal process, DSS can augment individuals' ability to correctly classify misinformation; that is, by providing context to the perceived content or highlighting other indicators of the information being misinformation (e.g., [9, 10, 11, 12, 16]). Switching

Joint Proceedings of the ACM IUI Workshops 2025, March 24–27, 2025, Cagliari, Italy

*Corresponding author.

† These authors contributed equally.

maged.mortaga@uni-luebeck.de (M. Mortaga); ma.sieger@uni-luebeck.de (M. Sieger); lilian.kojan@uni-luebeck.de (L. Kojan); hendrik.nunner@uni-luebeck.de (H. Nunner); leonard.stellbrink@uni-luebeck.de (L. Stellbrink); andre.calerovaldez@uni-luebeck.de (A. C. Valdez); tim.schrills@uni-luebeck.de (T. Schrills)

0009-0002-3683-3621 (M. Mortaga); 0009-0001-8365-8267 (M. Sieger); 0000-0002-4064-4447 (L. Kojan); 0000-0002-4083-3651 (H. Nunner); 0000-0003-2188-1743 (L. Stellbrink); 0000-0002-6214-1461 (A. C. Valdez); 0000-0001-7685-1598 (T. Schrills)

from completely automated detection and removal of posts to an augmentation-oriented approach changes the level of automation [17], i.e., leads to a reduced level of automation. However, lower levels of automation may influence how users experience their interaction with the system and, in turn, how they are able to perform in their task, i.e., misinformation detection.

Potential psychological consequences concern the users' sense of autonomy, the experienced workload, and their dependency on the system [18, 19]. For example, a reduction of automation can result in a higher workload for the user [20] as users have to manually check every post. On the other hand, a support system that achieves higher levels of automation greatly reduces workload, but might increase the risk of decreased user autonomy or over-reliance on the system [21, 18]. This may lead to users rejecting the system after experiencing failure [22] or discarding it altogether. We argue that experimental research is needed to understand the effects of automation levels in misinformation detection support on a performance level, as well as effects on user experience. Consequently, the objective of this study was to implement different DSS with different levels of automation and experimentally investigate the impact of these differences.

2. Background / Related work

DSS can provide context, highlight patterns, or flag indicators of misinformation to support users in the classification process. While augmenting information, the level of automation of a DSS may affect user experience, e.g., perceived workload, trust, and traceability of the system [17, 18]. Understanding the trade-offs between varying levels of automation is critical to designing DSS that balance user autonomy with effective misinformation detection. This balance depends on automation-based effects on user experience and performance.

High levels of automation are not generally preferable, as unwarranted trust, skill degradation, and diminished situational awareness may contribute to reduced user engagement and complacency [17]; similar concerns have been highlighted in [23]. In turn, reducing automation in misinformation detection systems can enhance human-AI integration by fostering user interaction, particularly in the decision-making stage (see [18]).

Human involvement is especially important as attributes of information such as tone, context, and perceived truthfulness [15] influence what can be perceived as truthful, a differentiation which cannot sufficiently be made by automated systems. Thus, limiting automation to evaluative support rather than fixed, fully automated recommendations or removal may be desirable. Stronger interaction with the content may further lead to more informed decision-making, as postulated in the Elaboration-Likelihood-Model (ELM) [24].

Increased system interaction during decision-making may also improve the perception of system traceability, ensuring users understand system processes and outcomes [25]. While this might foster trust in the system, processing this additional information may also increase workload [26]. The balance between ensuring transparency to foster user engagement and achieving high levels of automation to reduce workload is delicate and requires empirical examination. Here, it is crucial to see how the technological environment conditions affect this balance and not, as previous studies tend to do, to look at the differences in users [27].

Systems that lack transparency and do not allow the user to interact with them, risk losing user trust, which may lead to lower levels of user engagement. In contrast to that, excessively interactive systems may be more transparent for the user but reduce usability due to cognitive overload [28]. For misinformation detection, higher transparency without overwhelming the user is crucial to maintaining trust and supporting nuanced decision-making.

