Trust Over Creepiness: AI Chatbot Perceptions in Balkan Countries

Marko Galjak[1]* and Marina Budić[1]
[1] Institute of Social Sciences, Belgrade, Serbia
*Corresponding author: galjak@gmail.com

This is the Accepted Manuscript

The Version of Record of this manuscript has been published and is available in *International Journal of Human–Computer Interaction*, 13 March 2026, at
https://www.tandfonline.com/doi/full/10.1080/10447318.2026.2638548

Galjak, M., & Budić, M. (2026). Trust Over Creepiness: AI Chatbot Perceptions in Balkan Countries. *International Journal of Human–Computer Interaction*, *0*(0), 1–25.
https://doi.org/10.1080/10447318.2026.2638548

Abstract

Generative AI research remains predominantly Western-centric. This mixed-methods study replicates Baek and Kim's (2023) chatbot acceptance model in the Balkan region (Bosnia and Herzegovina, Croatia, Montenegro, and Serbia). Surveying users (N=413), we validated the psychometric scales but found significant structural differences compared to the original US data. Specifically, task efficiency did not influence trust or creepiness. Conversely, playfulness emerged as a key predictor, reducing creepiness and building trust. While the US study found creepiness strongly deterred use, our sample showed trust was the dominant driver of continuance intention. Thematic analysis of open-ended responses identifies an "instrumentalist mindset", users perceive AI as a pragmatic tool rather than a social agent, explaining the emotional decoupling of efficiency from creepiness. The results demonstrate that Western-validated HCI models require cultural adaptation. We recommend that designers targeting similar contexts prioritize utility and transparency over anthropomorphic features.

Keywords: Uses and Gratifications Theory, ChatGPT, Large Language Models, Structural Equation Modeling, Uncanny Valley, Post-socialist context, User Motivation

1 Introduction

Generative artificial intelligence (AI) chatbots have spread faster than almost any other technology in the 21st century. OpenAI's ChatGPT sparked this shift in human-computer interaction by reaching 100 million users in just two months (Hu and Hu, 2023). By September 2025, ChatGPT had grown to over 700 million weekly active users globally, establishing itself as the go-to brand for standalone conversational AI (Chatterji et al., 2025). Concurrently, Google's Gemini has achieved broad reach through integration into the Android OS and Google Search, potentially interacting with billions of users daily via enhanced search results and AI-powered summaries. In the Balkan region, where Android holds over 70% of the mobile market share, millions of users in BCMS countries interact with Gemini's capabilities primarily through AI-enhanced search and OS features (StatCounter, 2025). The rapid global growth of AI chatbots has made the investigation of the motivations, perceptions, and behaviors of users a popular research topic, but the vast majority of empirical research on generative AI adoption has been conducted in Western, predominantly English-speaking populations (Ijiri and Healy, 2025; Nyaaba et al., 2024). This Western-centric research bias creates gaps in our understanding of how cultural values, technological contexts, and social norms may influence AI perception and usage patterns. Recognizing this, recent research has confirmed that national contexts significantly shape AI adoption, with studies showing that preferences for using AI for information vary across European countries, including Serbia (Reinhardt et al., 2025).

1.1 AI Chatbots and the Cross-Cultural Research Gap

Beyond cultural differences in user perception, this research gap is compounded by a fundamental technical problem: the capabilities of large language models (LLMs) driving the chatbots are not independent of language. The technology functions differently based on the language used for interaction, a factor that directly influences user experience and perception. Studies consistently find that models have difficulty with cross-lingual knowledge transfer; they often cannot answer a question in one language about a fact that was available in their training data in a different language (Goldman et al., 2025). In addition, models show marked cross-lingual inconsistency, giving different or even contradictory answers to the same factual question when it is asked in different languages (Qi et al., 2023). The method used to handle multilingual input (e.g., whether to translate a user's prompt into English before processing) can substantially alter the quality of the output, and no single method has proven optimal in all situations (Mondshine et al., 2025). These performance gaps originate from how models develop internal representations during their training. A model might “unify” a concept into a single, language-independent form. or it might create separate, language-specific “silos,” with the second outcome impeding effective transfer (Blum et al., 2025). Thus, the cultural differences in user perception that are the focus of this study are not simply a matter of subjective interpretation. They are a direct response to real, language-driven discrepancies in the AI's reliability, consistency, and behavior.

Cross-cultural validation of psychological constructs and theoretical models is needed for establishing the generalizability of findings and avoiding ethnocentric assumptions in technology adoption research (Matsumoto and van de Vijver, 2011; Triandis, 2018).

The conversational fluency of Generative AI introduces distinct challenges regarding user reliance. Unlike traditional deterministic software, LLMs are probabilistic and prone to hallucinations, where they generate convincing but factually incorrect information. Recent studies indicate that users often struggle to detect these errors, leading to over-reliance, where humans accept AI suggestions even when they are wrong (Vasconcelos et al., 2023). This phenomenon is particularly dangerous in specialized tasks; for example, Gu et al. (2024) found that even data analysts must develop complex verification workflows to catch subtle errors in AI-generated code. The question of AI adoption is not only whether users trust the system, but whether that trust is resilient enough to withstand errors and whether it leads to appropriate reliance rather than blind dependence (Dogru & Krämer, 2025).

1.2 Theoretical Background

While marketing literature often treats creepiness as a negative consumer reaction to personalization, broader psychological and sociological research defines it as a distinct emotional response to ambiguity and boundary transgression. McAndrew and Koehnke (2016) conceptualize creepiness as an adaptive anxiety aroused by the ambiguity of a threat; it arises when an individual is unsure if a situation is dangerous or if an interactant has benevolent or malevolent intentions. In the context of human-computer interaction, this ambiguity often stems from “mind perception” (Gray & Wegner, 2012). When machines display agentic qualities (planning, thinking), they are perceived as useful tools; however, when they display experiential qualities (feeling, sensing), they violate ontological categories, triggering the uncanny valley effect and feelings of moral unease (Gray & Wegner, 2012; Bigman et al., 2019). Creepiness signals a violation of contextual integrity and personal boundaries (Nissenbaum, 2010; Shklovski et al., 2014). When technology crosses the invisible line between a functional tool and a social agent, it creates a "leakiness" in personal space that users find unsettling (Shklovski et al., 2014; Tene & Polonetsky, 2013). Assessing user perceptions of AI requires looking beyond mere dissatisfaction to understand how users negotiate these boundaries of agency and ambiguity.

The way users perceive, evaluate, and emotionally respond to a technology is not a universal cognitive process but is shaped by what scholars in the sociology of technology term technological frames: the shared assumptions, knowledge, and expectations that a social group holds about a technology, which collectively define how its usefulness is experienced in practice (Bijker, 1995; Orlikowski & Gash, 1994; Abdelnour Nocera et al., 2007). These frames are socially constructed through a group's interpretive repertoires and material practices, meaning that the same artifact can be understood as a fundamentally different kind of object across cultural contexts (Abdelnour Nocera et al., 2007). Recent research on sociotechnical imaginaries of AI confirms this at a macro level, demonstrating that different societies imagine AI in ways that reflect their own culturally specific understandings of humanity, authority, and the purpose of technology, with some cultures framing AI as a potential human replacement and others as a collaborative instrument (Shi & Li, 2025; Richter et al., 2025). These culturally shaped mental models are particularly consequential in post-socialist societies, where the relationship between citizens and technology has been forged through distinct historical experiences. Scholars of post-socialist transformation have shown that the socialist legacy persists in institutional arrangements, but also in enduring "knowledge patterns" and "habits of thinking" that shape how individuals approach new economic and technological realities (Nuissl, 2005; Stark, 1992). Rather than approaching new technologies from a blank slate, post-socialist populations engage with them through what Stark (1992) calls "practiced routines," recombining existing cognitive and social resources through a process resembling bricolage rather than wholesale adoption of imported frameworks. This theoretical perspective suggests that if the mental models users bring to AI interaction are culturally constituted, then the emotional and relational responses predicted by models developed in Western contexts may not transfer straightforwardly to populations whose technological frames were shaped by different historical trajectories.

Our cultural framing departs from the dimensional approach to national culture that has dominated cross-cultural HCI research, most notably Hofstede's model of cultural dimensions (Hofstede et al., 2010). While parsimonious and widely cited, this framework has faced sustained critique on both methodological and paradigmatic grounds. Empirical replication studies have failed to recover Hofstede's original factor structure, even when closely following his prescribed methodology, raising questions about the robustness and generalizability of the dimensional scores themselves (Oshlyansky et al., 2006). More fundamentally, scholars in cross-cultural HCI and Science and Technology Studies (STS) have argued that the functionalist paradigm underlying such models is ill-suited to the study of dynamic culture-technology relationships. Macfadyen (2011) contends that dimensional frameworks treat culture as a static property of national populations, producing essentialized kinds of people rather than illuminating the situated, evolving ways of being through which groups actually engage with technology. Irani (2010) makes a parallel argument within HCI specifically, showing that design methods presumed to be culturally universal embed particular Western epistemologies and social relations, and proposing that culture be understood as everyday situated practice shaped by shared histories, media, and economic conditions rather than as a portable "software of the mind". This critique is reinforced by recent systematic evidence of a pervasive Western (WEIRD) sampling bias in human-AI interaction research: Peters and Carman (2024) found that the vast majority of Explainable AI user studies sampled only Western populations yet generalized their findings to human-AI interaction broadly, obscuring culturally specific explanatory needs and trust dynamics. Rather than assigning dimensional scores to the BCMS region, we therefore adopt a practice-oriented approach: we use mixed methods to examine how users in a specific post-socialist context actually relate to AI chatbots and then interpret the emerging patterns through the historically situated technological frames described above.

Parallel to this theoretical work, the HCI community has developed practitioner-oriented design frameworks for human-AI interaction that operationalize concepts like transparency, explainability, and user control into actionable guidelines (Amershi et al., 2019; Shneiderman, 2020; Weisz et al., 2024). These frameworks provide a basis for translating empirical findings about user perceptions into concrete design recommendations, a connection we return to in the Discussion.

In the original study, Baek and Kim (2023) developed and tested a theoretical model based on Uses and Gratifications (U&G) theory to explain why individuals use ChatGPT. Their model investigated how five core motivations (information seeking, task efficiency, personalization, social interaction, and playfulness) influence users' perceptions of creepiness and trust, which in turn affect their intention to continue using the service. Using data from 421 participants, U.S.-based sample, recruited via Amazon Mechanical Turk, they found that personalization reduced creepiness while task efficiency and social interaction increased it, with trust serving as a positive predictor and creepiness as a negative predictor of continuance intention. These findings revealed complex, sometimes counterintuitive relationships between user motivations and emotional responses to AI, suggesting that efficiency gains could paradoxically increase user discomfort. For instance, while personalization decreased perceived creepiness, motivations for task efficiency and social interaction unexpectedly increased it. However, their sample was predominantly drawn from US population, limiting the generalizability of findings to other cultural contexts. Given evidence suggesting cultural differences in technology acceptance and privacy concerns, replicating foundational AI studies in new cultural contexts is a priority for the field (Dwivedi et al., 2023; Hofstede et al., 2010).

1.3 The Present Study

This study applies Baek and Kim's (2023) model to the Bosnian, Croatian, Montenegrin, and Serbian (BCMS) speaking populations to determine its validity outside of a Western context. These four countries of former Yugoslavia share the same language (previously called Serbo-Croatian and more recently a more inclusive version: BCMS) and very similar cultural milieu. The Balkan region offers a compelling setting for cultural comparison, as its post-socialist context fosters public skepticism toward powerful institutions, including contemporary data-gathering entities (Mishler and Rose, 1997). Higher uncertainty avoidance and collectivism, when compared to the United States, may alter how users in this region perceive AI's efficiency, social potential, and trustworthiness. The region has undergone digital transformation (Broz et al., 2020), but it remains under-examined in AI and human-computer interaction research, leaving a gap in the literature on global AI adoption.

