Workshop on Developing HCI Study Designs Resistant to the Bias of User Expectations

STEEVEN VILLA, LMU Munich, Germany
THOMAS KOSCH, HU Berlin, Germany
ALENA DENISOVA, University of York, UK
ROBIN WELSCH, Aalto University, Finland

In the medical field, patients often experience tangible benefits from treatments they expect will improve their condition, even if the treatment has no mechanism of effect. This phenomenon often obscuring scientific evaluation of human treatment is termed the "placebo effect." Latest research in human-computer interaction has shown that using cutting-edge technologies similarly raises expectations of improvement, culminating in placebo effects that undermine evaluation efforts for user studies. This workshop delves into the role of placebo effects in human-computer interaction for cutting-edge technologies such as artificial intelligence, its influence as a confounding factor in user studies, and identifies methods that researchers can adopt to reduce its impact on study findings. By the end of this workshop, attendees will be equipped to incorporate placebo control measures in their experimental designs.

CCS Concepts: • **Human-centered computing** → **User studies**; *HCI theory, concepts and models*; *Empirical studies in HCI*.

Additional Key Words and Phrases: Placebo, Expectation, User Studies, Evaluation

ACM Reference Format:
Steeven Villa, Thomas Kosch, Alena Denisova, and Robin Welsch. 2023. Workshop on Developing HCI Study Designs Resistant to the Bias of User Expectations. In *CHI '24: CHI conference on Human Factors in Computing Systems, May 11–16, 2024, Honolulu, Hawaii, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 MOTIVATION

Robust evaluation is a foundational pillar in user-centered design in Human-Computer Interaction (HCI). For HCI as a methodologically strict discipline, it is imperative to ensure that the technologies and systems developed truly cater to the needs and preferences of end-users and bring benefits to them. A thorough and genuine understanding of the real benefits informs designers and developers, enabling them to refine and optimize their designs. Moreover, such evaluations promote transparency, integrity, and trust, ensuring that HCI solutions are effective in theory and provide meaningful value in real-world applications.

From medicine, we know that a placebo (e.g., a sugar pill) improves a patient's subjective benefits without administering any active substance or performing any specific procedure [9, 10]. A placebo reduces pain [14] or assists the treatment of illnesses [1, 11], consequently offering an efficient medical treatment without a disease-specific mechanism of action. The mechanism of effect relies on the patient's expectation of the placebo's effectiveness [2], which results in a positive evaluation after treatment [15, 19] but also at times in objective physiological changes [7, 8, 18]. Thus, the placebo effect obscures the evaluation of new medical treatments, no matter how helpful they are. Therefore, medical

Manuscript submitted to ACM

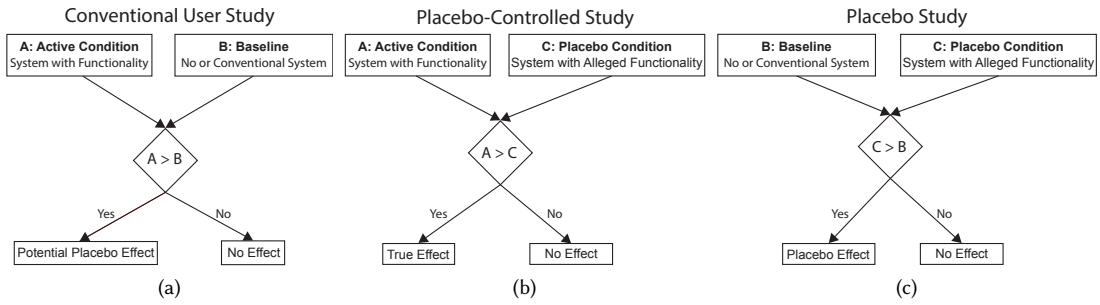

Fig. 1. Flow charts comparing conventional and placebo-controlled studies. **(a):** Conventional user studies comparing novel systems with a baseline do not consider the potential presence of a placebo, which may be responsible for the perceived superiority of the novel system. **(b):** Instead, comparing a novel system with a system that pretends an alleged novel functionality does reveal if the actual novel system has an effect [13].

