The Use of Large Language Models in HCI: A Critical Analysis of Synthetic Users

Joni Salminen
University of Vaasa
Vaasa, Finland
jonisalm@uwasa.fi

Danial Amin
University of Vaasa
Vaasa, Finland
danialam@uwasa.fi

Soon-Gyo Jung
Qatar Computing Research Institute, Hamad Bin Khalifa University
Doha, Qatar
sjung@hbku.edu.qa

Bernard Jansen
Qatar Computing Research Institute, Hamad Bin Khalifa University
Doha, Qatar
jjansen@acm.org

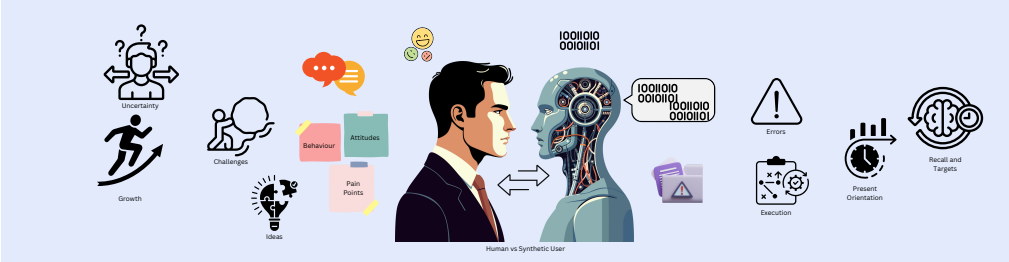

Figure 1: Contrasting Human and Synthetic Users in HCI Research: While synthetic users attempt to replicate human behaviors and attitudes, there are fundamental methodological questions about their validity in user research.

Abstract

While many user researchers remain skeptical of synthetic users generated by large language models (LLMs), their adoption is growing in industry. This conceptual article investigates the root causes driving synthetic user adoption, maps potential use cases, and identifies key risks. Our inquiry reveals that while synthetic users emerge from legitimate pressures in user research, they present fundamental methodological and epistemological challenges. Their normative biases and inability to generate novel insights make them particularly problematic for user research activities like usability testing and user interviews. However, the emergence of synthetic users underlines challenges in user research, including resource constraints, privacy concerns, and difficulties in demonstrating the return on investment. Rather than simply dismissing synthetic users, we argue that understanding them as a *symptom* of these underlying challenges can inform efforts to strengthen the HCI research methodology.

CCS Concepts

• **Human-centered computing** → **Interactive systems and tools**.

Keywords

synthetic users, artificial users, large language models in HCI

ACM Reference Format:
Joni Salminen, Danial Amin, Soon-Gyo Jung, and Bernard Jansen. 2025. The Use of Large Language Models in HCI: A Critical Analysis of Synthetic Users. In *The Augmented Humans International Conference 2025 (AHs 2025), March 16–20, 2025, Masdar City, Abu Dhabi, United Arab Emirates.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3745900.3746108

1 Introduction

User research in human-computer interaction (HCI) traditionally relies on direct engagement with human participants [16]. This paradigm is currently being challenged by *synthetic users* that are artificial entities generated by large language models (LLMs) that aim to emulate human behavior and interactions in a realistic manner (see Figure 1). Companies are increasingly adopting synthetic users to supplement or even replace human participants in their user or market research. A 2025 report found that 69% of market researchers have already used synthetic data in their research efforts[1], and 70% of market researchers believe that synthetic responses will

[1]https://backlinko.com/market-research-statistics

This work is licensed under a Creative Commons Attribution International 4.0 License.

AHs 2025, Masdar City, Abu Dhabi, United Arab Emirates
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1566-2/25/03
https://doi.org/10.1145/3745900.3746108

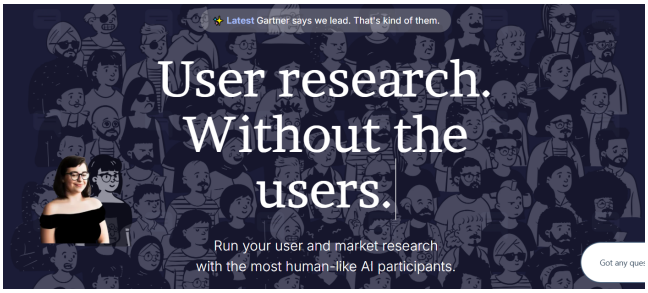

Figure 2: The homepage of Synthetic Users (https://www.syntheticusers.com/), a company offering synthetic user services.

constitute more than half of all data collection about consumers within the next three years[2].