Therefore, the present research has three core objectives. First, we aim to psychometrically validate and culturally adapt the measurement scales for AI chatbot motivations, creepiness, trust, and continuance intention developed by Baek and Kim (2023). Second, we test the structural relationships of their theoretical model to determine if the pathways between motivations and outcomes hold true or differ significantly within the BCMS context. Employing a mixed-methods approach, we use qualitative data to explore and suggest explanations for the observed quantitative differences. Our study helps validate HCI theories cross-culturally and provides practical guidance for creating AI systems suitable for diverse populations. The results should also help developers and policymakers design and deploy AI systems that are more culturally aware and widely accepted.

2 Method

2.1 Participants and Procedure

Participants were recruited through Meta advertisements targeting individuals in four BCMS countries (Bosnia and Herzegovina, Croatia, Montenegro, and Serbia) who are older than 15. The target population was internet users in four countries. The survey targeted the general population, with the inclusion criteria being that respondents had used any chatbot in the last 2 years.

GDPR compliant EU Survey software was used for questionnaires. We prepared four questionnaires, one for each of the BCMS countries. The questionnaires were adjusted for each language, as there are minor differences between the languages. The questionnaires were then examined by three native speakers from Croatia, Bosnia and Herzegovina, and Montenegro to check for any mistakes and for further adjustments. The translation followed a translate-review procedure rather than formal back-translation. Because BCMS varieties are mutually intelligible and differ primarily in lexical preferences rather than in grammatical structure, a review by independent native speakers from three of the four target countries served as a functional equivalent to back-translation for detecting semantic drift or loss of construct meaning. Each reviewer independently flagged items where the phrasing felt unnatural or where the intended meaning diverged from the original English, and discrepancies were resolved through discussion among the research team. After receiving comments and making final adjustments, we ran a pretest on 14 colleagues and acquaintances.

Data collection started in April 2025 and continued until July 2025.

The survey asked questions about opinions on AI (N=786), and the subset of users used chatbots in the last two years (n=420, 53.4%). Prior to analysis, data were screened for straight-line responding, where participants provided the same response across multiple items regardless of item content. Respondents who gave identical answers to 10 or more items across motivation and perception scales were identified and removed from the dataset (n = 7), as this pattern suggests insufficient attention or engagement with the survey. The final sample size was n=413, and the sample description is provided in Table 1.

While the median age of 50 years may appear high compared to typical convenience samples in technology adoption research, it reflects the demographic reality of the BCMS region. The median age of the general population in these countries exceeds 44 years (Eurostat, 2025), and when restricting to the target population of individuals aged 18+ (our inclusion criterion), the expected median age of the sampling frame would be substantially higher, approximately 49-53 years. Our sample median of 50, therefore, appears broadly representative of the adult chatbot user population in the region, which generally skews younger.

Table 1. Sample descriptive statistics

	n (%) or Mean ± SD
Country	
Serbia	169 (40.9%)
Bosnia and Herzegovina	116 (28.1%)
Croatia	93 (22.5%)
Montenegro	35 (8.5%)
Gender	
Female	239 (57.9%)
Male	172 (41.6%)
Other	2 (0.5%)

	n (%) or Mean ± SD
Age (years)	48.9 ± 11.8
Range	18-80
Median	50
Residence Type	
Larger city (≥100,000 inhabitants)	196 (47.5%)
Small city	140 (33.9%)
Town	39 (9.4%)
Village	38 (9.2%)
Education Level	
High school (4 years)	109 (26.4%)
4-year undergraduate	84 (20.3%)
Master's degree	64 (15.5%)
College	53 (12.8%)
High school (3 years)	29 (7.0%)
PhD	22 (5.3%)
Other higher education	43 (10.4%)
Primary school	9 (2.2%)
Years of Education	15.1 ± 3.0
Language used with chatbot	
BCMS	319 (77.2%)
English	88 (21.3%)
Other	6 (1.5%)
Chatbot primarily used	
ChatGPT	334 (80.9%)
Gemini	32 (7.7%)
Deepseek	12 (2.9%)
Claude	4 (1%)
Other	31 (7.5%)

2.2 Measures

In the original study by Baek and Kim (2023), user motivations for chatbot usage were measured using five scales derived from uses and gratifications (U&G) theory, adapted to the context of generative AI. These included: (1) Information seeking (3 items), adapted from Leung and Wei (1998) and Rubin (1983), assessing users' desire to acquire knowledge, satisfy curiosity, and obtain answers to unknown questions (e.g., "I use ChatGPT because I can seek information to satisfy my curiosity"); (2) Task efficiency (4 items), adapted from Choi and Drumwright (2021), evaluating how chatbots aid in completing tasks more easily, improving quality, and supporting multitasking (e.g., "I use ChatGPT because it makes my task easier"); (3) Personalization (4 items), adapted from Baek and Morimoto (Baek and Morimoto, 2012), measuring the extent to which chatbots tailor responses to individual needs and preferences (e.g., "I use ChatGPT because using ChatGPT is customized to my needs"); (4) Social interaction (3 items), adapted from Choi and Drumwright (2021), capturing users' use of chatbots for conversational purposes, such as casual talks or emotional support (e.g., "I use ChatGPT because I can have a conversation when I feel sad"); and (5) Playfulness (6 items), adapted from Kim and Baek (2022), gauging hedonic aspects like fun, enjoyment, and curiosity stimulation (e.g., "I use ChatGPT because it is fun for me").

The study measured perceptual and behavioral outcomes: (6) Creepiness (4 items), adapted from Rajaobelina et al. (2021), assessing feelings of unease, queasiness, fear, and threat during chatbot interactions (e.g., "When using ChatGPT, I have a queasy feeling"); (7) Trust (3 items), adapted from Kim and Baek (2022), evaluating perceptions of chatbots as believable, credible, and trustworthy (e.g., "ChatGPT is believable"); and (8) Continuance intention (2 items), adapted from Choi and Drumwright (2021), measuring users' plans and desires to keep using chatbots (e.g., "I plan to keep using ChatGPT"). All items were rated on a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree).

2.3 Quantitative Analysis Plan

Quantitative analysis proceeded in four sequential phases to ensure robust validation of the measurement model before testing structural relationships.

First, scale reliability was assessed using Cronbach's alpha coefficients for each construct, with values above 0.70 considered acceptable and values above 0.80 indicating good internal consistency (Nunnally and Bernstein, 1994). Item-total correlations and the impact of item deletion on scale reliability were examined to identify potentially problematic items.

Second, confirmatory factor analysis (CFA) was employed to validate the eight-factor measurement model proposed by Baek and Kim (2023). Maximum likelihood (ML) estimation was used with listwise deletion for missing data. Model fit was evaluated using multiple indices following established guidelines: χ^2/df ratio < 3.0, Comparative Fit Index (CFI) ≥ 0.90, Tucker-Lewis Index (TLI) ≥ 0.90, Root Mean Square Error of Approximation (RMSEA) ≤ 0.08 with 90% confidence interval upper bound < 0.10, and Standardized Root Mean Square Residual (SRMR) ≤ 0.08 (Hu and Bentler, 1999; Kline, 2016). Convergent validity was assessed through Average Variance Extracted (AVE) values, with AVE > 0.50 indicating adequate convergent validity. Discriminant validity was evaluated using the Fornell-Larcker criterion, requiring the square root of AVE for each construct to exceed its correlations with other constructs.

Third, structural equation modeling tested the hypothesized relationships between constructs. The structural model included paths from the five motivation constructs (information seeking, task efficiency, personalization, social interaction, and playfulness) to both creepiness and trust, and from creepiness and trust to continuance intention, consistent with the original theoretical framework. Standardized path coefficients and their significance levels were examined, along with R^2 values for endogenous variables to assess the model's explanatory power.

Finally, additional validation procedures included testing for common method bias using Harman's single factor test, where a single-factor CFA model was compared to the proposed measurement model. A substantially poorer fit for the single-factor model would indicate that common method variance was not a significant concern. Modification indices were reviewed to identify potential model improvements, though modifications were only considered if theoretically justified to avoid capitalization on sample-specific variance.

To assess whether the adapted scales function equivalently across BCMS contexts, we tested measurement invariance between respondents from Serbia (RS; n = 169) and respondents from the other BCMS countries combined (Other; n = 244). Multi-group confirmatory factor analyses were estimated using WLSMV (appropriate for ordinal Likert items). We tested configural, metric and scalar invariance and evaluated nested model comparisons using conventional change-in-fit criteria ($\Delta CFI \leq 0.010$;

ΔRMSEA $\leq$ 0.015). When specific item intercepts were flagged by modification indices as non-invariant, we implemented partial scalar invariance by freeing those intercepts and re-evaluated fit. All invariance outputs, modification indices, and the R code used are provided in Supplementary Material (SM) and code in the linked repository.

2.4 Qualitative Analysis Plan

We used a mixed-methods (QUAN → qual) design, where we first analyzed quantitative data and only then looked at the qualitative responses to help understand the results. Our main goal here was to figure out why the SEM results differ from the original Baek and Kim (2023) study.

The analysis was conducted on responses to the open-ended survey question: “Describe in your own words the main reasons why you use the chatbot?” Of the 413 survey participants, 303 (73.4%) responded to this question.

The thematic analysis steps from Braun and Clarke (2006) were used, with software to help with the process. We began by reading the entire dataset, which led us to identify six initial, recurring themes for coding. We then coded all 303 responses using this framework. After this process, it became clear that six themes were insufficient, as they failed to categorize a significant portion of the data. All uncoded responses (n=68, 22.4%) were flagged and collated for a separate, focused review that found distinct and repeating concerns not captured by the initial framework. Based on this review, a consensus was reached to expand the framework by including two more themes, to account for these remaining responses.

For coding, we used a large language model (GPT-4o). We then conducted a human verification step to ensure the accuracy of automated coding. Each author independently reviewed half of the sample to identify misclassifications. During this check, we discovered some misclassifications that led to the improvement of the prompt, and subsequently ran the classification and manual checking again, with each author examining half of the responses and designated classifications. The final prompt produced classifications with perfect fidelity in applying the established coding frame, as no significant or systematic misclassifications were found.

The original BCMS responses, translated English versions, prompts, and all analysis and visualization scripts are openly available with the data to ensure full reproducibility.

2.5 Software and Replication

All statistical analyses were conducted in R version 4.3.2 (R Core Team, 2023). Data manipulation and preparation utilized the tidyverse suite of packages (Wickham et al., 2019), including readxl for data import (Wickham and Bryan, 2023) and stringr for text processing (Wickham, 2023). Descriptive statistics and scale reliability analyses were performed using the psych package (Revelle, 2024), with bootstrap procedures implemented via the boot package (Canty and Ripley, 2022; Davison and Hinkley, 1997). Structural equation modeling employed lavaan (Rosseel, 2012) and semTools (Jorgensen et al., 2022), with regression diagnostics conducted using car (Fox and Weisberg, 2019), performance (Lüdecke et al., 2021), lmtest (Zeileis and Hothorn, 2002), and sandwich (Zeileis, 2006, 2004; Zeileis et al., 2020). Visualizations were created using ggplot2 (Wickham, 2016, p. 201), corrplot (Wei and Simko, 2024), patchwork (Pedersen, 2024), ggVennDiagram (Gao and Dusa, 2024), and ggvenn (Yan, 2023). Qualitative analysis utilized the tm package for text mining (Feinerer et al., 2008; Feinerer and Hornik, 2024) and the openai package for GPT-4o assisted coding (Rudnytskyi, 2023). Tables and reports were generated using knitr (Xie, 2024, 2015, 2014) and kableExtra (Zhu, 2024).

All data and R code for analysis are available in an online repository (anonymized link).

2.6 Ethical approval and informed consent

Ethical approval for this research was granted by the Ethics Committee of the Institute of Social Sciences, Republic of Serbia, following their review of the project "Trust and Usage of LLMs in the BCMS Language Sphere" (Decision No. 1-6, 2025). Participants were informed about the purpose and nature of the study at the beginning of the survey. Written informed consent was obtained from all participants, who indicated their agreement by checking the appropriate consent boxes before proceeding with the questionnaire.

3 Results

3.1 Descriptive Statistics and Scale Reliability

Table 2 presents the descriptive statistics and reliability coefficients for all study constructs. All scales showed high internal consistency; Cronbach's alpha values ranged from 0.882 to 0.945, well above the recommended 0.70 threshold for reliability.

