
Investigating Factors Behind User Trust in ChatGPT



ChatGPT, developed by OpenAI, is a versatile tool for various tasks, ranging from entertainment to health care queries. It can summarise large volumes of text, generate programming code, and assist with certain health care tasks. Despite these benefits, there are significant risks that can hinder its adoption in high-risk domains. These include potential inaccuracies, ethical issues (such as copyright and plagiarism), hallucination (producing false but plausible information), and biases from its pre-2021 training data. The discourse on ChatGPT's deployment recommends its use for tasks that humans can accurately supervise. Studies suggest using ChatGPT under human oversight to mitigate risks and ensure critical decisions remain human-made. This approach maximises efficiency by automating mundane tasks while preserving human expertise for complex ones. Successful integration of ChatGPT hinges on setting realistic performance expectations and understanding user perceptions of risk. The aim of [a recent study published in JMIR Human Factors](#) is to explore how perceived workload, satisfaction, performance expectancy, and risk-benefit perception influence users' trust in ChatGPT.

Factors Influencing User Trust in ChatGPT: A Conceptual Framework

User trust in ChatGPT is crucial and depends on its perceived accuracy and reliability. Positive user experiences can enhance trust and satisfaction, whereas inaccuracies can erode trust. Building and maintaining trust involves continuously validating the AI's outputs and acknowledging its limitations.

A conceptual framework based on established technology acceptance theories explores factors influencing user trust in ChatGPT. This includes performance expectancy, workload, satisfaction, risk-benefit perception, and trust. The framework hypothesizes that:

- Higher perceived workload negatively affects user trust.
- Higher perceived workload negatively affects user satisfaction.
- Higher user satisfaction positively affects trust.
- Higher performance expectancy positively affects trust.
- Positive risk-benefit perception positively affects trust.

Understanding these factors can aid in effectively integrating ChatGPT into various sectors by balancing its capabilities with informed user expectations.

Surveying User Perceptions: Methodology and Constructs

A semistructured, web-based questionnaire was distributed to U.S. adults who used ChatGPT (version 3.5) at least once a month. Data collection occurred from February to March 2023 via Qualtrics, managed by Centiment for its extensive reach and ability to ensure unique responses using fingerprinting technology. A preliminary soft launch with 40 responses was conducted to identify and resolve potential issues before full-scale dissemination.

The survey focused on five constructs: trust, workload, performance expectancy, satisfaction, and risk-benefit perception. Responses were measured on a 4-point Likert scale (1=strongly disagree to 4=strongly agree). The constructs were based on established models and theories such as TAM, UTAUT, the theory of planned behavior, and research on trust and security in digital environments.

1. Trust (T)

- T1: ChatGPT is competent in providing the information and guidance I need
- T2: ChatGPT is reliable in providing consistent and dependable information
- T3: ChatGPT is transparent
- T4: ChatGPT is trustworthy in the sense that it is dependable and credible
- T5: ChatGPT will not cause harm, manipulate its responses, or create negative consequences for me
- T6: ChatGPT will act with integrity and be honest with me
- T7: ChatGPT is secure and protects my privacy and confidential information

2. Workload (WL)

- WL1: Using ChatGPT was mentally demanding
- WL2: I had to work hard to use ChatGPT

3. Performance expectancy (PE)

- PE1: ChatGPT can help me achieve my goals
- PE2: ChatGPT can reduce my workload
- PE3: ChatGPT improves my work efficiency
- PE4: ChatGPT helps me make informed and timely decisions

4. Satisfaction (S)

- S: I am satisfied with ChatGPT

5. Risk-benefit perception (R)

- R: The benefits of using ChatGPT outweigh any potential risks

These theoretical foundations are critical for understanding the constructs and questions used in the survey, drawing from key research in information systems, human-computer interaction, and psychology.

Survey Insights: User Behaviour and Model Validation

Among the 607 survey participants, the usage frequency of ChatGPT varied: 29.9% used it monthly, 26.1% weekly, 24.5% more than once per week, and 19.4% almost daily. Educationally, 33.6% had at least a high school diploma, and 43.1% held a bachelor's degree. Participants primarily used ChatGPT for information (36%), amusement (33.4%), and problem-solving (22.2%). A smaller number used it for health-related inquiries (7.2%) and miscellaneous activities like brainstorming and content creation (1%).

The survey measured trust, workload, performance expectancy, satisfaction, and risk-benefit perception using a 4-point Likert scale. The model explained 2% of the variance in satisfaction and 64.6% in trust. Reliability was high for all constructs, with Cronbach's alpha and ρ values above 0.7 and average variance extracted above 0.5. The model fit was good, with an RMSEA value of 0.07, indicating the model adequately represents the data.

Key findings:

- Hypothesis 1: Increased perceived workload decreases trust in ChatGPT. This was not supported as the effect was not statistically significant.
- Hypothesis 2: Increased perceived workload decreases satisfaction with ChatGPT. This was supported.
- Hypothesis 3: Increased satisfaction with ChatGPT increases trust. This was supported.
- Hypothesis 4: Increased performance expectancy increases trust. This was supported.
- Hypothesis 5: Increased risk-benefit perception increases trust. This was supported.

The structural model shows the path coefficients for these relationships, confirming most hypotheses except for the impact of workload on trust.

Human Factors and Implications for Responsible AI

This study is among the first to explore how human factors such as workload, performance expectancy, risk-benefit perception, and satisfaction influence trust in ChatGPT. The results indicate that these factors significantly impact trust, with performance expectancy exerting the strongest influence, underscoring its critical role. Additionally, satisfaction was found to mediate the relationship between workload and trust, and there was a positive correlation between trust in ChatGPT and the perception of risk versus benefits. These findings align with the Biden-Harris Administration's efforts to advance responsible AI research and the development of the AI Risk Management Framework (AI RMF), which emphasises trustworthy AI systems.

Reducing user workload is vital for enhancing satisfaction, which in turn improves trust in ChatGPT. This finding is consistent with the AI RMF's focus on creating equitable and accountable AI systems. The study's results align with existing literature indicating that higher perceived workload leads to lower satisfaction and trust, similar to patterns observed in job satisfaction and well-being studies in various contexts. Higher user satisfaction correlates with increased trust in ChatGPT, consistent with previous research on digital transaction services and mobile transaction apps. This underscores the importance of ensuring user satisfaction to foster trust in innovative technologies like AI chatbots. The study also found a strong positive correlation between performance expectancy and trust in ChatGPT, extending findings from studies on

wearables and mobile banking. The mediating role of satisfaction between workload and trust is a notable contribution, highlighting the importance of performance expectancy in fostering trust.

The positive correlation between risk-benefit perception and trust aligns with prior studies on chatbot usage for digital shopping and customer service, confirming similar dynamics within the context of ChatGPT. However, the study has several limitations. It focused specifically on ChatGPT, so the results may not generalise to other AI chatbots. The reliance on self-reported data introduces potential response biases and measurement errors. The cross-sectional design captures data at one point in time, and longitudinal studies are needed for a comprehensive understanding of trust dynamics over time. Additionally, the sample consisted of active ChatGPT users, possibly excluding the perspectives of non-users.

Overall, this study highlights the significant factors influencing trust in ChatGPT: performance expectancy, satisfaction, workload, and risk-benefit perceptions. These insights contribute to the broader goal of responsible AI practices, emphasising user-centric design and safety. Future research should adopt longitudinal designs and include diverse user perspectives to deepen the understanding of trust in AI technologies.

Source: [JMIR Human Factors](#)

Image Credit: [iStock](#)

Published on : Wed, 29 May 2024