Finally, the ELM emphasizes that argument quality and information literacy significantly influence users' ability to identify false information. Central route processing, requiring engagement with high-quality arguments, mitigates susceptibility to misinformation [29]. Peripheral cues, such as the credibility of a source, can also influence perception, but are less resistant to manipulation. As demonstrated in empirical studies, users with higher information literacy are less likely to be deceived, particularly on structured platforms like news websites [29]. Based on that, DSS in misinformation detection should aim for a level of automation that supports central route processing. To the best of our knowledge, the psychological consequences of low automation in DSS for misinformation detection have not been studied experimentally before.

3. Present Research

Building on prior research, this study investigates how the type of automation in DSS influences users' performance and experience in detecting misinformation. Specifically, we explore how different levels of automation—recommendation-based (System R) as opposed to evaluation-based (System E)—affect accuracy, interaction frequency, trust, and subjective information processing awareness (SIPA). We hypothesize that:

- **H1:** There is a difference in users' accuracy in detecting misinformation between System E and System R.
- **H2:** There is a difference in the number of user messages between System E and System R.
- **H3:** Users' trust will differ between System E and System R.
- **H4:** Users' SIPA varies depending on which system they interact with.

Previous studies [21, 30] have shown that trust in automated systems plays a crucial role in determining how users interact with such tools (**H3**). We aim to advance the development of human-centered AI by using a controlled experiment to explore how varying levels of automation in DSS can shape not only users' performance (**H1, H2**), but also their automation-related user experience (**H3, H4**).

4. Methods

For our study, we developed a web application simulating a social media platform. Participants interacted with two AI-based DSS systems—System E and System R—to classify information as misinformation or accurate information. Both AI-based systems used OpenAI's Assistant API [1] (*gpt-3.5-turbo*). They had identical interfaces (see Figure 1), differing only in their initial message. For System E, the message was: *"Hello! There are many different opinions and pieces of information about the post shown above. How can I help you?"*. For System R, the message was: *"Hello! The post shown above (may/probably/definitely) (does not) contain misinformation. How can I help you?"*. The exact message for System R depended on the shown post, but was consistent for each participant viewing the same post.

To identify subject areas with high levels of misinformation, we selected posts based on topics from reputable fact-checking websites such as Snopes, Correctiv, AFP, and Reuters. We created posts using information from these sources, adding false content to some to create misinformation posts. Twenty potential posts (10 real, 10 misinformation) were created, from which 10 (6 real, 4 misinformation) were selected for the study based on a consensus between the authors. The posts were originally written in German but translated for clarification (as presented in this paper and on OSF [2]).

Study participants were recruited by students enrolled in a user research course, each of whom was tasked with recruiting three participants through personal contacts and university online forums[3]. Of 108 participants recruited (May 29-June 17, 2024), 99 remained after excluding one underage and eight for survey speeding. The sample was relatively young (range: 18-59; median age = 22) and predominantly male (male: 54; female: 43; another gender: 2).

We conducted a randomized within-subject study. Participants first completed a pre-test survey to collect sociodemographic data, social media usage habits, AI experience, and affinity for technology interaction (ATI) [31]. They were then presented with both AI-based DSS in random order. On

[1] https://platform.openai.com/docs/assistants/overview (last accessed 24.02.2025)
[2] https://osf.io/fh79m/?view_only=313bc29356ad42c0b4575627c5f830b5
[3] This study was conducted in accordance with the Declaration of Helsinki. It was approved by the University of Lübeck's ethics committee.