.

studies use placebos as controls when testing a new treatment. Treatments are only considered effective if gains exceed the improvement of participants treated with a placebo control condition, i.e., an improvement caused by a treatment must exceed change caused by expectations alone. This placebo control is widely used in other scientific fields assessing human responses, including psychological treatment [2], sports science [16], and visualization research [4]. However, it is not the norm when evaluating the effectiveness of novel interfaces in Human-Computer Interaction (HCI).

In HCI studies, the expectations of improvements can be manipulated by the participant's awareness of interacting with a novel technology, functionality, or Artificial Intelligence (AI). Instructions might even explicitly describe the evaluated systems' enhanced usability or user experience and, hence, increasing participants' expectations [3]. Thus, a placebo effect in HCI would consist of a participant's favorable evaluation of effectiveness after the interaction (see Figure 1).

HCI studies have shown a placebo can improve usability and user experience without deploying a functional system. In games, fake power-up elements that make no difference to gameplay [6], and sham descriptions of AI adaptation increase the self-reported game immersion [5]. In social media, providing control settings for prioritizing items in one's news feed can result in higher subjective ratings of user satisfaction, even when these control settings do not influence anything [20]. Kosch et al. [13] could show that a non-functional supportive AI technology can create a placebo effect related to performance gains and relieved workload measures. Pataranutaporn et al. [17] showed that participant beliefs about an AI can be primed. In this context, participants who were about to interact with a mental health AI chatbot were divided into three groups. Each group was informed that the AI had different characteristics: (1) caring and empathetic, (2) only acting benevolently with manipulative motives, (3) using an algorithmic code with no inner motive. Afterward, all participants engaged with the same generative AI model. However, the authors found that participants who were told the AI was benevolent reported the AI to be significantly more trustworthy, empathetic, and more effective in providing mental health advice compared to participants primed to believe it was neutral or manipulative. Kloft et al. [12] the impact of user expectations on human-AI interactions, revealing that heightened expectations, regardless of AI presence, enhance performance due to placebo effects. Interestingly, negative AI descriptions do not alter these biased and robust performance expectations, highlighting the complex interplay between user beliefs and AI interactions. This can have a negative impact on human behavior when using allegedly enhancing technologies. Villa et al. [21]

explored the placebo effect within the context of human augmentation technologies, demonstrating a sustained belief of improvement after using a sham augmentation system and increased risk-taking linked to heightened expectancy.

Previous research presented strong results, considering that most of these studies were conducted in a highly controlled manner, often devoid of an experimenter's influence. Considering other context variables,e.g., the authority of the experimenter, the perceived value of the product, the setting of the room [22] which are present in typical user studies, can increase the placebo effect, the empirical literature and the knowledge base of HCI research may be biased.

To summarize, user expectations can elicit placebo effects that undermine HCI evaluation. Yet, HCI research lags behind psychology and medicine in examining the placebo effect. A placebo effect in HCI creates a self-fulfilling design process – An expectation of enhanced usability or user experience brings real benefits and favorable evaluation in the absence of better system functionality. We argue that protocols and practices in HCI must be carefully examined in light of the placebo effect. This workshop enables researchers and practitioners to control for the placebo effect and establishes new standards for evaluating usability and user experience in HCI. We present how current research methods can be improved and outline theories in usability and user experience that can fall victim to placebo effects. The workshop will pave the way for the adoption of placebo control in HCI and sketch robust evaluation techniques for researchers and practitioners in HCI with regard to placebo effects.

2 ORGANIZERS

Steeven Villa

Steeven Villa is a Ph.D. Student in human-computer interaction at LMU Munich, where he studies the limits of human augmentation. Previously, he researched how to create realistic haptic sensations in VR and AR using wearable devices at the Rainbow Team from INRIA Rennes. He has also researched mid-air ultrasound haptics, computer graphics, and the Universidade Federal Rio Grande do Sul. His research involves conceptualizing and developing novel technologies to enhance human natural cognitive, motor, and sensory skills and evaluating such augmentation technologies' societal and behavioral implications.