Synthetic users are *marketed* as alternatives to traditional user research participants, promising reduced costs and increased efficiency [12]. Some synthetic user services promote their solutions even as "user research without the user" (see Figure 2), which seems a paradoxical claim. Alerted by this idea of removing the user from user research, many HCI researchers have expressed skepticism toward synthetic users, questioning their validity [2, 17, 25], as well as arguing that the idea of synthetic users goes against the principles of user research and user-centered design (UCD), which predominantly postulate that insights about users should be tractable to real user data, not simulation [9].

From an intellectual standpoint, the current discussion on synthetic users tends to gravitate toward two extremes: *outright dismissal* or *uncritical acceptance*. Neither of these polarized stances advances a balanced understanding of this emerging phenomenon. Thus, there is a gap in analyzing why synthetic users exist and what their emergence reveals about the current state of HCI research methods. Against this backdrop, we investigate three research questions (RQs):

- **RQ1:** *What are possible root causes driving the emergence of synthetic users in HCI?*
- **RQ2:** *What are possible use cases for synthetic users in HCI?*
- **RQ3:** *What are known limitations and key issues in synthetic users?*

The aim of our conceptual analysis is to understand the reasons driving synthetic user use and their implications for user research. This effort is particularly important given the rapid adoption of synthetic users in commercial user research, which needs to be informed about the limitations of this technology.

2 Conceptual Underpinnings

Market research and UX design have traditionally used "target groups", "personas", and "segments" as proxy users or "surrogates". These are fictional profiles that represent real user groups and help understand user needs and behavior [8, 10, 14]. In the context of HCI, continuing this trend, *synthetic users represent artificial entities designed to replicate human user behavior and interactions in research and testing contexts.*

Conceptually, synthetic users are similar to these segmented entities in the sense that they act as representatives of real users; for example, personas, especially data-driven personas [19], are typically based on a rigorous analysis of primary user data, in which said data is segmented and then written up as persona profiles [22].

Synthetic user services typically operate through (1) a proprietary synthetic segmentation engine that processes and synthesizes available data, (2) prompt design mechanisms that shape the synthetic user's responses, and (3) a user interface (UI) that provides human-like interactions to designers and other decision-makers using the synthetic users [28, 30]. When no real user data is provided, synthetic user services may require a user definition to get started. For example, when given a user description of "a low-income, immigrant parent in Oakland who does not speak English" [26], the system generates simulated responses and produces artificial interview transcripts and research notes. While this workflow *superficially* resembles traditional user research processes, it operates without involving any real participants. The LLM generates plausible-seeming but entirely synthetic responses that align, as per the LLM's pattern-seeking, with the specified user attributes. This *facsimile* of authentic user research fundamentally differs from traditional user research in which insights emerge from *actual* human interactions and experiences, often dealing with *targeted* rather than generic user groups [27].

3 Possible Root Causes Driving the Adoption of Synthetic Users in HCI

The adoption of synthetic users is primarily driven by pull and push factors. Pull factors relate to the convenience of using synthetic user services to get "realistic-sounding" user feedback, while the push factors relate to persistent practical challenges in the user research profession.

Concerning the *push factors*, according to an industry report [18], more than half (64%) of UX teams face time constraints, nearly half (46%) face budget constraints, and nearly a third (31%) find it difficult to recruit participants. The appeal of synthetic users is further augmented by growing privacy concerns and regulatory frameworks that make traditional data collection more cumbersome [23]. User researchers must invest significant time and resources to understand and comply with privacy regulations [11]. This includes creating compliant consent forms, establishing data governance policies, and collaborating with legal teams[3]. Additionally, the use of synthetic users does not require institutional review board (IRB) approval, which can be an onerous process. All of these efforts can be circumvented when using synthetic users based on generic patterns about people.