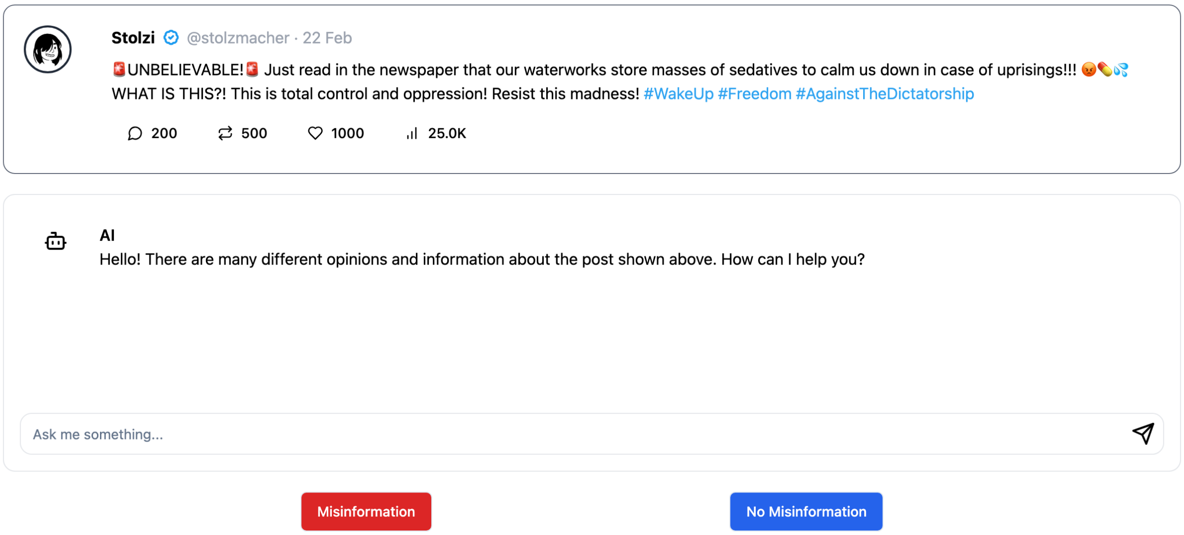

Figure 1: User interface of web application of System E. The interface shown is adjusted for height and space reasons for illustration purposes. The actual UI includes a larger chat box to display more previous chat messages.

each system, they evaluated five randomly selected posts to determine whether the posts contained misinformation, allowing us to measure their performance, i.e., their misinformation detection accuracy and the number of messages exchanged with the system. After engaging with each DSS, participants completed a post-test survey assessing their trust in the system (Human Computer Trust, HCT) [30] and the experienced traceability (subjective information processing awareness, SIPA) [32].

5. Results

The descriptive analysis revealed comparable performance between both systems across multiple measures (see Table 1). Detection accuracy was similarly high for both systems, with each participant correctly classifying at least one post. The average number of user messages per post was higher for System E, though some participants did not interact with either system at all. Trust and subjective sipa were similarly high for both systems, with satisfactory scale reliability according to McDonald's Omega.

Table 1

Performance Measures						User Experience Measures					
Measure	**Sys**	*Mdn*	*M*	*SD*	**min-max**	**Measure**	**Sys**	*Mdn*	*M*	*SD*	ω_h
Detection Accuracy	E	80%	76%	21%	20-100%	HCT	E	4.07	4.01	0.78	0.67
	R	80%	82%	19%	20-100%		R	4.13	4.10	0.84	0.81
Messages per Post	E	1.20	1.29	1.05	0.0-4.4	SIPA	E	4.00	3.75	1.01	0.79
	R	0.80	1.14	1.22	0.0-8.8		R	4.00	3.87	1.11	0.91

Note. Descriptive statistics comparing evaluative (E) and recommender (R) systems. Left: Performance measures. Right: user experience measures (HCT: human-computer trust, SIPA: subjective information processing awareness, measured on 1-6 scales). ω_h: McDonald's Omega indicating scale reliability.

Statistical analysis revealed that the differences in performance between systems were statistically significant (see Table 2). Notably, participants achieved higher detection accuracy (H1) with System R compared to System E, while showing increased message interaction (H2) with System E. There were no statistically significant differences in trust (H3) and information processing awareness (H4).

These findings suggest that the choice of AI-based DSS design meaningfully impacts both user performance and interaction patterns.

Table 2

Measure	Mean		Wilcoxon Test			
	E	R	*V*	*z*	*p*	*r*
H1: Detection Accuracy	76%	82%	569.00	-2.55	.010*	-.26
H2: Messages per Post	1.29	1.14	1580.00	2.23	.026*	.22
H3: Trust	4.01	4.10	1796.50	-1.49	.137	-.15
H4: Info. Process. Awareness	3.75	3.87	1802.00	-0.99	.324	-.10

Note. Summary of hypotheses tests comparing evaluative (E) and recommender (R) systems. (*) indicates $p < .05$.