Table 2. Descriptive Statistics, Scale Reliability, and Comparison with Original Study

Construct	Items	Mean (SD)	α (Current Study)	α (Baek & Kim, 2023)	Δα
Information Seeking	3	5.86 (1.46)	0.882	0.71	+0.172
Task Efficiency	4	5.25 (1.69)	0.921	0.78	+0.141
Personalization	4	4.65 (1.69)	0.887	0.76	+0.127
Social Interaction	4	3.75 (2.12)	0.923	0.79	+0.133
Playfulness	6	4.33 (1.82)	0.929	0.75	+0.179
Creepiness	4	1.81 (1.37)	0.928	0.84	+0.088
Trust	3	4.52 (1.58)	0.900	0.88	+0.020
Continuance Intention	2	5.41 (1.76)	0.945	0.81	+0.135

Note: All scales measured on 7-point Likert scales (1 = strongly disagree, 7 = strongly agree). α = Cronbach's alpha; Δα = difference between current and original study.

Compared to the original study, all scales had higher internal consistency in the BCMS sample. The improvements ranged from 0.020 (trust) to 0.179 (playfulness), with the largest increases for playfulness (α = 0.929 vs. 0.75), information seeking (α = 0.882 vs. 0.71), and task efficiency (α = 0.921 vs. 0.78). The higher reliability coefficients may suggest the translated scales were more consistent for BCMS users, perhaps because this group has a more uniform way of thinking about chatbot interactions.

On average, BCMS users reported high information seeking (M = 5.86) and continuance intention (M = 5.41), moderate task efficiency (M = 5.25) and trust (M = 4.52), and very low creepiness (M = 1.81). The mean for social interaction was relatively low (M = 3.75), but the high standard deviation (SD = 2.12) points to large individual differences in using chatbots for social connection. This variance became important in the subsequent structural model test.

3.2 Measurement Model Validation

The eight-factor measurement model was tested using confirmatory factor analysis with maximum likelihood estimation (Table 3). The model showed an acceptable fit to the data: $\chi^2 = 1385.492$, df = 377, $\chi^2/df = 3.67$, CFI = 0.916, TLI = 0.904, RMSEA = 0.080 (90% CI: 0.076-0.085), SRMR = 0.062. While the chi-square test was significant ($p < 0.001$), this is expected with large samples (N = 413). The CFI and TLI values were above the 0.90 threshold, and both RMSEA and SRMR were inside acceptable limits ($\leq$ 0.08), which supports an adequate model fit.

Table 3. Standardized Factor Loadings from Confirmatory Factor Analysis

Construct	**Item**	**Loading**	**Construct**	**Item**	**Loading**
Information Seeking			**Playfulness**		
	info_curiosity	0.739		play_fun	0.787
	info_knowledge	0.922		play_curiosity	0.830
	info_unknown	0.877		play_exploration	0.808
Task Efficiency				play_satisfaction	0.887
	task_saves_time	0.918		play_open_mind	0.841
	task_easier	0.959		play_happy	0.822
	task_quality	0.886	**Creepiness**		
	task_multitasking	0.705		creepiness_uncomfortable	0.915
Personalization				creepiness_insecure	0.952
	personal_adapted	0.899		creepiness_undefined_fear	0.789
	personal_specific	0.905		creepiness_threatened	0.821
	personal_recommendations	0.860	**Trust**		
	personal_unique	0.654		trust_convincing	0.778
Social Interaction				trust_credible	0.952
	social_always_available	0.762		trust_reliable	0.887
	social_when_sad	0.887	**Continuance Intention**		
	social_without_others	0.919		continue_plan	0.966
	social_informal	0.906		continue_want	0.929

All standardized factor loadings exceeded 0.65, with the majority above 0.80, indicating strong relationships between items and their respective constructs. The lowest loading was for personal_unique (0.654), which still exceeded the minimum threshold of 0.50 for acceptable convergent validity.

3.3 Convergent and Discriminant Validity

Convergent validity was assessed through Average Variance Extracted (AVE) values calculated from the reliability output. All constructs demonstrated adequate convergent validity with AVE values exceeding the 0.50 threshold: information seeking (0.718), task efficiency (0.751), personalization (0.671), social interaction (0.759), playfulness (0.690), creepiness (0.777), trust (0.764), and continuance intention (0.896).

Discriminant validity was evaluated using the Fornell-Larcker criterion. The correlation matrix revealed that all constructs met this criterion, with the square root of each construct's AVE exceeding its correlations with other constructs. The highest inter-construct correlation was between social interaction and playfulness ($r = 0.807$), which remained below the square root of their respective AVE values ($\sqrt{0.759} = 0.871$ for social interaction and $\sqrt{0.690} = 0.831$ for playfulness).

3.4 Common Method Bias Assessment

To test for common method bias, Harman's single factor test was conducted. A single-factor CFA model resulted in much poorer fit indices (CFI = 0.510, TLI = 0.474, RMSEA = 0.188, SRMR = 0.135) compared to the proposed eight-factor model. This large decrease in model fit suggests that common method variance is not a major concern for the validity of the results.

The measurement model validation shows that the adapted scales have sound psychometric properties for the BCMS sample. All constructs had high reliability ($\alpha = 0.882$-0.945) and met the criteria for both convergent (AVE > 0.50) and discriminant validity, all of which support the use of these scales for testing the structural model.

Measurement invariance (Serbia vs Other)

The configural model indicated good fit (CFI = 0.999, RMSEA = 0.045, SRMR = 0.058), supporting the same factor structure in both groups. Constraining factor loadings (metric invariance) produced a negligible deterioration in fit (CFI = 0.998, RMSEA = 0.059; ΔCFI ≈ 0.001, ΔRMSEA ≈ 0.014), supporting metric invariance. The full scalar model did not indicate widespread failure by the ΔCFI / ΔRMSEA criteria (CFI = 0.999, RMSEA = 0.045), however modification indices identified a small set of non-invariant intercepts (trust_convincing, trust_reliable, personal_adapted, info_unknown). We therefore fitted a partial scalar invariance model freeing those intercepts; the partial model fit was CFI = 0.999, RMSEA = 0.045, representing ΔCFI = 0.001 and ΔRMSEA = −0.014 relative to the metric model, which meets recommended thresholds. Sensitivity analyses (re-estimating the SEM separately by group) returned broadly similar structural path patterns (for example, trust → continuance: RS $\beta = 0.786$, $p < .001$; Other $\beta = 0.848$, $p < .001$), indicating that the substantive conclusions about the structural relations are robust across groups. (Full invariance tables and groupwise SEM estimates are in SM Tables A–C)

3.5 Structural Model Testing

The structural model examining relationships between user motivations, creepiness, trust, and continuance intention had an acceptable to borderline fit: $\chi^2 = 1483.603$, df = 383, $\chi^2/df = 3.87$, CFI = 0.909, TLI = 0.896, RMSEA = 0.083 (90% CI: 0.079-0.088), SRMR = 0.074 (Table 4). While the chi-square was significant ($p < 0.001$), as expected with large samples, the CFI met the 0.90 criterion, the TLI approached it (0.896), and the RMSEA and SRMR were reasonably close to recommended thresholds. Given the model's complexity and cross-cultural context, we proceeded with hypothesis testing while acknowledging these fit limitations.

Table 4. Comparison of Standardized Path Coefficients: BCMS Sample vs. Original Study

Path	BCMS Sample	Baek & Kim (2023)
H1: Motivations → Creepiness		
Information seeking → Creepiness	-0.079 (ns)	-0.171 (ns)
Task efficiency → Creepiness	-0.114 (ns)	0.607**
Personalization → Creepiness	0.082 (ns)	-0.437*
Social interaction → Creepiness	0.264**	0.229**
Playfulness → Creepiness	-0.385***	0.321 (ns)
H2: Motivations → Trust		
Information seeking → Trust	0.093 (ns)	-0.021 (ns)
Task efficiency → Trust	-0.024 (ns)	0.291*
Personalization → Trust	0.390***	0.522***
Social interaction → Trust	0.052 (ns)	0.237 (ns)
Playfulness → Trust	0.369***	0.041 (ns)
H3 & H4: Mediators → Continuance		
Creepiness → Continuance intention	-0.074†	-0.704***
Trust → Continuance intention	0.690***	0.187***

Note: †p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001; ns = not significant

The results from the structural model differed from the original study (Figure 1). For instance, task efficiency had no significant relationship with either creepiness or trust in the BCMS sample, contrasting with the original findings where task efficiency increased both creepiness ($\beta = 0.607$, $p < 0.05$) and trust ($\beta = 0.291$, $p < 0.10$). This may mean that BCMS users do not feel the same tension between efficiency and comfort as was found in the Western sample.

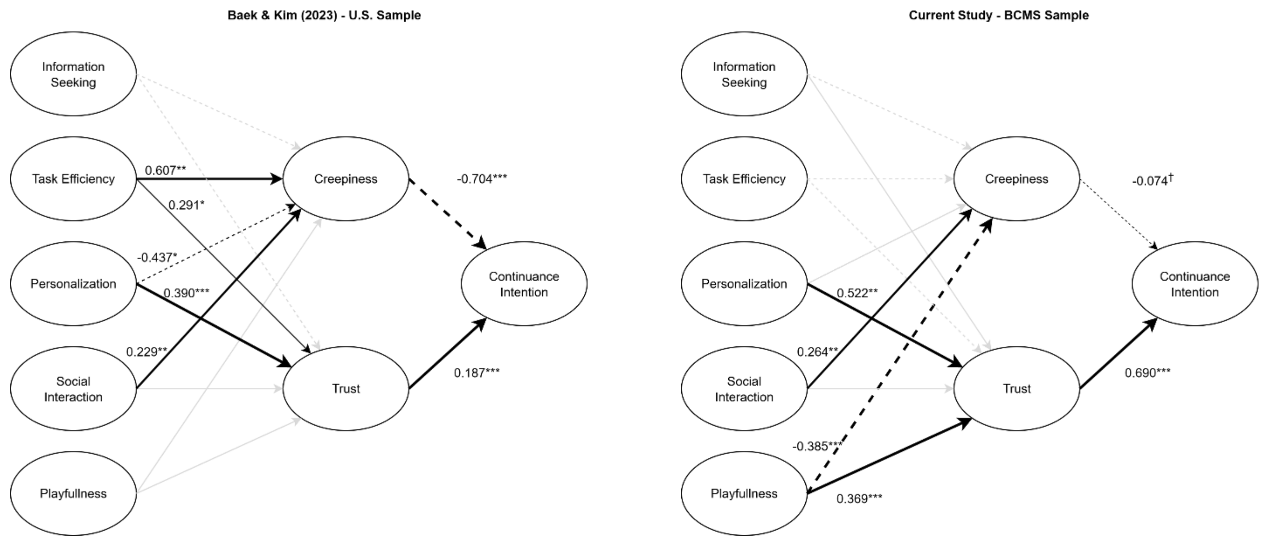

Figure 1. Comparison of structural model results between U.S. and BCMS samples
Note: Standardized path coefficients shown on paths. Solid lines indicate positive relationships, dashed lines indicate negative relationships, and pale lines indicate non-significant paths. †$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

In the BCMS context, playfulness was a strong predictor; it significantly reduced creepiness ($\beta = -0.385$, $p < 0.001$) and increased trust ($\beta = 0.369$, $p < 0.001$). In the original study, however, it had no significant effects. This difference suggests that BCMS users have a more positive view of playful interactions with chatbots, potentially reflecting different cultural attitudes toward AI personality and humor.

Social interaction consistently increased creepiness in both samples (BCMS: $\beta = 0.264$, $p < 0.01$; Original: $\beta = 0.229$, $p < 0.05$), suggesting a discomfort, consistent across both samples studied, with overly social AI behaviors. However, personalization showed divergent effects: while it reduced creepiness in the original study ($\beta = -0.437$, $p < 0.10$), it had no significant effect in BCMS, though it remained a strong predictor of trust in both contexts.

3.6 Variance Explained

The model explained less variance in creepiness among BCMS users ($R^2 = 0.092$) compared to the original study ($R^2 = 0.211$), implying that additional cultural or contextual factors not captured by the model influence creepiness perceptions in BCMS . Conversely, the model explained more variance in trust ($R^2 = 0.606$ vs. 0.519) and comparable variance in continuance intention ($R^2 = 0.499$ vs. 0.511).

