QUEST: Questionnaire and Experiment Structuring Tool

Lukas Struck
University of Bonn

Florin Martius
University of Bonn

Kay Schröder
University of Applied Sciences Düsseldorf

Christian Tiefenau
University of Bonn

Matthew Smith
University of Bonn, Fraunhofer FKIE

Abstract

Questionnaire-based experiments require researchers to keep study designs, preregistrations, analysis plans, ethics applications, and other research artifacts consistent throughout the research process. Existing survey and experiment platforms mainly support study execution, while study designs are maintained separately, leading to duplicated effort, implementation drift, and increased supervision complexity. We propose a *"survey follows study design"* workflow in which researchers specify the study design once in a structured representation from which surveys and related artifacts can be derived. By explicitly linking study designs to research artifacts, this approach reduces duplication and implementation drift while enabling automation of tasks such as generating preregistrations and study documentation. We evaluated the workflow with our prototype QUEST in 13 sessions with researchers from Usable Security and Privacy (USP) and Human–Computer Interaction (HCI). Preliminary findings suggest that researchers value the design-first workflow and its support for maintaining consistency across research artifacts.

1 Introduction

Survey-based experiments are common in USP and HCI. Their quality depends heavily on explicit design decisions, such as hypotheses, variables, conditions, measures, assignment logic, and analysis plans, all of which must remain consistent throughout research lifecycle. Usually, researchers define the study design separately from the a survey or experiment execution tool, for example in text documents, spreadsheets, or project notes. To run the study, they must then translate this design into a survey or experiment execution tool. Existing tools split the same study into partial representations across execution, design rationale, and related research artifacts, such as preregistration. [1–9, 11–13]. This fragmentation creates friction for researchers, collaborators, and supervisors through duplicated effort, implementation drift and increased supervision complexity.

To prevent this friction, we propose a *"survey follows study design"* workflow for survey-based experiments. This workflow treats study design as the source of truth from which the survey flow, survey content, and other research artifacts are derived. It consists of three stages: First, researchers define a study design, from which survey flow derives, and researcher input survey content. This connectivity reduces redundant work and implementation drift for researchers, while also allowing related research artifacts that are considered "chores", like preregistrations, to be automated.

To support and evaluate this workflow in practice, we developed our prototype QUEST. To evaluate whether QUEST fits researchers' workflow, we conducted interviews (n=8) and think-aloud studies (n=5). While the analysis is ongoing, preliminary findings suggest that our prototype fits researchers' workflows and the linking of study design with experiment execution benefits the study process.

2 Workflow and Tool Proposal

The process begins with the definition of study design, where researchers specify hypotheses, variables, conditions, and the structure of condition assignment. Based on this design, they then construct the survey flow, while maintaining explicit links to the underlying design decisions. As the survey is developed, questions and informational content are added to the flow and associated with the relevant design elements. These connections between study design, survey structure, and survey content provide researchers with greater oversight, making it easier to identify errors and inconsistencies and to clearly trace which questions belong to which experimental

USENIX Symposium on Usable Privacy and Security (SOUPS) 2026.
August 23–26, 2026, Hannover, Germany.

conditions and hypotheses.

Our current prototype QUEST supports common survey-experiment structures and basic question and content types, such as factorial mixed methods designs and multiple-choice and open text questions. To facilitate collaborative study development, it provides features such as comments, proposals, and live synchronization. Since the study design is represented explicitly and remains linked to its execution throughout the survey creation process, downstream research activities can be supported more effectively. For example, hypotheses and design specifications can be used directly for power analyses, relevant information can be leveraged to derive a data management plan, and parts of a preregistration can be generated automatically. However, QUEST is limited in its functionality regarding study design, input methods, and automations and only intended as a proof-of-concept for the workflow.

3 User Study

Setup. We conducted a preliminary evaluation to understand whether the design-first workflow fits researchers' practices and where the limits of QUEST are. The evaluation included 13 sessions with researchers from USP and HCI. The sessions consisted of 8 interviews and 5 think-aloud studies; two participants took part in both formats. Two graduated researchers from our working group piloted the interviews. Participants were PhD students and postdocs from Germany, Switzerland, and the USA, consequently the sessions were held in English and German.