Thomas Kosch

Thomas Kosch is a professor of the Human-Computer Interaction Group at the Humboldt University of Berlin. His research focuses on implicit AI-driven physiological interfaces, user sensing, and user state predictions for human augmentation. Before, he was a professor at the Human-Centered Computing Group at Utrecht University and a postdoctoral researcher at TU Darmstadt, where he was conceptualizing physiological interaction and unobtrusive user sensing. He is experienced in designing user studies, quantitative and qualitative methods, machine learning, and prototyping.

Alena Denisova

Alena Denisova is a Lecturer (Assistant Professor) at the University of York, UK. She is actively involved in collaborative and interdisciplinary projects that involve conceptualizing and measuring the user experience of video games and designing and building educational and persuasive interactive media. Her research explores the role of the 'placebo effect' of technology in shaping player experiences, perceived challenge and uncertainty in video games, and, more recently, emotionally impactful player experiences - understanding how these experiences are shaped with the view to inform the design of games that promote these experiences.

Robin Welsch

Robin Welsch is an assistant professor at Aalto University, FI, researching human-computer interaction to improve theories and methods in engineering psychology. His current research interests include AR/VR, artificial intelligence, and human augmentation. Before joining Aalto University, he was an interim professor of General Psychology and Human Factors at TU Chemnitz, GER, worked as a postdoctoral researcher in the Human-Centered Ubiquitous Media group at LMU Munich, GER, and was a visiting scientist at VU Amsterdam, NL. He received his PhD in experimental psychology from the University of Mainz, GER, in 2020.

3 PRE-WORKSHOP PLANS

We will target members of the HCI community who are interested in enhancing the methodological and empirical foundations of HCI research. Due to the international and multidisciplinary nature of our organizing team, we can connect with local institutions of higher learning, special interest groups, and other related conferences. Due to the proximity of some of our organizing committee to the CHI23 venue, we will also seek to recruit first-year PhD researchers and local Master's students in an effort to foster more robust research practices from the early stage.

To help attendees understand the context of the topic at hand, our workshop will provide a collection of reference works from various fields (e.g., psychology and medicine) and examples of published HCI works that account for placebo conditions. We will disseminate information and materials via the website for the workshop. The information also includes the workshop's purpose, motivation, and potential outcomes. In addition, the website serves as a medium for advertising and recruiting potential workshop attendees. The participants will receive regular email updates regarding the workshop.

3.1 Website

The website for the workshop will include a description and objectives, a call for participation, and a list of proposed topics. It will also contain a link to the submission system, a detailed schedule, organizational information, and details about the organizers. To ensure that participants have sufficient time to read the accepted workshop papers and engage in constructive conversation at the workshop, they will be posted online before the conference.

4 WORKSHOP STRUCTURE & ACTIVITIES

Workshop Format: To make participation in our workshop possible for those who are still unable to physically attend CHI due to financial logistical or other constraints, we will conduct a hybrid workshop with a streaming option. The workshop will be a one-day event comprised of two blocks. The first block will establish the workshop's framework and foundation, including theoretical background and a keynote. The second block will consist of participant presentations, followed by a round of feedback for each presentation.

Workshop Timeline:

5 ASYNCHRONOUS ENGAGEMENT

Our website will provide the presenter's slides, papers, video recordings, and outcomes. Following the workshop, participants and those with an interest can view the materials. We will also offer a Discord server for people to discuss their findings, collaborate on projects, and provide feedback on previous work. Each presentation or group assignment

Table 1. Workshop Timeline

–	Block I: Foundations and Background.
10:00	Welcome and Introduction of the workshop.
10:15	Keynote: Placebo Effect on Medicine and Psychology, Why it is important?.
11:00	Keynote, Questions.
11:15	Break.
11:30	Keynote: Sham Systems and How Placebo Conditions would Look like in HCI.
12:15	Keynote, Questions.
12:30	Lunch Break.
–	Block II: Placebo Conditions in the Practice.
14:00	Introduction of the round of presentations.
14:15	First block of Presentations and Feedback.
15:00	Break.
15:30	Second block of Presentations and Feedback.
16:15	Wrap up and Conclusions.

on the Discord server will have its own channel, with all relevant resources (slides, papers, and results) linked there for easy access.