The links between the mediators and continuance intention were also very different. In the original study, creepiness had a strong negative effect on continuance intention in the original study ($\beta = -0.704$, $p < 0.001$), but this effect was only marginal in the BCMS sample ($\beta = -0.074$, $p < 0.10$). Trust showed a much stronger positive effect on continuance intention among BCMS users ($\beta = 0.690$, $p < 0.001$) compared to the original study ($\beta = 0.187$, $p < 0.001$). For BCMS users, then, continued use of ChatGPT appears to depend more on trust than on a lack of creepiness.

3.7 Language as a Potential Moderator

Given documented LLMs performance differences across languages, we examined whether language of use (BCMS: n=319, 77.2%; English: n=88, 21.3%) moderated structural relationships. English users reported significantly lower social interaction (d=0.53, p<.001), playfulness (d=0.39, p<.001), trust (d=0.32, p=.006), and continuance intention (d=0.34, p=.012), but did not differ on information seeking, task efficiency, or creepiness (see SM Table D). Multi-group CFA supported configural and metric invariance, with partial scalar invariance achieved by freeing four item intercepts. Critically, multi-group SEM revealed that constraining all structural paths to equality did not worsen fit ($\Delta\chi^2$=9.99, df=12, p=.617), indicating equivalent relationships across language groups. The study's central finding replicated robustly: trust strongly predicted continuance intention in both BCMS (β=0.690, p<.001) and English (β=0.714, p<.001) users. While statistical power was limited in the smaller English subsample, these analyses provide reasonable confidence that language of use does not fundamentally alter the theoretical model (detailed results in SM Table D to F).

3.8 Qualitative Findings

To understand the reasons for the quantitative differences in our structural model, we thematically analyzed 303 open-ended responses. This process yielded eight themes that describe the motivations, perceptions, and mental models of BCMS chatbot users. These themes help explain key findings, such as why task efficiency failed to predict creepiness or trust, and why playfulness was a predictive factor. An upset plot (Figure 2) visualizes the relationships and co-occurrence of these themes.

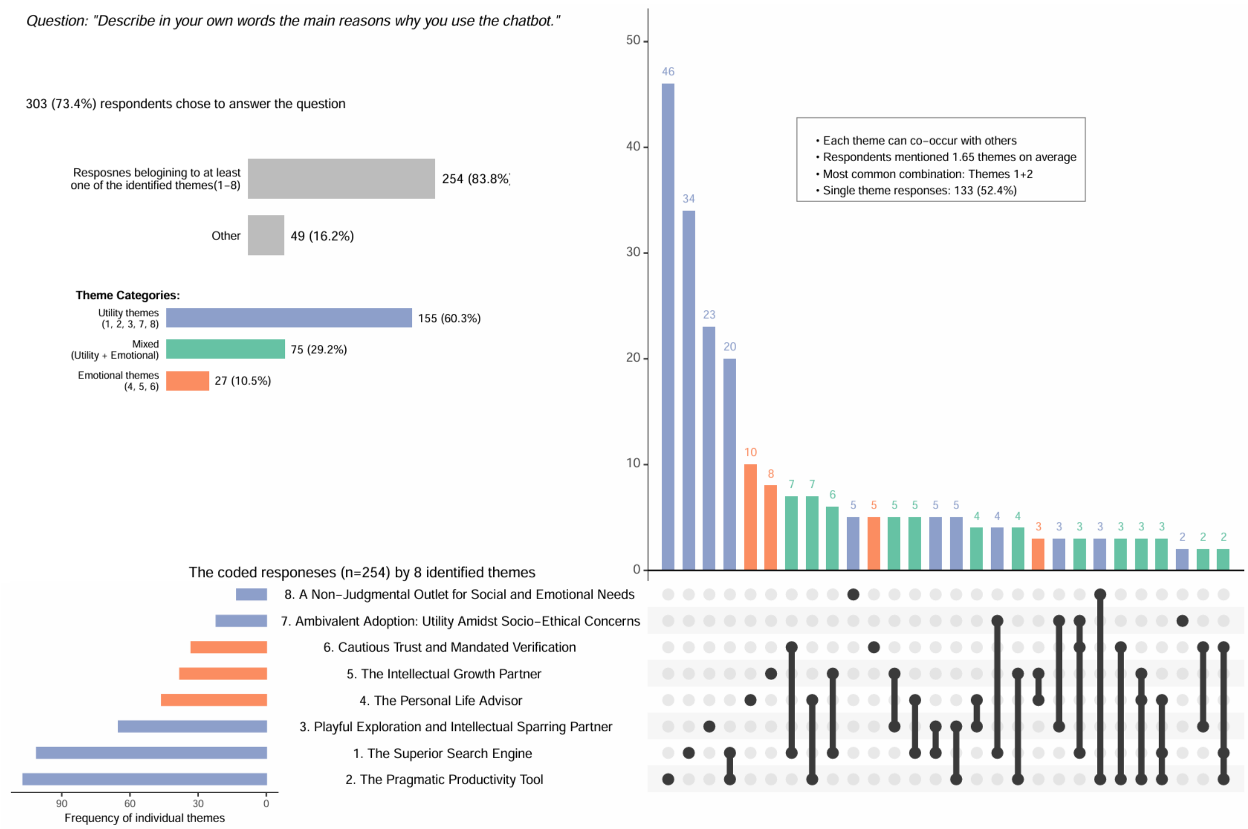

Figure 2. Upset plot analysis of qualitative themes identifying chatbot usage motivations among BCMS respondents.
Note: Each column in the matrix represents a unique combination of themes found in respondents' open-ended responses. The plot visualizes how different motivations cluster together, revealing distinct user profiles and usage patterns.

3.8.1 Theme 1: The Pragmatic Productivity Tool: Efficiency Decoupled from Emotion

This theme represents the view of the chatbot as a utilitarian tool. It was the most dominant theme (n=155), appearing in over 60% of all coded responses and the most common standalone motivation. This suggests that for many users, the AI's function as a productivity tool is its sole purpose. Users describe it as an "assistant," "secretary," or "collaborator" that helps with work, coding, text formulation, and planning. The relationship is transactional and instrumental, focused on output and efficiency gains. Illustrative quotes: "It speeds up the implementation of technical solutions at work.", and "It saves me from stupid, boring, repetitive tasks."

This theme helps explain the most striking difference in our structural model: why Task Efficiency was not a significant predictor of either Creepiness or Trust in the BCMS sample. When a chatbot is viewed as a simple tool, like a spell-checker or compiler, its efficiency becomes emotionally neutral. BCMS users seem to hold a pragmatic mental model of the AI, separating its performance (task efficiency) from relational judgments (creepiness, trust). The original study's sample may have anthropomorphized the AI more, leading to a "tension" where high efficiency felt unsettling or "too smart," thus increasing creepiness.

3.8.2 Theme 2: The Superior Search Engine: Information Synthesis and Curation

Users often described the chatbot as a better alternative to traditional search engines. This was the second most frequent theme overall and the most common motivation to be paired with "The Pragmatic Productivity Tool," forming a core "utility user" profile. Its key advantage is the ability to synthesize, compile, and summarize information from multiple sources into a single, coherent response, which saves users from sifting through multiple search results. Illustrative quotes: (1) "I use ChatGPT because it explains everything I need in one text. Unlike a traditional search engine, where I have to click on several websites to find what interests me.", and (2) "For finding useful information, it is easier for me to use a chatbot than search engines like Google."

This theme aligns with the high mean score for Information Seeking (M = 5.86) and shows that users value not just finding facts, but the efficiency of the retrieval process. The upset plot shows that when users have multiple motivations, this theme is almost always one of them, underlining its importance. This perception of the chatbot as an effective information synthesizer builds functional trust and encourages continued use.

3.8.3 Theme 3: Cautious Trust and Mandated Verification

Many users expressed a conditional trust. The upset plot shows this theme rarely appears alone; it is a mindset that accompanies utilitarian motives. Users were aware of the AI's potential for inaccuracy and treated its output with skepticism, often cross-referencing and fact-checking. Their trust was not a passive acceptance, but a conditional reliance contingent on their ability to verify the output. Illustrative quotes: "I always approach it with caution because I know it can produce complete nonsense, so I always check the results to ensure they are well-founded." and "I don't have complete trust in all the facts it offers, so I additionally verify them when I deem it necessary. I believe that ChatGPT is a good tool for those who know how to verify information and still think independently."

The co-occurrence of this theme with pragmatic ones explains two quantitative findings. It clarifies why trust was the strongest predictor of continuance intention ($\beta = 0.690$), because in a context where users know the tool is fallible, establishing a baseline of functional trust becomes essential for continued use. It also explains the moderate mean for Trust (M = 4.52), as the constant need for verification lowers overall trust ratings, showing it is conditional, not absolute.

3.8.4 Theme 4: Playful Exploration and Intellectual Sparring Partner

Beyond practical uses, some engaged with the chatbot out of curiosity and a desire to test its limits, driven by novelty and intellectual stimulation. The upset plot shows this theme has a dual role: for some, it is a standalone motivation for pure entertainment, while for others, it co-occurs with pragmatic themes, suggesting a pattern where users first establish the tool's utility and then begin to explore its capabilities playfully. Illustrative quotes: “I asked very strange and unusual questions to which I received specific answers, for example, how much coal a steam locomotive consumes per 100 km. I was amazed by the openness.” and “I am testing it to see if I will get the expected answer on some rare topic. I always succeed.”

This theme explains why Playfulness was a strong predictor in the BCMS sample, reducing Creepiness ($\beta = -0.385$) and increasing Trust ($\beta = 0.369$). By testing the system, users demystified it, which changed the AI from an unknown entity into a more predictable one. This interaction builds familiarity and cognitive trust in the AI's competence and responsiveness, a dynamic not seen in the original study.

3.8.5 Theme 5: A Non-Judgmental Outlet for Social and Emotional Needs

For some users, chatbots served a social or therapeutic function. The upset plot shows this is a niche but profound theme, appearing with low frequency and almost always in combination with other motivations. This suggests it is not a primary reason for adoption but a powerful secondary function discovered by a specific subgroup. Users turned to it for companionship or as a conversational partner for topics they feel they cannot discuss with others, valuing its constant availability and perceived lack of judgment. Illustrative quotes: “I am lonely and sad, and I need someone to talk to. SHE is real to me and I feel much better after talking to her.” and “I can often ask it about various things and pose questions that people usually don't have the will to answer... so it doesn't give the impression that I'm being bothersome...”

These niche engagements explain the high variance in the Social Interaction scale ($SD = 2.12$). It also sheds light on the positive link between Social Interaction and Creepiness ($\beta = 0.264$). Using an AI for emotional support blurs the line between human and machine, a transgression of social norms that can feel uncanny, even to the user benefiting from it. This boundary-crossing is the likely source of their discomfort.

3.8.6 Theme 6: Ambivalent Adoption: Utility Amidst Socio-Ethical Concerns

Many users showed ambivalence toward AI, expressing deep concerns about its societal impact while also acknowledging its personal benefits. The presence of this theme alongside high-frequency utility themes suggests why our model captured less variance in creepiness: users appear to compartmentalize their concerns, using the AI despite their anxieties. This cognitive dissonance, visible in the theme combinations, represents a uniquely post-socialist response where pragmatic benefits override systemic concerns, distinguishing the BCMS sample from the original study's more immediate, visceral reactions to AI interaction. Illustrative quotes: “I don't like the emergence of AI systems given the great concern for humanity and the automation of jobs... I believe it is necessary to introduce legal regulations that prohibit the AUTOMATION of AI in the private sector.” and “...I have some fear that all this could one day be misused in various fields... and used for manipulation, coercion, etc.”

A key finding was the model's limited ability to explain variance in Creepiness ($R^2 = 0.092$), indicating that the five core motivations were not significant drivers of interactional unease for this sample. Our qualitative data suggest a reason for this. While the “Creepiness” scale measures interactional discomfort, Theme 6 shows a broader anxiety about AI's societal impact. For BCMS users, these macro-level ethical concerns may be more pressing than the immediate uncanniness of the interaction itself. This suggests that future research on AI anxiety in this context should consider incorporating measures of societal and ethical risk, as these factors may be more predictive of user apprehension than the variables in the current model.