Concerning the *pull factors*, the tension between convenience and research validity emerges as a central theme in the application of synthetic users [3, 12]. While user research's purpose is to gain deep insights into users' circumstances, synthetic users are often promoted for their convenience in repetitive tasks and scenarios. This is compatible with the *principle of least effort* that postulates

[2]https://www.customerexperiencedive.com/news/synthetic-data-consumer-research-customer-journey-qualtrics/732408/

[3]https://www.rallyuxr.com/post/is-your-user-research-ethical-best-practices-for-privacy-and-consent

that people adopt easier-to-use tools over time, even when those tools would yield suboptimal outcomes [5, 17, 23]. Moreover, artificial narratives posed by synthetic users might produce outcomes comparable to a random real user's perspective [21], especially given the fluency, realism, and coherence of LLM-generated texts that aim to recreate perspectives and experiences of real people [7]. Harshly put, LLMs can provide "great quotes" for PowerPoint presentations that sound plausible. For some teams and individuals, especially under time and financial resource constraints, these synthetic user perspectives might appear "close enough" to "get the job done". Also, the general difficulty of assessing user research outcomes–particularly in determining "return on investment" (ROI) or establishing clear cause-effect relationships between research results and design choices–adds to the appeal of synthetic users because if the value of user research is "anyway unmeasurable", what harm can using synthetic users do?

4 Potential Use Cases for Synthetic Users in HCI

We consider here four use cases (UCs) that give the reader an *exploratory idea* of synthetic user applications, while illustrating that applying synthetic users is not a "black and white" matter.

UC01: Early ideation and hypothesis generation. Synthetic users can generate diverse perspectives for initial exploration and refinement through a more rigorous collection of user data. Creativity has been noted as a key application area for human-AI collaboration [20], and because creativity deals, in major part, with imagining "what could be", the imaginary (hallucinative) properties of LLMs pose less of a problem. In other words, the argument is: users *could* think about this matter in these ways.

UC02: Competitive analysis and market research. For example, when LLMs would carry out "deep research" [33] on customers' sentiments, e.g., based on publicly available product reviews, they could take the role of different customer types to represent their perspectives. In other words, using publicly available information about people, synthetic users extrapolate what people would generally think about a topic.

UC03: Usability testing. If designers have a well-defined understanding of a user group (e.g., general accessibility requirements among visually impaired users), these requirements could be included to guide the LLM to take the role of the "real" user. Another idea is to use synthetic users to pilot test user research plans and instruments *in silico* to identify potential risks prior to launching the study. In the process, synthetic users could help practitioners learn research techniques or practice interview skills.

UC04: Replacing traditional user research with synthetic alternatives. This use case is the least recommended one, but we mention it because several synthetic user services actually promote it as an option. It appears to fundamentally misunderstand the purpose of user research, namely to investigate real human experiences that cannot be synthesized from existing patterns and that often yield *unexpected* findings.

5 Known Limitations and Key Issues

Next, we will briefly investigate the challenges of synthetic users from a conceptual perspective, discussing four key issues (KIs).

KI01: Validation. Validation is a key concern. The validation of synthetic user systems remains largely proprietary, with companies claiming periodic comparison with live user testing sessions to ensure reliability. However, the specific methodologies are rarely disclosed. Current synthetic user systems operate as proprietary black boxes, making independent verification impossible. As Chapman postulates [6], it is likely that LLMs do not do a decent job of simulating specific groups of people in specific contexts—contrary to the purpose of user research in the first place.

KI02: Circularity. Because LLMs are pattern synthesis engines [4], they fundamentally cannot produce insights beyond the patterns present in their training data. In user research, this means that the outputs often reflect *stereotypical and generic* user characterizations [3] rather than specific, real-world experiences. In other words, the extant internet knowledge that the LLMs are based on eliminates the chance of serendipity, which is especially problematic when testing novel products for which there is no available "internet knowledge" (though the LLM might have content on the product category). Essentially, synthetic users reflect known patterns, making them unsuitable for discovering *new* usability issues or unexpected user needs.

KI03: Convincing Mimicry. The textual and narrative quality of synthetic users tends to be outstanding, such that their outputs *appear* convincing [24]. This increases the temptation to use them in real scenarios because they *seem* to make sense; they *seem* realistic, complete, and logical. Yet all this can be attributed to the LLM's general tendency of being "good with words". The "convincing mimicry" problem implies that designers may be more likely to trust synthetic user output precisely because it appears fluent and well-articulated, even when the content lacks validity. This creates a dangerous disparity between perceived and actual reliability. From brainstorming and the "hypothetical", there is a dangerously close leap to holding fiction as truth, as it becomes ingrained into decision makers' cognitive processes—what started as a "harmless exercise" could then transform into a truism that misleads the design process.