6 POST-WORKSHOP PLANS & WORKSHOP PROCEEDINGS

Following the workshop, we will post short videos recorded during the keynotes, presentations, and discussions along with the workshop slides, annotations, and other types of outcomes to our website and social media. We will encourage researchers to use the workshop's discussions and feedback to revise their research statements and position papers after the event. We will help researchers get their completed papers and reports posted on arXiv and our preprint server. Following approval from each speaker, we will upload video recordings of their pitches and the keynote address to a

repository or similar platform. The organizers intend to consolidate the most important points from the group work and moderated discussion into a manuscript crafted with the workshop attendees.

7 CALL FOR PARTICIPATION

A sugar pill is an example of a placebo. It makes a patient feel better without giving them any active drugs or doing any specific procedure. The effect mechanism relies on the patient's expectation of the placebo's effectiveness, which results in a positive evaluation after treatment and, at times, in objective physiological changes. In HCI, user study participants know the new system or method and can tell the difference between the experimental conditions. Instructions might even explicitly describe the evaluated systems' enhanced usability or user experience and, hence, set participants' expectations. Thus, a placebo effect in HCI would consist of a participant's favorable evaluation of effectiveness after the interaction. This workshop helps researchers and practitioners control for the placebo effect and sets new standards for evaluating usability and user experience in HCI.

Submissions should be in the two-column ACM format and no more than three pages long, not counting references. Information about submitting papers can be found on the workshop website[1]. The talks and presentations will be hybrid. We will record the presentations and publish them on the website. Participants will be selected based on the merit of their contribution to the workshop. We encourage authors to make their research available on arXiv[2] after the workshop. At least one author of each accepted submission must attend the workshop. All participants must register for the workshop and at least one conference day. We solicit the following types of submissions: position papers or research statements. Your position paper can address the topic of the placebo effect in general or demonstrate how a placebo control is used or planned to be used in your own experimental design.

REFERENCES

[1] Henry K. Beecher. 1955. The powerful placebo. *Journal of the American Medical Association* 159, 17 (12 1955), 1602–1606. https://doi.org/10.1001/jama.1955.02960340022006 arXiv:https://jamanetwork.com/journals/jama/articlepdf/303530/jama_159_17_006.pdf

[2] Walter Boot, Daniel Simons, Cary Stothart, and Cassie Stutts Berry. 2013. The Pervasive Problem With Placebos in Psychology Why Active Control Groups Are Not Sufficient to Rule Out Placebo Effects. *Perspectives on Psychological Science* 8 (07 2013), 445–454. https://doi.org/10.1177/1745691613491271

[3] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. 2019. 23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. 1–15. https://doi.org/10.1145/3290605.3300733

[4] Michael Correll. 2020. What Do We Actually Learn from Evaluations in the "Heroic Era" of Visualization? arXiv:2008.11250 [cs.HC]

[5] Alena Denisova and Paul Cairns. 2015. The placebo effect in digital games: Phantom perception of adaptive artificial intelligence. In *Proceedings of the 2015 annual symposium on computer-human interaction in play*. 23–33.

[6] Alena Denisova and Eliott Cook. 2019. Power-Ups in Digital Games: The Rewarding Effect of Phantom Game Elements on Player Experience. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 161–168. https://doi.org/10.1145/3311350.3347173

[7] Paul Enck, Ulrike Bingel, Manfred Schedlowski, and Winfried Rief. 2013. The placebo response in medicine: minimize, maximize or personalize? *Nature reviews Drug discovery* 12, 3 (2013), 191–204.