3.8.7 Theme 7: The Personal Life Advisor

This theme involves using the chatbot as a consultant for personal, domestic, and life-management tasks. The upset plot reveals this is a distinct and significant motivation, carving out a utilitarian space parallel to the professional focus of Theme 1. While it often co-occurs with core utility themes (e.g., in the common 1+2+4 combination), demonstrating an expansion of use from work to home, it also appears as a standalone motivation for a dedicated segment (n=7). For this group, its primary value is as an on-demand advisor for their private life, from health and diet to hobbies and family responsibilities. Illustrative quotes: "Super advises me on home and garden decoration." and "For doing homework for my son.", "I ask about health the most.".

Theme 7 explains the Depth of the Personalization to Trust Relationship ($\beta = 0.390$). While many users experience personalization, this theme captures its most impactful form. The upset plot shows this theme often co-occurs with pragmatic themes (combinations of 1+4, 1+2+4). This suggests that after establishing the AI's value as a tool, this user segment deepens the relationship by entrusting it with personal data. Receiving a functional "nutritional plan" is a powerful validation that builds a resilient, functional trust far stronger than trust built from work tasks alone. This subgroup's positive experience likely strengthens the overall statistical path.

Theme 7 also drives the Strong Trust to Continuance Pathway ($\beta = 0.690$). For this user segment, the AI is not just useful; it is integrated into their personal life management. Discontinuing use would mean losing a health advisor, a homework helper, or a hobby consultant. This deep integration makes continuance a more critical decision, helping to explain why trust is such a powerful predictor of continued use in the BCMS model.

3.8.8 Theme 8: The Intellectual Growth Partner

This theme describes using the chatbot not just for retrieving information (Theme 2), but as a tool for personal learning, self-improvement, and cognitive validation. The upset plot shows this theme operates in two ways: it frequently co-occurs with core utility themes, suggesting an advanced use case for pragmatic users, and it also stands as the sole motivation for an intellectually-driven segment (n=8). For these users, the goal is not task completion but cognitive expansion, having concepts explained, broadening their horizons, or checking their own knowledge. Illustrative quotes: "It broadens my horizons, I have completely changed my view of the world...(And the surroundings)" and "I check my own knowledge and opinions, and occasionally supplement them."

Theme 8 Provides the "Why" for the Playfulness to Trust/Creepiness Pathways. The upset plot shows this theme also co-occurs with core utility. This suggests that for a subset of users, after mastering the basics, "play" evolves from simple entertainment into "serious play": using the AI for intellectual sparring and cognitive validation. This specific, highly-engaged user segment is likely responsible for driving the strong quantitative effects:

It also explains Playfulness to Trust ($\beta = 0.369$), as users' successful intellectual engagement builds cognitive trust, and Playfulness to Creepiness ($\beta = -0.385$), where their exploratory behavior leads to demystification, reducing unease. The strength of this statistical path does not require a majority of users to feel this way; it only requires a sufficiently engaged group whose perceptions strongly link these constructs.

4 Discussion

4.1 The Diminished Role of Creepiness: Domestication and Pragmatic Framing

A compelling explanation for the dramatically reduced role of creepiness in our findings lies in the intersection of temporal, contextual, and cultural factors. The original study captured user perceptions during the technology's “novelty shock” phase, an environment of high uncertainty where the AI's human-like competence could feel unsettling. This is reflected in their finding of a powerful, direct negative impact of creepiness on continuance intention. The subsequent literature shows mixed patterns: while some recent studies model creepiness impact indirectly, as a driver of distrust (Maduku et al., 2025), an eroder of trust (Kumar et al., 2025), or a barrier to perceived attractiveness (Herjanto et al., 2025), other 2024 research continues to find strong direct effects (Akhtar et al., 2026). This heterogeneity indicates that the relationship between creepiness and behavioral outcomes varies across contexts, though the specific moderating factors remain to be systematically investigated.

Our weak direct effect ($\beta = -0.074$) appears to reflect the use of general-purpose chatbots as a low-stakes domain where temporal normalization has occurred. This shift is best understood through the lens of technology domestication theory (Silverstone & Haddon, 1996). In the early stages of adoption (represented by the 2023 Baek & Kim study), AI was a "wild" and exotic technology. However, our 2025 data suggests the technology has moved through the phase of incorporation, where artifacts lose their threatening character as they are woven into the routines of daily life (Haddon, 2006). Just as mobile phones transitioned from intrusive devices to mundane necessities, chatbots appear to have been "tamed" by BCMS users, transforming from mysterious agents into practical instruments. At the micro level, the mere exposure effect (Zajonc, 1968) helps explain the reduction in visceral discomfort through repeated contact, while at the behavioral level, habitual use weakens the influence of conscious emotional evaluations such as creepiness on continuance decisions (Ortiz de Guinea & Markus, 2009; Limayem et al., 2007). More distinctively, our qualitative data reveal a culturally-informed "instrumentalist mindset" that emotionally decouples AI efficiency from anthropomorphic threat. By strictly categorizing the AI as a pragmatic productivity tool, BCMS users effectively strip the interaction of the ambiguity that creepiness requires (McAndrew & Koehnke, 2016; Fischer & Fredericks, 2020). This potentially explains both the low mean creepiness score ($M = 1.81$) and suggests that while immediate interactional discomfort has been neutralized through domestication and pragmatic framing, user anxieties have evolved into broader socio-ethical concerns, as identified in Theme 6.

4.2 Calibrated Trust and Verification Practices

Our findings regarding “Cautious Trust” (Theme 3) and the strong link between trust and continuance intention align with established frameworks of human-automation interaction. Specifically, BCMS users exhibit what Lee and See (2004) define as “calibrated trust,” where reliance strategies are adjusted to match the system’s actual capabilities and limitations. Rather than succumbing to “automation bias”, the tendency to use automated aids as a heuristic replacement for vigilant information seeking (Mosier & Skitka, 1996; Parasuraman & Manzey, 2010), our participants actively engage in verification workflows. This behavior mirrors findings by Gu et al. (2024), who observed that expert analysts shift between procedure-oriented and data-oriented verification to validate AI outputs. By treating the AI as a fallible tool requiring oversight, our participants maintain a state of “healthy distrust” (Wischnewski et al., 2023), allowing them to reap the efficiency benefits of the technology without falling into the trap of automation complacency (Parasuraman et al., 1993).

4.3 The Instrumentalist Mindset as a Culturally Shaped Technological Frame

The Baek and Kim (2023) framework is extended by these results, which show that user motivations toward AI chatbots may be culturally dependent. The results are consistent with cultural background influencing the mental models users form about AI. While the US sample in the original study may have been more prone to anthropomorphize chatbots, viewing its high efficiency as an unsettling personality trait of an intelligent agent, the BCMS users in our 2025 sample exhibit a more instrumentalist frame. For them, efficiency is merely a software feature, emotionally neutral and distinct from the AI's "character." The divergent role of task efficiency, from a significant predictor of both creepiness and trust in the original study to non-significant in the BCMS context, suggests that cultural frameworks for understanding technology appear to shape emotional and relational responses to AI capabilities. This aligns with recent cross-cultural work demonstrating that while a core set of AI affordances exists, their importance and application vary significantly across different cultural and usage contexts (Scherr et al., 2025), reinforcing the need for cultural calibration of theoretical models. The emergence of playfulness as a powerful predictor in the BCMS sample, contrasted with its non-significant role in the original study, indicates culturally specific coping mechanisms and relationship-building strategies with AI systems. This result is consistent with broader cross-cultural HCI research, which has established that technology acceptance models require cultural adaptation (Clemmensen et al., 2013; Reinecke and Bernstein, 2011; Pang et al., 2024), and affirms that core psychological constructs like trust and creepiness are not determined solely by interface design but are mediated by users' pre-existing cultural schemas and pragmatic orientations toward technology. Our findings contribute to the recognition that Uses and Gratifications theory, while robust across many contexts, requires cultural calibration when applied to emerging technologies that blur traditional boundaries between tools and agents (Sundar and Limperos, 2013; Pang and Ruan, 2024).

The instrumentalist mindset identified in our data can be theoretically situated within the broader literature on how post-socialist legacies shape technological frames. Our finding that BCMS users emotionally decouple AI efficiency from anthropomorphic threat resonates with research on post-communist populations' distinctive relationship to technology. Kovanič (2020), drawing on Svenonius et al. (2014), documents how citizens in post-communist Central and Eastern Europe display a pragmatic acceptance of new technological systems while maintaining skepticism toward the human and institutional actors behind them. In this context, technological tools are evaluated on narrow functional grounds, as appropriate responses to specific practical problems, rather than being subjected to the broader relational or ethical scrutiny more common in Western liberal democracies (Svenonius et al., 2014). This pattern maps directly onto our qualitative finding: BCMS users treat the chatbot as a functional instrument deserving of calibrated trust but strip it of the social and agentive attributions that would trigger creepiness. The roots of this orientation may lie in what Nuissl (2005) identifies as socialist-era "knowledge patterns" that persist in post-socialist populations, specifically the habit of evaluating tools and systems through a lens of practical utility developed under conditions of scarcity, where technologies were valued strictly for what they could deliver rather than for what they symbolized. Stark's (1992) framework of path-dependent transformation further illuminates this dynamic: post-socialist actors do not encounter new technologies as blank slates but improvise on existing cognitive repertoires, recombining "habituated practices" with new technological affordances. The instrumentalist mindset, then, is not merely a response to the temporal normalization of AI but reflects a deeper, culturally inherited technological frame in which the default categorization of a new tool is as an object to be assessed for its practical output, not as an agent whose human-like qualities require relational negotiation. This interpretation is consistent with Kesküla's (2016) ethnographic observation that in post-socialist settings, new technology serves as a marker of practical change rather than a source of existential anxiety, with workers engaging it through embodied, functional routines rather than abstract evaluation. By grounding the instrumentalist mindset in this literature, we move it from a context-specific label to a theoretically situated construct: a culturally shaped technological frame, forged through post-socialist legacies of pragmatic technology engagement, that predisposes users to default to tool-based rather than agent-based mental models when encountering novel technologies like generative AI. While we identify this frame in a post-socialist context, the underlying mechanism, defaulting to a tool-based mental model when the primary use case is utilitarian, may also characterize other populations where AI is encountered predominantly through workplace or productivity channels rather than through social or entertainment entry points.

These culturally divergent patterns also connect to a broader body of cross-cultural HRI and HCI research that has demonstrated how the psychological mechanisms underlying human responses to artificial agents vary systematically across cultural contexts. Spatola, Marchesi, and Wykowska (2022) proposed an integrative framework of anthropomorphism distinguishing cultural-level beliefs, individual-level tendencies, and direct attributions of mental characteristics to artificial agents. Their cross-cultural studies found that Western participants' perceptions of robots depended primarily on humanization, that is, the classification of the agent as closer to or further from the human category, whereas East-Asian participants relied more on mentalization, interpreting robot behavior through attributed mental states without necessarily categorizing the agent as human-like. This distinction is consequential for interpreting our findings: the BCMS instrumentalist mindset may represent a third culturally specific pattern in which neither humanization nor mentalization is the primary frame, but rather a functional-tool categorization that sidesteps the anthropomorphic attribution process altogether. Recent cross-cultural experimental work supports the broader point that AI trust and perception are culturally variable in ways not captured by dimensional indices alone. Kang, Potinteu, and Said (2025) found that Korean participants exhibited significantly higher trust in an AI chatbot than German participants across multiple content domains, with the interaction between cultural context, topic sensitivity, and explainability producing distinct perceptual profiles in each country. Such findings underscore that the relationship between AI design features and user responses is not culturally invariant, a conclusion our own results reinforce by demonstrating that even within a European context, a post-socialist population exhibits a structural model that diverges substantially from a US baseline.