6. Discussion

The primary goal of our research was to explore how AI-based DSS with different levels of automation affect user behavior and experience, as reflected in user accuracy, user engagement, trust, and SIPA. Supporting H1, System R led to a significantly higher accuracy than System E, showing the effectiveness of a recommendation-based approach. In line with H2, users engaged more frequently with system E, potentially encouraging users to engage more thoroughly with the post. Interestingly, H3 and H4 were not supported, as there were no significant differences between the two systems in terms of user trust and SIPA.

Our results indicate that while recommendation systems improve accuracy and reduce interaction, evaluative systems foster greater engagement without affecting users' trust or perception of information processing. In addition, we found a trade-off between the level of automation and user involvement in misinformation detection.

One explanation for the higher number of interactions in System E is the way users might process the information provided by the system. Based on the ELM, it is feasible to assume that a higher number of interactions with the system is due to a higher level of elaboration, i.e., central route processing of the provided information [29]. While engaging with more pieces of information, specifically with different arguments regarding a decision as provided by System E, can lead to a more thorough engagement with the content, it can also lead to heightened feelings of insecurity [33]. Lack of confidence in their own decision-making ability might explain the higher number of errors after engaging with System E. Recent studies also show that even if confidence is not decreased, more information can lead to poor decision-making and dissatisfaction [34]. Therefore, future research should record users' confidence in their decision-making as well as their satisfaction with their decision-making. Another aspect of central route processing is that stronger engagement with the information increases the stability of users' decisions over time [35]. Attitude stability should be tested in a follow-up survey after the initial interaction with a DSS.

Expanding on the idea that higher engagement may lead to insecurity and decision errors, the lower accuracy but higher interaction in System E could also stem from participants' attempts to justify their time investment. When users engage more extensively with a DSS, they may feel compelled to integrate their own judgment into the system's output, even when the system is highly reliable. [36] demonstrated that users working with AI-based DSS often perform less accurately than the system itself due to these strategies of combining human and machine input.

This phenomenon may reflect a psychological tendency to rationalize time spent by placing greater weight on personal contributions, potentially leading to decisions that differ from the system's suggestion. To control for different user strategies in utilizing AI-based DSS, future research should be designed to keep resource investment between systems equal. For instance, implementing a fixed delay before participants can respond may help equalize conditions and reduce the potential urge to contribute based on varying time investments.

6.1. Conclusion

Our research shows that in the dynamic field of misinformation detection, the deliberate design of automation levels and user integration plays an important role in enabling users to detect misinformation.

We were able to demonstrate that users interacted less with a recommendation-based system that had a higher level of automation than with an evaluative system. Surprisingly, participants achieved higher accuracy in detecting misinformation with a recommendation-based system. The subjective and behavioral effects may be due to a central or peripheral route of processing, as described in the ELM.

Further research should explore the users' perception of confidence when detecting misinformation, as well as the long-term effects of automation levels on user certainty after interacting with the system. In addition, more studies are needed that focus on different system designs, e.g., by modifying system reliability and task difficulty to address potential ceiling effects.

Acknowledgments

Supported by infoXpand (German Federal Ministry of Education and Research, Funding Code: 031L0300C)

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT for text revision, paraphrasing, rewording, grammar and spelling checks, as well as for generating stimuli. Additionally, GitHub Co-Pilot was utilized to support the development of the experimental environment. These tools were also employed for text translation and structuring taxonomies. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] Pew Research Center, Americans' Changing Relationship With Local News, 2024. URL: https://www.pewresearch.org/journalism/2024/05/07/americans-changing-relationship-with-local-news/.

[2] The European Parliament, Media & News Survey 2023, The European Parliament, 2023. URL: https://europa.eu/eurobarometer/surveys/detail/3153.

[3] S. Andı, How and why do consumers access news on social media, Reuters institute digital news report (2021) 2020–06. URL: https://web.archive.org/web/20240913105929/https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021/how-and-why-do-consumers-access-news-social-media.