[8] Paul Enck and Sibylle Klosterhalfen. 2019. Does sex/gender play a role in placebo and nocebo effects? Conflicting evidence from clinical trials and experimental studies. *Frontiers in neuroscience* 13 (2019), 160.

[9] Damien G Finniss, Ted J Kaptchuk, Franklin Miller, and Fabrizio Benedetti. 2010. Biological, clinical, and ethical advances of placebo effects. *The Lancet* 375, 9715 (2010), 686–695.

[10] Asbjørn Hróbjartsson and Peter C. Gøtzsche. 2001. Is the Placebo Powerless? *New England Journal of Medicine* 344, 21 (2001), 1594–1602. https://doi.org/10.1056/NEJM200105243442106 arXiv:https://doi.org/10.1056/NEJM200105243442106 PMID: 11372012.

[11] Ted J Kaptchuk. 1998. Powerful placebo: the dark side of the randomised controlled trial. *The lancet* 351, 9117 (1998), 1722–1725. https://doi.org/10.1016/S0140-6736(97)10111-8

[1] https://posthci.com/placebows

[2] https://arxiv.org

[12] Agnes M. Kloft, Robin Welsch, Thomas Kosch, and Steeven Villa. 2023. "AI enhances our performance, I have no doubt this one will do the same": The Placebo Effect Is Robust to Negative Descriptions of AI. https://doi.org/10.48550/arXiv.2309.16606 arXiv:2309.16606 [cs.HC]

[13] Thomas Kosch, Robin Welsch, Lewis Chuang, and Albrecht Schmidt. 2023. The Placebo Effect of Artificial Intelligence in Human-Computer Interaction. *ACM Transaction Computer-Human Interaction* 29, 6, Article 56 (jan 2023), 32 pages. https://doi.org/10.1145/3529225

[14] Louis Lasagna, Frederick Mosteller, John M von Felsinger, and Henry K Beecher. 1954. A study of the placebo response. *The American journal of medicine* 16, 6 (1954), 770–779.

[15] Guy Montgomery and Irving Kirsch. 1996. Mechanisms of placebo pain reduction: an empirical investigation. *Psychological science* 7, 3 (1996), 174–176.

[16] M Ojanen et al. 1994. Can the true effects of exercise on psychological variables be separated from placebo effects? *International Journal of Sport Psychology* 25, 1 (1994), 63–80.

[17] Pat Pataranutaporn, Ruby Liu, Ed Finn, and Pattie Maes. 2023. Influencing human–AI interaction by priming beliefs about AI can increase perceived trustworthiness, empathy and effectiveness. *Nature Machine Intelligence* (2023), 1–11. https://doi.org/10.1038/s42256-023-00720-7

[18] Manfred Schedlowski, Paul Enck, Winfried Rief, and Ulrike Bingel. 2015. Neuro-bio-behavioral mechanisms of placebo and nocebo responses: implications for clinical trials and clinical practice. *Pharmacological reviews* 67, 3 (2015), 697–730.

[19] Steve Stewart-Williams and John Podd. 2004. The placebo effect: dissolving the expectancy versus conditioning debate. *Psychological bulletin* 130, 2 (2004), 324.

[20] Kristen Vaccaro, Dylan Huang, Motahhare Eslami, Christian Sandvig, Kevin Hamilton, and Karrie Karahalios. 2018. The illusion of control: Placebo effects of control settings. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[21] Steeven Villa, Thomas Kosch, Felix Grelka, Albrecht Schmidt, and Robin Welsch. 2023. The Placebo Effect of Human Augmentation: Anticipating Cognitive Augmentation Increases Risk-Taking Behavior. *Computers in Human Behavior* (2023), 107787. https://doi.org/10.1016/j.chb.2023.107787

[22] Tor D Wager and Lauren Y Atlas. 2015. The neuroscience of placebo effects: connecting context, learning and health. *Nature Reviews Neuroscience* 16, 7 (2015), 403–418.