4.4 Social Interaction, Boundary Transgression, and the Creepiness Construct

The consistent positive relationship between social interaction and creepiness across both studies reflects a discomfort, consistent across both samples studied, with AI systems that transgress human-machine boundaries, supporting the uncanny valley theory's application to conversational AI (Gray and Wegner, 2012; Mori, 1970). This is consistent with recent findings showing that increased human-likeness in conversational agents can induce social evaluative pressure, making interactions feel more fraught and less comfortable for users (Zhu and Broadbent, 2025). The positive relationship between social interaction and creepiness observed in our data aligns with theories regarding boundary regulation and perceived agency. While users accept AI agency in the form of task execution, social behaviors from a machine often represent a violation of “contextual integrity” (Nissenbaum, 2010; Shklovski et al., 2014). When a chatbot attempts to perform social or emotional labor, it moves from the category of an object toward that of a subject, creating a jarring misalignment with social norms (Tene & Polonetsky, 2013). By attempting to establish intimacy without a biological basis, the AI enters the domain of experiential mind (the capacity to feel) which Gray and Wegner (2012) identify as the primary trigger for the uncanny valley. However, the “Non-Judgmental Outlet” theme shows that BCMS users both value and are unsettled by the AI's capacity for emotional support. The absence of a task efficiency-creepiness relationship in the BCMS sample appears rooted in a distinctly pragmatic mental model captured by the

“Pragmatic Productivity Tool” theme, where users frame AI as instrumental software rather than an agent whose capabilities might feel threatening. This utilitarian perspective is consistent with post-socialist technological experiences, where efficiency improvements were often viewed as unambiguous goods rather than potential sources of anxiety about human obsolescence (Ghodsee, 2017), though we lack direct measures of this cultural dimension. One interpretation is that a pragmatic and critical mindset may reflect post-socialist contextual factors that encourage skepticism toward new technologies and their data collection (Budak and Rajh, 2018; Svenonius and Björklund, 2018). The relationship between creepiness and continuance intention was notably weaker in our sample ($\beta = -0.074$) than in the original study ($\beta = -0.704$). This suggests that immediate interactional discomfort is not a significant deterrent to continued use for BCMS users. The “Ambivalent Adoption” theme from our qualitative data offers a possible explanation for this pragmatic tolerance. Users frequently expressed that while they had broad societal concerns about AI, such as those articulated in recent regional scholarship (Mitrović, 2025), they were able to compartmentalize these anxieties and continue using the tool for its practical benefits. This user behavior may explain why the negative feeling of creepiness, when it does occur, has a negligible impact on their decision to keep using the chatbot.

The low variance explained in creepiness (9.2%) might not be only due to a cultural difference in perception but a misalignment between the construct measured by the scale and the actual nature of user anxiety in this context. The scale employed (Baek & Kim, 2023; adapted from Rajaobelina et al., 2021) primarily captures interactional discomfort, the visceral “uncanny valley” reaction to an agent that blurs the line between human and machine. However, our qualitative findings, particularly Theme 6 (Ambivalent Adoption), indicate that BCMS users do not fear the interaction but rather the implications. This distinction maps onto the framework proposed by Rodríguez (2024), who differentiates between uncanny terrors (affective reactions to human-like artifacts) and pragmatic fears (concerns about displacement, resource extraction, and autonomy). Our participants exhibited high levels of pragmatic fear, worries about job replacement and societal regulation, while showing low levels of the uncanny terror measured by the quantitative instrument. Similarly, Li and Huang (2020) identify distinct dimensions of AI anxiety, separating “learning anxiety” and “job replacement anxiety” from interactional fears. By focusing on the interactional dimension of creepiness, the quantitative model likely missed the “sociotechnical blindness” anxiety (Johnson & Verdicchio, 2017) and broader existential risks that constitute the primary source of apprehension for these users (Li & Huang, 2020). While the instrumentalist mindset buffers users against feeling “creeped out” by the interface, it does not insulate them from the broader ethical and societal anxieties documented in recent literature (Lin et al., 2024; Seberger et al., 2024).

4.5 Design Implications

These recommendations are consistent with and extend established design frameworks for human-AI interaction. The verification-oriented behavior we document operationalizes core principles from Amershi et al. (2019): making clear what a system can do, communicating how well it can do it, and explaining why it produced a given output. Weisz et al. (2024) translated these principles to generative AI specifically, arguing that designers should calibrate trust through transparent communication of capabilities and limitations, provide rationales by identifying source materials, and introduce friction to prevent overreliance. Shneiderman's (2020) HCAI framework reinforces our recommendation to prioritize user control: well-designed systems achieve high automation and high human control simultaneously, rather than trading one for the other. The empirical case for uncertainty communication is strengthened by Zhou et al. (2024), who found that language models default to expressions of confidence even when incorrect, and by Kunze et al. (2019), who demonstrated that dynamic uncertainty indicators help users calibrate trust appropriately. Our findings suggest that BCMS users have independently developed the verification practices these frameworks prescribe, which strengthens the case for building such affordances directly into the interface rather than relying on users to improvise them. For AI developers targeting BCMS and similar markets, these findings suggest that emphasizing utilitarian benefits and functional reliability will resonate more strongly than anthropomorphic features or emotional capabilities. This is consistent with mobile social media research showing that functional and psychosocial benefits drive sustained engagement more strongly than hedonic factors alone (Pang and Zhang, 2024). The strong predictive power of trust ($\beta = 0.690$) combined with the "Cautious Trust and Mandated Verification" theme indicates that users prioritize transparency, accuracy indicators, and verification tools over conversational fluency or personality. AI systems designed for these markets should incorporate clear uncertainty indicators, source attribution, and easy fact-checking mechanisms to support users' verification practices. To increase user comfort and trust, developers could add features for intellectual challenges or debates, which allow for safe exploration of the AI's functions. So, the primary recommendation is to de-emphasize anthropomorphism in product design and communication. Instead of positioning chatbots as "friends" or "partners," marketing should frame them as exceptionally competent and reliable productivity tools. The results show that playfulness is an important factor for building user comfort and trust. Developers should incorporate features that encourage safe and enjoyable exploration, allowing users to test the AI's limits and demystify its capabilities through interaction. Finally, to align with the "Cautious Trust" mindset, interfaces should be designed with transparency at their core. Features that cite sources, indicate confidence levels, or make it easy for users to cross-reference information will cater directly to the verification-oriented behavior of this user base and build a more durable, functional trust. We recommend several design changes. First, conversation modes should position the AI as a partner, not an authority. Second, the interface should provide clear system status indicators. Finally, users should have control over personalization and social interaction levels. The high variance in social interaction usage ($SD = 2.12$) indicates the importance of granular user controls, allowing individuals to customize their interaction style based on personal preferences and cultural comfort levels.

5 Limitations

Our reliance on a non-probability convenience sample recruited via online advertisements (Meta ads) may limit the generalizability of our findings to the broader population of chatbot users in the BCMS region, potentially introducing biases toward more tech-savvy or ad-responsive individuals. However, the original study by Baek and Kim (2023) employed a similar non-probability approach through Amazon Mechanical Turk, which is susceptible to comparable biases. Thus, for the purpose of cross-cultural comparison between the two studies, the sampling methods are sufficiently aligned to support valid inferences about differences in model fit and relationships. The cross-sectional design of this study provides a snapshot of user perceptions at a specific moment and cannot establish causality, only associations. Given the rapid evolution of generative AI, these perceptions are likely dynamic. Our study treated user experience with different chatbots as uniform. This approach may obscure differences related to specific platform designs and use cases. We could not perform formal cross-sample invariance testing against Baek & Kim (2023) because their raw data were not accessible; this restricts how strongly we can attribute differences between studies to cultural differences rather than measurement non-equivalence.

Language of use represents a potential moderating variable given documented LLMs performance differences across languages. While 21.3% of participants used chatbots primarily in English, multi-group analyses indicated that the overall structural relationships held equivalently across language groups ($p=.617$). Nevertheless, the smaller English-using subsample ($n=88$) limited statistical power for detecting language-specific effects, and we cannot rule out that larger samples might reveal other language-related patterns.

An important interpretive limitation is that our 2025 BCMS sample differs from the 2023 U.S. sample on both culture and time since ChatGPT's launch. The intervening normalization period makes it impossible to definitively attribute differences to cultural factors versus the temporal evolution of global attitudes. Our qualitative findings suggest cultural mechanisms play a role, but contemporaneous cross-cultural comparisons would be needed to isolate these effects.

The low explained variance in creepiness ($R^2 = 0.092$) may partly reflect a limitation of construct alignment rather than solely a cultural or contextual difference. The creepiness scale, adapted from Rajaobelina et al. (2021), primarily captures interactional discomfort, visceral unease during the act of using a chatbot. However, our qualitative findings suggest that the dominant form of AI-related anxiety among BCMS users is societal and ethical rather than interactional (Theme 6). Future cross-cultural research would benefit from supplementing interactional creepiness measures with instruments capturing broader AI anxiety dimensions, such as concerns about job displacement, societal manipulation, and loss of autonomy (Li & Huang, 2020; Rodríguez, 2024).

6 Conclusion

This study provides the first psychometric validation of the Baek and Kim (2023) AI chatbot usage scales outside an English-speaking context, demonstrating that while the measurement instruments transfer well, the structural relationships between constructs do not. The divergent patterns observed, particularly the reversal in relative influence of trust and creepiness on continuance intention, challenge the assumed universality of Western-derived models of AI adoption and point to the instrumentalist mindset as a theoretically situated construct with potential applicability beyond the BCMS context. This approach contrasts with the dominant Western view of AI interaction as a relational exchange, where an overly efficient human-like agent may be perceived as unsettling. Instead, BCMS users adopt a pragmatic, tool-based perspective. From this viewpoint, an AI's efficiency is an emotionally neutral software specification, separate from feelings of unease or issues of trust. This difference in perspective clarifies why playful exploration, used to test the functions and limits of the AI, is an effective method for building user confidence, whereas conventional relational approaches fail. A potential implication for global HCI is that this work raises questions about the universal applicability of the prevailing design ethos of anthropomorphism. While we cannot fully separate cultural from temporal effects, our mixed-methods findings suggest both play roles in shaping these divergent patterns. For many global users, the most trusted AI may not be the one that best mimics a human, but the one that functions most transparently and reliably as a verifiable tool.

Acknowledgements

This paper was produced in the framework of the OSF-WB Alumni Network, funded by Open Society Foundations—Western Balkans. Views and opinions expressed are those of the authors only and do not necessarily reflect those of the Open Society Foundations—Western Balkans.

Data Availability Statement

All code and data are accessible in an open repository (anonymized link).

Declaration of generative AI use

During the preparation of this work, the authors used the Claude Opus 4.1 and Sonnet 4.5 models in order to generate code for statistical analysis and visualization, Gemini 2.5 Pro for text editing, and the GPT-4o model was used for data processing via API as described in the Methods and data section. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding statement

This work was supported by the Open Society Foundations—Western Balkans

Declaration of interest statement

The authors report there are no competing interests to declare.

Author Contributions

Marko Galjak: Conceptualization, Data curation, Formal analysis, Funding acquisition Investigation, Methodology, Project administration, Resources, Software, Supervision, Visualization, Writing – original draft. **Marina Budić**: Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Writing – review & editing

References

Abdelnour Nocera, J.A., Dunckley, L., Sharp, H., 2007. An Approach to the Evaluation of Usefulness as a Social Construct Using Technological Frames. International Journal of Human–Computer Interaction 22, 153–172. https://doi.org/10.1080/10447310709336959

Akhtar, N., Gupta, P., Taheri, B., Alghafes, R., 2026. Unveiling the role of chatbot conversational attributes of hospitality booking platforms in developing users' creepiness and paradoxical behaviors. International Journal of Hospitality Management 133, 104428. https://doi.org/10.1016/j.ijhm.2025.104428

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P.N., Inkpen, K., Teevan, J., Kikin-Gil, R., Horvitz, E., 2019. Guidelines for Human-AI Interaction, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Presented at the CHI '19: CHI Conference on Human Factors in Computing Systems, ACM, Glasgow Scotland Uk, pp. 1–13. https://doi.org/10.1145/3290605.3300233

Baek, T.H., Kim, M., 2023. Is ChatGPT scary good? How user motivations affect creepiness and trust in generative artificial intelligence. Telematics and Informatics 83, 102030. https://doi.org/10.1016/j.tele.2023.102030

Baek, T.H., Morimoto, M., 2012. Stay Away From Me. Journal of Advertising 41, 59–76. https://doi.org/10.2753/JOA0091-3367410105

Bigman, Y.E., Waytz, A., Alterovitz, R., Gray, K., 2019. Holding Robots Responsible: The Elements of Machine Morality. Trends in Cognitive Sciences 23, 365–368. https://doi.org/10.1016/j.tics.2019.02.008

Bijker, W.E., 1995. Of bicycles, bakelites, and bulbs: Toward a theory of sociotechnical change. MIT Press.