KI04: Normative Bias. The normative bias presents itself in synthetic user responses, with generated personas often defaulting to progressive, normative values stemming from the imposed guardrails, and describing users in an overwhelmingly positive light [28]. This reflects broader biases in the LLM training data rather than representative user perspectives. LLMs' tendency toward normative responses and positivity bias [28] risks reinforcing existing assumptions and dominant belief and value systems rather than challenging them, even when users would disagree with their content. Designers can easily confuse synthetic user responses with validated user needs.

6 Discussion

Addressing RQ1 , we postulate that synthetic users emerge from a convergence of practical and systemic pressures. Privacy regulations, while designed to protect real users, inadvertently create incentives that support artificial instead of real interactions. Also, the challenge of demonstrating clear value from user research can make synthetic alternatives more appealing because the fluency of LLMs can create compelling and persuasive narratives that appear

Table 1: Summary of synthetic users for HCI research: use cases and key issues.

Use Cases		Key Issues	
UC01	*Early ideation and hypothesis generation:* Synthetic users can potentially generate diverse perspectives for initial exploration in design projects.	KI01	*Validation:* Validation methodologies remain proprietary black boxes, making independent verification impossible for synthetic user systems.
UC02	*Competitive analysis and market research:* Synthetic users can access public information and "repackage" it from first-person perspective.	KI02	*Circularity:* LLMs as pattern synthesis engines cannot produce insights beyond their training data, resulting in stereotypical rather than specific user characterizations.
UC03	*Usability testing:* Can represent well-known patterns of user behaviors and pain points, e.g., those pertaining to accessibility.	KI03	*Convincing Mimicry:* The high textual quality of synthetic user outputs creates a dangerous disparity between perceived and actual reliability of the generated insights.
UC04	*Replacing traditional user research:* Misunderstands the fundamental purpose of user research to investigate real human experiences that contain unexpected elements (serendipity).	KI04	*Normative Bias:* Generated personas often default to progressive, normative values and overwhelmingly positive descriptions, reflecting biases in LLM training data and/or guardrails imposed by the AI companies.

more logical than "messy" and complex real user data. The difficulty in measuring user research outcomes contributes to this trend. If an organization struggles to establish clear links between user research findings and made decisions, the perceived cost-benefit ratio of traditional research methods becomes questionable. In this context, any user narrative can serve as design inspiration, and an artificial narrative might appear as valid as a randomly selected real user perspective. Finally, the principle of least effort [34] suggests organizations and people will gravitate toward easier solutions that produce reasonable outcomes at face value, even when these solutions are not optimal.

Addressing RQs 2 and 3, the identified use cases and risks rely on how the synthetic users are implemented ("devil is in the details"). Technical implementations of synthetic user services vary. On one end of the spectrum, the synthetic users are based on publicly available "internet knowledge", are used without primary user data, and are subject to any number of biases and stereotypes. On the other end of the spectrum, synthetic user services can provide a functionality to upload primary user data and then leverage retrieval augmented generation (RAG) or other techniques to ensure that the synthetic users correspond with the uploaded user data [15].

Proponents often suggest using synthetic users "in combination with" real user research [1]. Even though this appears to offer a reasonable compromise, it introduces specific risks. The *weakest link principle* [29] suggests that flawed information from synthetic users could compromise overall research quality, as designers and researchers cannot fully control the insights that influence their decisions. This is because of the *intractability problem* in UX design—we cannot control which information designers precisely remember and use [32]. Another risk is that the people involved in the use of synthetic users "forget" over time that these users are fictitious; this is akin to the mere exposure or anchoring biases, in that people may be consciously aware of the fact that the synthetic user is presenting "made-up information" but this made-up information nonetheless affects their thinking process, tacitly nudging some decisions to directions that are aligned with said information. There are also risks of practitioners developing unrealistic expectations about the capabilities of synthetic user services, which would be consistent with "magical properties" associated with algorithms and artificial intelligence [13, 31]. So, it is not at all "risk free" to deploy synthetic users, even with informative "warning labels".

We offer the following, tentative guidelines for the practical implementation of synthetic users in user research:

Guidelines about synthetic users

DO (POSSIBLY) USE:

- For *existing, well-known products* with
- *existing, well-known requirements*
- to pilot test user research plans and possible responses coming from
- a relatively *homogeneous user population.*

DO NOT USE:

- For *fringe or marginalized populations* with
- *novel products or services* that have
- *unpredictable requirements* or use cases to
- replace authentic user research with actual human subjects.