[4] L. H. Butler, P. Lamont, D. L. Y. Wan, T. Prike, M. Nasim, B. Walker, N. Fay, U. K. H. Ecker, The (Mis)Information Game: A social media simulator., Behavior research methods (2023). URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85164456019&doi=10.3758%2fs13428-023-02153-x&partnerID=40&md5=4b44ea60ce2546414cdd9316359cc385. doi:10.3758/s13428-023-02153-x.

[5] S. van der Linden, J. Roozenbeek, J. Compton, Inoculating Against Fake News About COVID-19, Frontiers in Psychology 11 (2020). doi:10.3389/fpsyg.2020.566790.

[6] S. van der Linden, Misinformation: Susceptibility, spread, and interventions to immunize the public, Nature Medicine 28 (2022) 460–467. doi:10.1038/s41591-022-01713-6.

[7] K. O'Hara, D. Stevens, Echo Chambers and Online Radicalism: Assessing the Internet's Complicity in Violent Extremism, Policy & Internet 7 (2015) 401–422. doi:10.1002/poi3.88.

[8] N. Bontridder, Y. Poullet, The role of artificial intelligence in disinformation, Data & Policy 3 (2021) e32. doi:10.1017/dap.2021.20.

[9] Y. Abdallah, N. Salhab, A. E. Falou, Towards Fake News Identification using Machine Learning, in: 1st International Conference in Advanced Innovation on Smart City, ICAISC 2023 - Proceedings, IEEE, Jeddah, Saudi Arabia, 2023, pp. 1–6. doi:10.1109/ICAISC56366.2023.10085321.
[10] W. Gao, M. Ni, Y. Huang, TGNN: Topic-aware Graph Neural Network for Identifying Fake News from Social Media, in: 2022 IEEE 2nd International Conference on Electronic Technology, Communication and Information, ICETCI 2022, 2022, pp. 110–113. doi:10.1109/ICETCI55101.2022.9832149.
[11] J. Alghamdi, S. Luo, Y. Lin, A comprehensive survey on machine learning approaches for fake news detection, Multimedia Tools and Applications 83 (2023) 51009–51067. doi:10.1007/s11042-023-17470-8.
[12] F. Jahanbakhsh, Y. Katsis, D. Wang, L. Popa, M. Muller, Exploring the use of personalized ai for identifying misinformation on social media, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 2023, pp. 1–27.
[13] M. Hameleers, T. van der Meer, R. Vliegenthart, Civilized truths, hateful lies? incivility and hate speech in false information – evidence from fact-checked statements in the US, Inf. Commun. Soc. 25 (2022) 1596–1613.
[14] E. K. Vraga, L. Bode, Defining misinformation and understanding its bounded nature: Using expertise and evidence for describing misinformation, Polit. Commun. 37 (2020) 136–144.
[15] Z. Adams, M. Osman, C. Bechlivanidis, B. Meder, (Why) Is Misinformation a Problem?, Perspectives on Psychological Science 18 (2023) 1436–1463. doi:10.1177/17456916221141344.
[16] K. Hartwig, S. Schmid, T. Biselli, H. Pleil, C. Reuter, Misleading information in crises: exploring content-specific indicators on twitter from a user perspective, Behaviour & Information Technology 0 (2024) 1–34. URL: https://doi.org/10.1080/0144929X.2024.2373166. doi:10.1080/0144929X.2024.2373166.
[17] R. Parasuraman, T. B. Sheridan, C. D. Wickens, A model for types and levels of human interaction with automation, IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 30 (2000) 286–297. doi:10.1109/3468.844354.
[18] L. Onnasch, C. D. Wickens, H. Li, D. Manzey, Human performance consequences of stages and levels of automation: An integrated meta-analysis, Human Factors: The Journal of the Human Factors and Ergonomics Society 56 (2014) 476–488. URL: https://journals.sagepub.com/doi/10.1177/0018720813501549. doi:10.1177/0018720813501549.
[19] M. R. Endsley, Supporting human-ai teams: Transparency, explainability, and situation awareness, Computers in Human Behavior 140 (2023) 107574.
[20] R. Parasuraman, V. Riley, Humans and Automation: Use, Misuse, Disuse, Abuse, Human Factors 39 (1997) 230–253. doi:10.1518/001872097778543886.
[21] J. D. Lee, K. A. See, Trust in automation: Designing for appropriate reliance, Human Factors 46 (2004) 50–80. URL: https://doi.org/10.1518/hfes.46.1.50_30392. doi:10.1518/hfes.46.1.50_30392. arXiv:https://doi.org/10.1518/hfes.46.1.50_30392, pMID: 15151155.
[22] D. Manzey, J. Reichenbach, L. Onnasch, Human performance consequences of automated decision aids, J. Cogn. Eng. Decis. Mak. 6 (2012) 57–87.
[23] V. Alexander, C. Blinder, P. J. Zak, Why trust an algorithm? performance, cognition, and neurophysiology, Computers in Human Behavior 89 (2018) 279–288. URL: https://www.sciencedirect.com/science/article/pii/S0747563218303480. doi:https://doi.org/10.1016/j.chb.2018.07.026.
[24] R. Petty, J. Cacioppo, The elaboration likelihood model of persuasion, Advances in hydroscience 19 (1986) 124–205.
[25] T. Schrills, T. Franke, How do users experience traceability of ai systems? examining subjective information processing awareness in automated insulin delivery (aid) systems, ACM Trans. Interact. Intell. Syst. 13 (2023). URL: https://doi.org/10.1145/3588594. doi:10.1145/3588594.
[26] K. Akash, K. Polson, T. Reid, N. Jain, Improving human-machine collaboration through transparency-based feedback – part i: Human trust and workload model, IFAC-PapersOnLine 51 (2019) 315–321.