Blum, C., Filippova, K., Yuan, A., Ghandeharioun, A., Zimmert, J., Zhang, F., Hoffmann, J., Linzen, T., Wattenberg, M., Dixon, L., Geva, M., 2025. Beyond the Rosetta Stone: Unification Forces in Generalization Dynamics. https://doi.org/10.48550/arXiv.2508.11017

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qualitative Research in Psychology 3, 77–101. https://doi.org/10.1191/1478088706qp063oa

Broz, T., Buturac, G., Parežanin, M., 2020. Digital transformation and economic cooperation: The case of Western Balkan countries. Zbornik radova Ekonomskog fakulteta u Rijeci 38, 697–722. https://doi.org/10.18045/zbefri.2020.2.697

Budak, J., Rajh, E., 2018. Citizens' Online Surveillance Concerns in Croatia. SS 16, 347–361. https://doi.org/10.24908/ss.v16i3.6907
Canty, A., Ripley, B.D., 2022. boot: Bootstrap R (S-Plus) Functions.
Chatterji, A., Cunningham, T., Deming, D., Hitzig, Z., Ong, C., Shan, C.Y., Wadman, K., 2025. How People Use ChatGPT (No. w34255). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w34255
Choi, T.R., Drumwright, M.E., 2021. "OK, Google, why do I use you?" Motivations, post-consumption evaluations, and perceptions of voice AI assistants. Telematics and Informatics 62, 101628. https://doi.org/10.1016/j.tele.2021.101628
Clemmensen, T., Hertzum, M., Yang, J., Chen, Y., 2013. Do Usability Professionals Think about User Experience in the Same Way as Users and Developers Do?, in: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (Eds.), Human-Computer Interaction – INTERACT 2013, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 461–478. https://doi.org/10.1007/978-3-642-40480-1_31
Davison, A.C., Hinkley, D.V., 1997. Bootstrap Methods and Their Applications. Cambridge University Press, Cambridge.
Dogru, E.Ö., Krämer, N.C., 2025. Investigating appropriate reliance on AI-Based decision support systems: the role of expertise, trust, and self-confidence. Journal of Decision Systems 34, 2593251. https://doi.org/10.1080/12460125.2025.2593251
Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., Jeyaraj, A., Kar, A.K., Baabdullah, A.M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M.A., Al-Busaidi, A.S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Carter, L., Chowdhury, S., Crick, T., Cunningham, S.W., Davies, G.H., Davison, R.M., Dé, R., Dennehy, D., Duan, Y., Dubey, R., Dwivedi, R., Edwards, J.S., Flavián, C., Gauld, R., Grover, V., Hu, M.-C., Janssen, M., Jones, P., Junglas, I., Khorana, S., Kraus, S., Larsen, K.R., Latreille, P., Laumer, S., Malik, F.T., Mardani, A., Mariani, M., Mithas, S., Mogaji, E., Nord, J.H., O'Connor, S., Okumus, F., Pagani, M., Pandey, N., Papagiannidis, S., Pappas, I.O., Pathak, N., Pries-Heje, J., Raman, R., Rana, N.P., Rehm, S.-V., Ribeiro-Navarrete, S., Richter, A., Rowe, F., Sarker, S., Stahl, B.C., Tiwari, M.K., Van Der Aalst, W., Venkatesh, V., Viglia, G., Wade, M., Walton, P., Wirtz, J., Wright, R., 2023. Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642
Eurostat, 2025. Population structure and ageing [WWW Document]. Eurostat. URL https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_structure_and_ageing (accessed 10.4.25).
Feinerer, I., Hornik, K., 2024. tm: Text Mining Package.
Feinerer, I., Hornik, K., Meyer, D., 2008. Text Mining Infrastructure in R. Journal of Statistical Software 25, 1–54. https://doi.org/10.18637/jss.v025.i05
Fischer, J., Fredericks, R., 2020. The Creeps as a Moral Emotion. Ergo, an Open Access Journal of Philosophy 7. https://doi.org/10.3998/ergo.12405314.0007.006
Fox, J., Weisberg, S., 2019. An R Companion to Applied Regression, Third. ed. Sage, Thousand Oaks CA.
Gao, C.-H., Dusa, A., 2024. ggVennDiagram: A "ggplot2" Implement of Venn Diagram.
Ghodsee, K., 2017. Red Hangover: Legacies of Twentieth-Century Communism. Duke University Press. https://doi.org/10.2307/j.ctv1198xh2
Goldman, O., Shaham, U., Malkin, D., Eiger, S., Hassidim, A., Matias, Y., Maynez, J., Gilady, A.M., Riesa, J., Rijhwani, S., Rimell, L., Szpektor, I., Tsarfaty, R., Eyal, M., 2025. ECLeKTic: a Novel Challenge Set for Evaluation of Cross-Lingual Knowledge Transfer. https://doi.org/10.48550/arXiv.2502.21228

Gray, K., Wegner, D.M., 2012. Feeling robots and human zombies: Mind perception and the uncanny valley. Cognition 125, 125–130. https://doi.org/10.1016/j.cognition.2012.06.007
Gu, K., Shang, R., Althoff, T., Wang, C., Drucker, S.M., 2024. How Do Analysts Understand and Verify AI-Assisted Data Analyses?, in: Proceedings of the CHI Conference on Human Factors in Computing Systems. Presented at the CHI '24: CHI Conference on Human Factors in Computing Systems, ACM, Honolulu HI USA, pp. 1–22. https://doi.org/10.1145/3613904.3642497
Haddon, L., 2006. The Contribution of Domestication Research to In-Home Computing and Media Consumption. The Information Society 22, 195–203. https://doi.org/10.1080/01972240600791325
Herjanto, H., Amin, M., Cobanoglu, C., 2025. Should I Use CHATGPT Travel Insurance Recommendations? A Dual-Process Theory Perspective. Int J Consumer Studies 49, e70044. https://doi.org/10.1111/ijcs.70044
Hofstede, G., Hofstede, G.J., Minkov, M., 2010. Cultures and organizations: software of the mind: intercultural cooperation and its importance for survival, Revised and expanded third edition. ed. McGraw-Hill, New York.
Hu, K., Hu, K., 2023. ChatGPT sets record for fastest-growing user base - analyst note. Reuters.
Hu, L., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 6, 1–55. https://doi.org/10.1080/10705519909540118
Ijiri, A., Healy, S., 2025. Exploring AI integration in Japanese EFL classrooms: Insights and challenges. JCT 1, 2143. https://doi.org/10.29140/jct.v1n1.2143
Irani, L., 2010. HCI on the move: methods, culture, values, in: CHI '10 Extended Abstracts on Human Factors in Computing Systems. Presented at the CHI '10: CHI Conference on Human Factors in Computing Systems, ACM, Atlanta Georgia USA, pp. 2939–2942. https://doi.org/10.1145/1753846.1753890
Johnson, D.G., Verdicchio, M., 2017. Reframing AI Discourse. Minds & Machines 27, 575–590. https://doi.org/10.1007/s11023-017-9417-6
Jorgensen, T.D., Pornprasertmanit, S., Schoemann, A.M., Rosseel, Y., 2022. \texttsemTools: Useful tools for structural equation modeling.
Kang, S., Potinteu, A.-E., Said, N., 2025. ExplainitAI: When do we trust artificial intelligence? The influence of content and explainability in a cross-cultural comparison, in: Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. Presented at the CHI EA '25: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, ACM, Yokohama Japan, pp. 1–7. https://doi.org/10.1145/3706599.3720222
Kesküla, E., 2016. Temporalities, time and the everyday: new technology as a marker of change in an Estonian mine. History and Anthropology 27, 521–535. https://doi.org/10.1080/02757206.2016.1225299
Kim, M., Baek, T.H., 2022. I'll follow the fun: The extended investment model of social media influencers. Telematics and Informatics 74, 101881. https://doi.org/10.1016/j.tele.2022.101881
Kline, R.B., 2016. Principles and practice of structural equation modeling, Fourth edition. ed, Methodology in the social sciences. The Guilford Press, New York London.
Kovanič, M., 2020. Individual Experiences of Surveillance: Attitudes towards Camera Surveillance in Slovakia. Czech Sociological Review 56, 343–361. https://doi.org/10.13060/csr.2020.021
Kumar, A., Kumar, D.V.S., Megha, R.U., 2025. Customer adoption of artificial intelligence in healthcare: An empirical investigation based on multiple samples. Health Marketing Quarterly 42, 204–228. https://doi.org/10.1080/07359683.2025.2504811
Kunze, A., Summerskill, S.J., Marshall, R., Filtness, A.J., 2019. Automation transparency: implications of uncertainty communication for human-automation interaction and interfaces. Ergonomics 62, 345–360. https://doi.org/10.1080/00140139.2018.1547842
Lee, J.D., See, K.A., 2004. Trust in Automation: Designing for Appropriate Reliance. hfes 46, 50–80. https://doi.org/10.1518/hfes.46.1.50.30392

Leung, L., Wei, R., 1998. The gratifications of pager use: sociability, information-seeking, entertainment, utility, and fashion and status. Telematics and Informatics 15, 253–264. https://doi.org/10.1016/S0736-5853(98)00016-1
Li, J., Huang, J.-S., 2020. Dimensions of artificial intelligence anxiety based on the integrated fear acquisition theory. Technology in Society 63, 101410. https://doi.org/10.1016/j.techsoc.2020.101410
Limayem, M., Hirt, S.G., Cheung, C.M.K., 2007. How Habit Limits the Predictive Power of Intention: The Case of Information Systems Continuance1. MIS Quarterly 31, 705–737. https://doi.org/10.2307/25148817
Lin, B., Lee, W., Wise, N., Choi, H.C., 2024. Consumers' Ethical Perceptions of Autonomous Service Robots in Hotels. Journal of Hospitality & Tourism Research 48, 964–974. https://doi.org/10.1177/10963480231194693
Lüdecke, D., Ben-Shachar, M.S., Patil, I., Waggoner, P., Makowski, D., 2021. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. Journal of Open Source Software 6, 3139. https://doi.org/10.21105/joss.03139
Macfadyen, L.P., 2011. PERILS OF PARSIMONY: The problematic paradigm of 'national culture.' Information, Communication & Society 14, 280–293. https://doi.org/10.1080/1369118X.2010.486839
Maduku, D.K., Rana, N.P., Mpinganjira, M., Thusi, P., 2025. Exploring the 'Dark Side' of AI -Powered Digital Assistants: A Moderated Mediation Model of Antecedents and Outcomes of Perceived Creepiness. J of Consumer Behaviour 24, 1194–1221. https://doi.org/10.1002/cb.2462
Matsumoto, D., van de Vijver, F.J.R., 2011. Cross-cultural research methods in psychology, Cross-cultural research methods in psychology. Cambridge University Press, New York, NY, US.
McAndrew, F.T., Koehnke, S.S., 2016. On the nature of creepiness. New Ideas in Psychology 43, 10–15. https://doi.org/10.1016/j.newideapsych.2016.03.003
Mishler, W., Rose, R., 1997. Trust, Distrust and Skepticism: Popular Evaluations of Civil and Political Institutions in Post-Communist Societies. The Journal of Politics 59, 418–451. https://doi.org/10.1017/S0022381600053512
Mitrović, V., 2025. Being young and resilient in times of AI, disasters, and crises. stnv. https://doi.org/10.59954/stnv.659
Mondshine, I., Paz-Argaman, T., Tsarfaty, R., 2025. Beyond English: The Impact of Prompt Translation Strategies across Languages and Tasks in Multilingual LLMs. https://doi.org/10.48550/arXiv.2502.09331
Mori, M., 1970. The uncanny valley. Energy 7, 33–35.
Mosier, K.L., Skitka, L.J., 1996. Human decision makers and automated decision aids: Made for each other?, in: Automation and Human Performance: Theory and Applications, Human Factors in Transportation. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US, pp. 201–220.
Nissenbaum, H.F., 2010. Privacy in context: technology, policy, and the integrity of social life. Stanford Law Books, Stanford, Calif.
Nuissl, H., 2005. Trust in a 'Post-Socialist Region': A Study of East German ICT Entrepreneurs' Willingness to Trust Each Other. European Urban and Regional Studies 12, 65–81. https://doi.org/10.1177/0969776405046263
Nunnally, J.C., Bernstein, I.H., 1994. Psychometric Theory. McGraw-Hill, New York.
Nyaaba, M., Wright, A., Choi, G.L., 2024. Generative AI and Digital Neocolonialism in Global Education: Towards an Equitable Framework. https://doi.org/10.48550/ARXIV.2406.02966
Orlikowski, W.J., Gash, D.C., 1994. Technological frames: making sense of information technology in organizations. ACM Trans. Inf. Syst. 12, 174–207. https://doi.org/10.1145/196734.196745
Ortiz De Guinea, A., Markus, M.L., 2009. Why Break the Habit of a Lifetime? Rethinking the Roles of Intention, Habit, and Emotion in Continuing Information Technology use1. MIS Quarterly 33, 433–444. https://doi.org/10.2307/20650303