7 Conclusion

Generative AI can produce realistic user descriptions without using primary user data quickly and cost-effectively, but when these are not based on real user data, they can be misleading. The adoption of synthetic user services is driven by both *pull* and *push* factors: real user research is hard, and using synthetic users is easy. Synthetic users are gaining traction, particularly when research timelines are compressed and access to representative users is limited. Despite many HCI researchers dismissing synthetic users, their growing adoption suggests that synthetic users tackle perceived issues in current user research practices. So, rather than dismissing synthetic users entirely, the HCI community should seek to understand what their emergence reveals about shortcomings in user research practices. The HCI community may also need to pay more attention to demonstrating the value of conventional user research.

References

[1] Neeraj Arora, Ishita Chakraborty, and Yohei Nishimura. 2025. AI–Human Hybrids for Marketing Research: Leveraging Large Language Models (LLMs) as Collaborators. *Journal of Marketing* 89, 2 (2025), 43–70. doi:10.1177/00222429241276529 arXiv:https://doi.org/10.1177/00222429241276529

[2] A. Baki Kocaballi. 2023. Conversational AI-Powered Design: ChatGPT as Designer, User, and Product. *arXiv e-prints* (2023), arXiv–2302.

[3] James Bisbee, Joshua D Clinton, Cassy Dorff, Brenton Kenkel, and Jennifer M Larson. 2024. Synthetic replacements for human survey data? the perils of large language models. *Political Analysis* 32, 4 (2024), 401–416.

[4] Danilo Campos. 2023. What if Bill Gates is right about AI? — redeem-tomorrow.com. https://redeem-tomorrow.com/what-if-bill-gates-is-right-about-ai. [Accessed 11-02-2025].

[5] Senthil Chandrasegaran, Peter Lloyd, et al. 2025. Synthetic users: insights from designers' interactions with persona-based chatbots. *AI EDAM* 39 (2025), e2.

[6] Chris Chapman. 2023. Debunking LLM "IQ Test" Results — quantuxblog.com. https://quantuxblog.com/debunking-llm-iq-test-results. [Accessed 11-02-2025].

[7] Yoonseo Choi, Eun Jeong Kang, Seulgi Choi, Min Kyung Lee, and Juho Kim. 2024. Proxona: Leveraging LLM-Driven Personas to Enhance Creators' Understanding of Their Audience. doi:10.48550/arXiv.2408.10937 arXiv:2408.10937.

[8] Henry J. Claycamp and William F. Massy. 1968. A Theory of Market Segmentation. *Journal of Marketing Research* 5, 4 (1968), 388–394. doi:10.2307/3150263

[9] Alan Cooper. 1999. *The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity* (1 edition ed.). Sams - Pearson Education, Indianapolis, IN.

[10] Sabrina Duda. 2018. Personas—Who Owns Them. In *Omnichannel Branding: Digitalisierung als Basis erlebnis- und beziehungsorientierter Markenführung*, Vittoria von Gizycki and Carola Anna Elias (Eds.). Springer Fachmedien Wiesbaden, Wiesbaden, 173–191. doi:10.1007/978-3-658-21450-0_8

[11] Lucas Franke, Huayu Liang, Sahar Farzanehpour, Aaron Brantly, James C. Davis, and Chris Brown. 2024. An Exploratory Mixed-Methods Study on General Data Protection Regulation (GDPR) Compliance in Open-Source Software. doi:10.48550/arXiv.2406.14724 arXiv:2406.14724 [cs].

[12] Perttu Hämäläinen, Mikke Tavast, and Anton Kunnari. 2023. Evaluating large language models in generating synthetic hci research data: a case study. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, 1–19.

[13] Bernard J. Jansen, Joni Salminen, Soon-gyo Jung, and Hind Almerekhi. 2022. The illusion of data validity: Why numbers about people are likely wrong. *Data and Information Management* (Sept. 2022), 100020. doi:10.1016/j.dim.2022.100020

[14] Angus Jenkinson. 1994. Beyond segmentation. *Journal of targeting, measurement and analysis for marketing* 3, 1 (1994), 60–72.