[27] M. Sieger, A. Calero Valdez, T. Schrills, Quantifying information dissemination: Structured review on methods and variables for sharing behavior on social network sites, 2024. URL: https://openaccess.cms-conferences.org/publications/book/978-1-964867-07-6/article/978-1-964867-07-6_0. doi:10.54941/ahfe1004858.
[28] A. Bertrand, T. Viard, R. Belloum, J. R. Eagan, W. Maxwell, On selective, mutable and dialogic xai: a review of what users say about different types of interactive explanations, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23, Association for Computing Machinery, New York, NY, USA, 2023. URL: https://doi.org/10.1145/3544548.3581314. doi:10.1145/3544548.3581314.
[29] C.-Y. Chen, S.-L. Chang, Belief in or identification of false news according to the elaboration likelihood model, 2021. URL: https://api.semanticscholar.org/CorpusID:232068287.
[30] M. Madsen, S. Gregor, Measuring Human-Computer Trust, in: 11th Australasian Conference on Information Systems, volume 53, ACIS, Brisbane, Australia, 2000, pp. 6–8.
[31] T. Franke, C. Attig, D. Wessel, A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale, International Journal of Human–Computer Interaction 35 (2019) 456–467. doi:10.1080/10447318.2018.1456150.
[32] T. Schrills, M. Zoubir, M. Bickel, S. Kargl, T. Franke, Are Users in the Loop? Development of the Subjective Information Processing Awareness Scale to Assess XAI, in: Proceedings of the ACM CHI Workshop on Operationalizing Human-Centered Perspectives in Explainable AI (HCXAI 2021), ACM, HCXAI Workshop, Yokohama, Japan, 2021, pp. 1–9.
[33] K. Desender, A. Boldt, N. Yeung, Subjective confidence predicts information seeking in decision making, Psychological Science 29 (2018) 761 – 778. URL: https://api.semanticscholar.org/CorpusID:4567205.
[34] P. Ward, Choice, uncertainty, and decision superiority: Is less ai-enabled decision support more?, IEEE Transactions on Human-Machine Systems 53 (2023) 781–791. URL: https://api.semanticscholar.org/CorpusID:259714483.
[35] D. Cyr, M. Head, E. Lim, A. Stibe, Using the elaboration likelihood model to examine online persuasion through website design, Information & Management 55 (2018) 807–821. URL: https://www.sciencedirect.com/science/article/pii/S0378720617303488. doi:https://doi.org/10.1016/j.im.2018.03.009.
[36] M. L. Bartlett, J. S. McCarley, Benchmarking Aided Decision Making in a Signal Detection Task, Human Factors 59 (2017) 881–900. doi:10.1177/0018720817700258. arXiv:28796974.