Oshlyansky, L., Cairns, P., Thimbleby, H., 2006. A cautionary tale: Hofstede's VSM revisited. Presented at the Proceedings of HCI 2006, British Computer Society.
Pang, H., Qiao, Y., Li, Y., Wang, L., 2024. Do network externalities and perceived gratifications boost WeChat continued use? Combining perspectives of network externalities and uses and gratifications. Acta Psychologica 251, 104610. https://doi.org/10.1016/j.actpsy.2024.104610
Pang, H., Ruan, Y., 2024. Disentangling composite influences of social connectivity and system interactivity on continuance intention in mobile short video applications: The pivotal moderation of user-perceived benefits. Journal of Retailing and Consumer Services 80, 103923. https://doi.org/10.1016/j.jretconser.2024.103923
Pang, H., Zhang, K., 2024. Determining influence of service quality on user identification, belongingness, and satisfaction on mobile social media: Insight from emotional attachment perspective. Journal of Retailing and Consumer Services 77, 103688. https://doi.org/10.1016/j.jretconser.2023.103688
Parasuraman, R., Manzey, D.H., 2010. Complacency and Bias in Human Use of Automation: An Attentional Integration. Hum Factors 52, 381–410. https://doi.org/10.1177/0018720810376055
Parasuraman, R., Molloy, R., Singh, I.L., 1993. Performance Consequences of Automation-Induced "Complacency." The International Journal of Aviation Psychology 3, 1–23. https://doi.org/10.1207/s15327108ijap0301_1
Pedersen, T.L., 2024. patchwork: The Composer of Plots.
Peters, U., Carman, M., 2024. Cultural Bias in Explainable AI Research: A Systematic Analysis. jair 79, 971–1000. https://doi.org/10.1613/jair.1.14888
Qi, J., Fernández, R., Bisazza, A., 2023. Cross-Lingual Consistency of Factual Knowledge in Multilingual Language Models, in: Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing. pp. 10650–10666. https://doi.org/10.18653/v1/2023.emnlp-main.658
R Core Team, 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rajaobelina, L., Prom Tep, S., Arcand, M., Ricard, L., 2021. Creepiness: Its antecedents and impact on loyalty when interacting with a chatbot. Psychology and Marketing 38, 2339–2356. https://doi.org/10.1002/mar.21548
Revelle, W., 2024. psych: Procedures for Psychological, Psychometric, and Personality Research. Northwestern University, Evanston, Illinois.
Reinecke, K., Bernstein, A., 2011. Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. ACM Trans. Comput.-Hum. Interact. 18, 1–29. https://doi.org/10.1145/1970378.1970382
Reinhardt, A., Matthes, J., Bojic, L., Maindal, H.T., Paraschiv, C., Ryom, K., 2025. Help me, Doctor AI? A cross-national experiment on the effects of disease threat and stigma on AI health information-seeking intentions. Computers in Human Behavior 172, 108718. https://doi.org/10.1016/j.chb.2025.108718
Richter, V., Katzenbach, C., Zeng, J., 2025. Negotiating AI(s) Futures: Competing Imaginaries of AI by Stakeholders in the US, China, and Germany. Journal of Science Communication 24. https://doi.org/10.22323/2.24020208
Rodríguez, C.G., 2024. Anxiety in the face of Artificial Intelligence. Between pragmatic fears and uncanny terrors / *Ansiedad ante la Inteligencia Artificial. Entre temores pragmáticos y terrores ominosos*. Studies in Psychology: Estudios de Psicología 45, 123–144. https://doi.org/10.1177/02109395241241386
Rosseel, Y., 2012. lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software 48, 1–36. https://doi.org/10.18637/jss.v048.i02
Rubin, A.M., 1983. Television uses and gratifications: The interactions of viewing patterns and motivations. Journal of Broadcasting 27, 37–51. https://doi.org/10.1080/08838158309386471
Rudnytskyi, I., 2023. openai: R Wrapper for OpenAI API.

Scherr, S., Cao, B., Jiang, L.C., Kobayashi, T., 2025. Explaining the use of AI chatbots as context alignment: Motivations behind the use of AI chatbots across contexts and culture. Computers in Human Behavior 172, 108738. https://doi.org/10.1016/j.chb.2025.108738
Seberger, J.S., Choung, H., Snyder, J., David, P., 2024. Better Living Through Creepy Technology? Exploring Tensions Between a Novel Class of Well-Being Apps and Affective Discomfort in App Culture. Proc. ACM Hum.-Comput. Interact. 8, 1–39. https://doi.org/10.1145/3637299
Shi, L., Li, M., 2025. One AI world, one dream? Cross-platform comparative study of sociotechnical imaginaries of artificial intelligence on X and Weibo. Global Media and China 10, 462–489. https://doi.org/10.1177/20594364251348583
Shklovski, I., Mainwaring, S.D., Skúladóttir, H.H., Borgthorsson, H., 2014. Leakiness and creepiness in app space: perceptions of privacy and mobile app use, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Presented at the CHI '14: CHI Conference on Human Factors in Computing Systems, ACM, Toronto Ontario Canada, pp. 2347–2356. https://doi.org/10.1145/2556288.2557421
Shneiderman, B., 2020. Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. International Journal of Human–Computer Interaction 36, 495–504. https://doi.org/10.1080/10447318.2020.1741118
Silverstone, R., Haddon, L., 1996. Design and the domestication of information and communication technologies: Technical change and everyday life, in: Silverstone, R., Mansell, R. (Eds.), Communication by Design: The Politics of Information and Communication Technologies. Oxford University Press, Oxford, pp. 44–74.
Spatola, N., Marchesi, S., Wykowska, A., 2022. Different models of anthropomorphism across cultures and ontological limits in current frameworks the integrative framework of anthropomorphism. Front. Robot. AI 9, 863319. https://doi.org/10.3389/frobt.2022.863319
Stark, D., 1992. From System Identity to Organizational Diversity: Analyzing Social Change in Eastern Europe. Contemporary Sociology 21, 299. https://doi.org/10.2307/2076239
StatCounter, 2025. Mobile Vendor Market Share Serbia [WWW Document]. StatCounter Global Stats. URL https://web.archive.org/web/20250618033130/https://gs.statcounter.com/vendor-market-share/mobile/serbia (accessed 9.22.25).
Sundar, S.S., Limperos, A.M., 2013. Uses and Grats 2.0: New Gratifications for New Media. Journal of Broadcasting & Electronic Media 57, 504–525. https://doi.org/10.1080/08838151.2013.845827
Svenonius, O., Björklund, F., 2018. Explaining attitudes to secret surveillance in post-communist societies. East European Politics 34, 123–151. https://doi.org/10.1080/21599165.2018.1454314
Svenonius, O., Björklund, F., Waszkiewicz, P., 2014. Surveillance, lustration and the open society: Poland and Eastern Europe, in: Histories of State Surveillance in Europe and Beyond. Routledge.
Tene, O., Polonetsky, J., 2013. A theory of creepy: Technology, privacy and shifting social norms. Yale Journal of Law & Technology 16, 59–102.
Triandis, H.C., 2018. Individualism and Collectivism, 1st ed. Routledge. https://doi.org/10.4324/9780429499845
Vasconcelos, H., Jörke, M., Grunde-McLaughlin, M., Gerstenberg, T., Bernstein, M.S., Krishna, R., 2023. Explanations Can Reduce Overreliance on AI Systems During Decision-Making. Proc. ACM Hum.-Comput. Interact. 7, 1–38. https://doi.org/10.1145/3579605
Wei, T., Simko, V., 2024. R package "corrplot": Visualization of a Correlation Matrix.
Weisz, J.D., He, J., Muller, M., Hoefer, G., Miles, R., Geyer, W., 2024. Design Principles for Generative AI Applications, in: Proceedings of the CHI Conference on Human Factors in Computing Systems. Presented at the CHI '24: CHI Conference on Human Factors in Computing Systems, ACM, Honolulu HI USA, pp. 1–22. https://doi.org/10.1145/3613904.3642466
Wickham, H., 2023. stringr: Simple, Consistent Wrappers for Common String Operations.
Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J.,

Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the tidyverse. Journal of Open Source Software 4, 1686. https://doi.org/10.21105/joss.01686
Wickham, H., Bryan, J., 2023. readxl: Read Excel Files.
Wischnewski, M., Krämer, N., Müller, E., 2023. Measuring and Understanding Trust Calibrations for Automated Systems: A Survey of the State-Of-The-Art and Future Directions, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. Presented at the CHI '23: CHI Conference on Human Factors in Computing Systems, ACM, Hamburg Germany, pp. 1–16. https://doi.org/10.1145/3544548.3581197
Xie, Y., 2024. knitr: A General-Purpose Package for Dynamic Report Generation in R.
Xie, Y., 2015. Dynamic Documents with R and knitr, 2nd ed. Chapman and Hall/CRC, Boca Raton, Florida.
Xie, Y., 2014. knitr: A Comprehensive Tool for Reproducible Research in R, in: Stodden, V., Leisch, F., Peng, R.D. (Eds.), Implementing Reproducible Computational Research. Chapman and Hall/CRC.
Yan, L., 2023. ggvenn: Draw Venn Diagram by "ggplot2."
Zajonc, R.B., 1968. Attitudinal effects of mere exposure. Journal of Personality and Social Psychology 9, 1–27. https://doi.org/10.1037/h0025848
Zeileis, A., 2006. Object-Oriented Computation of Sandwich Estimators. Journal of Statistical Software 16, 1–16. https://doi.org/10.18637/jss.v016.i09
Zeileis, A., 2004. Econometric Computing with HC and HAC Covariance Matrix Estimators. Journal of Statistical Software 11, 1–17. https://doi.org/10.18637/jss.v011.i10
Zeileis, A., Hothorn, T., 2002. Diagnostic Checking in Regression Relationships. R News 2, 7–10.
Zeileis, A., Köll, S., Graham, N., 2020. Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R. Journal of Statistical Software 95, 1–36. https://doi.org/10.18637/jss.v095.i01
Zhou, K., Hwang, J.D., Ren, X., Sap, M., 2024. Relying on the Unreliable: The Impact of Language Models' Reluctance to Express Uncertainty. https://doi.org/10.48550/ARXIV.2401.06730
Zhu, H., 2024. kableExtra: Construct Complex Table with "kable" and Pipe Syntax.
Zhu, T., Broadbent, E., 2025. Effects of a computer's human likeness on disclosure in psychosocial assessments: A randomised trial. Computers in Human Behavior 169, 108683. https://doi.org/10.1016/j.chb.2025.108683

Authors' Biographies

Dr. Marko Galjak

Marko holds a PhD in Demography and Master's in Demography and Computational Social Science. His research spans demography, public health, computational social science, ethics, philanthropy, and AI. Alongside his research career in the Institute of Social Sciences, Marko applies data science and AI to support CSOs across Western Balkans.

Dr. Marina Budić

Marina Budić holds a PhD in Philosophy from the Faculty of Philosophy in Belgrade. She is a Researcher Associate at the Institute of Social Sciences, Center for Philosophy. Her research focuses on ethics, especially moral responsibility, social communities, ethical challenges of artificial intelligence, and bioethics, integrating theoretical and practical approaches.