[15] Soon-Gyo Jung, Joni Salminen, Kholoud Khalil Aldous, and Bernard J. Jansen. 2025. PersonaCraft: Leveraging language models for data-driven persona development. *International Journal of Human-Computer Studies* 197 (March 2025), 103445. doi:10.1016/j.ijhcs.2025.103445

[16] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research methods in human-computer interaction.* Morgan Kaufmann, Boston.

[17] Jie Li. 2024. How Far Can We Go with Synthetic User Experience Research? *Interactions* 31, 3 (May 2024), 26–29. doi:10.1145/3653682

[18] maze.co. 2023. Continuous Research Trends Report 2023 | Maze — maze.co. https://maze.co/resources/continuous-research-report/. [Accessed 11-02-2025].

[19] Jennifer (Jen) McGinn and Nalini Kotamraju. 2008. Data-driven persona development. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press, Florence, Italy, 1521. doi:10.1145/1357054.1357292

[20] Lucas Memmert, Julian Mies, and Eva Bittner. 2024. Brainstorming with a Generative Language Model: The Role of Creative Ability and Tool-Support for Brainstorming Performance. (2024). https://aisel.aisnet.org/icis2024/aiinbus/aiinbus/7/

[21] Ha Nguyen, Victoria Nguyen, Saríah López-Fierro, Sara Ludovise, and Rossella Santagata. 2024. Simulating climate change discussion with large language models: considerations for science communication at scale. In *Proceedings of the Eleventh ACM Conference on Learning@ Scale*. ACM, 28–38.

[22] Lene Nielsen. 2019. *Personas - User Focused Design* (2nd ed. 2019 edition ed.). Springer, New York, NY.

[23] Isaac Noble, Ivan Vendrov, Xin Xu, and Deepak Ramachandran. 2023. Realistic but Non-Identifiable Synthetic User Data Generation.. In *EvalRS@ KDD*.

[24] Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*. 1–22.

[25] Mirjana Prpa, Giovanni Maria Troiano, Matthew Wood, and Yvonne Coady. 2024. Challenges and Opportunities of LLM-Based Synthetic Personae and Data in HCI. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, 1–5. doi:10.1145/3613905.3636293

[26] Niloufar Salehi. 2024. I tried out SyntheticUsers, so you don't have to — niloufars.substack.com. https://niloufars.substack.com/p/i-tried-out-syntheticusers-so-you. [Accessed 11-02-2025].

[27] Joni Salminen, Soon-gyo Jung, Ahmed Kamel, Willemien Froneman, and Bernard J Jansen. 2022. Who is in the sample? An analysis of real and surrogate users as participants in user study research in the information technology fields. *PeerJ Computer Science* 8 (2022), e1136.

[28] Joni Salminen, Chang Liu, Wenjing Pian, Jianxing Chi, Essi Häyhänen, and Bernard J Jansen. 2024. Deus Ex Machina and Personas from Large Language Models: Investigating the Composition of AI-Generated Persona Descriptions. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.

[29] Martina Angela Sasse, Sacha Brostoff, and Dirk Weirich. 2001. Transforming the 'weakest link'—a human/computer interaction approach to usable and effective security. *BT technology journal* 19, 3 (2001), 122–131.

[30] Joongi Shin, Michael A. Hedderich, Bartłomiej Jakub Rey, Andrés Lucero, and Antti Oulasvirta. 2024. Understanding Human-AI Workflows for Generating Personas. In *Designing Interactive Systems Conference*. ACM, IT University of Copenhagen Denmark, 757–781. doi:10.1145/3643834.3660729

[31] David A. Siegel. 2010. The Mystique of Numbers: Belief in Quantitative Approaches to Segmentation and Persona Development. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. ACM, New York, NY, USA, 4721–4732. doi:10.1145/1753846.1754221 event-place: Atlanta, Georgia, USA.

[32] Laura van de Braak, Ronald de Haan, Mark Dingemanse, Ivan Toni, Iris van Rooij, and Mark Blokpoel. 2024. Intractability obstacles to explanations of communication.

[33] Junde Wu, Jiayuan Zhu, and Yuyuan Liu. 2025. Agentic Reasoning: Reasoning LLMs with Tools for the Deep Research. doi:10.48550/arXiv.2502.04644 arXiv:2502.04644 [cs].

[34] George K. Zipf. 1949. *Human Behaviour and the Principle of Least Effort.* Addison-Wesley.