In the interviews, we presented QUEST using a fictional mixed-methods study based on Kariryaa et al. [10]. The mixed-methods demo was chosen because it allowed us to show most features without forcing artificial tangents. After the demo, we conducted semi-structured interviews about workflow fit, expected benefits, and limitations. Each session consisted of a core part that took 25 minutes, after which participants had the options to continue with a free discussion about their research workflows, study designs, and how to model them, with most sessions taking around 50 minutes in total.

To evaluate where QUESTs limits are, participants of the think-aloud studies used QUEST with their own study designs from a past study. These sessions focused on missing features, friction points, and issues rather than on validating the implementation as mature software. Each think-aloud session took around 50 minutes. In the following, we present preliminary results.

Findings. In the interviews, participants consistently reported that the workflow fit their research practices and that they would benefit from such a tool as QUEST, because the centralization of experiment definition reduces implementation drift and improves clarity by connecting survey execution elements to the survey design. Automations were well liked because they addressed annoying manual work and reduced the need to re-enter or reconstruct design information. Collaboration features, especially comments and proposals, were valued for supervision and collaborative study design. In the think-aloud studies, participants gave examples on how to further improve QUEST to fit their needs, including more hypothesis types and input methods.

4 Conclusion and Future Work

The evaluation suggests that researchers value a workflow that maintains explicit linkage between study design and survey implementation. Despite the limitations of the current QUEST, participants expressed interest in adopting a tool based on this approach, indicating that the underlying workflow is a substantial support in survey-based research.

Across sessions, researchers consistently reported that the design-first workflow aligned with their existing practices, and highlighted the benefits of maintaining continuity throughout the research lifecycle.

While QUEST currently supports only a subset of the envisioned functionality, the results demonstrate the feasibility and promise of the approach. Future work could expand support for more complex study designs, improve input methods, and evaluate the workflow with a larger and more diverse group of researchers.

5 Ethics and Transparency

We designed the study to minimize potential harms to participants and did not foresee any risks. Participants may have disclosed engagement in questionable or suboptimal research practices during the interviews. However, the purpose of this work was not to evaluate or assign blame to individual researchers, but to understand how research practices are currently implemented and managed in practice. All participants provided informed consent through an online consent form, and the study received approval from our institution's IRB. Participants did not receive compensation for their participation. To protect participant privacy, personally identifiable information is removed after transcript verification. Because the Usable Security and Privacy and HCI research communities are relatively small, researchers may be identifiable through their statements or experiences. We therefore take care to anonymize quotations and reported findings and, to the best of our ability, ensure that published data cannot be linked to individual participants. Large language models were used to assist with the phrasing of parts of this poster, and agentic coding tools were used during the development of the QUEST prototype.

References

[1] Home | AsPredicted — aspredicted.org. https://aspredicted.org/, 2026. [Accessed 24-05-2026].

[2] LimeSurvey — Free Online Survey Tool — limesurvey.org. https://www.limesurvey.org/, 2026. [Accessed 24-05-2026].

[3] OSF — osf.io. https://osf.io/, 2026. [Accessed 24-05-2026].

[4] Qualtrics XM: The Leading Experience Management Software - Qualtrics — qualtrics.com. https://www.qualtrics.com/, 2026. [Accessed 24-05-2026].

[5] SurveyMonkey: The World's Most Popular Survey Platform — surveymonkey.com. https://surveymonkey.com/, 2026. [Accessed 24-05-2026].

[6] Christopher Wickens Daniel Li Chen, Martin Walter Schonger. oTree — otree.org. http://otree.org/, 2014. [Accessed 24-05-2026].

[7] Joshua R de Leeuw, Rebecca A Gilbert, and Björn Luchterhandt. JsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *J. Open Source Softw.*, 8(85):5351, May 2023.

[8] Felix Henninger, Yury Shevchenko, Ulf Mertens, Pascal J Kieslich, and Benjamin E Hilbig. lab.js: A free, open, online experiment builder, 2024.

[9] Pingfan Hu, Bogdan Bunea, and John Paul Helveston. surveydown: An open-source, markdown-based platform for programmable and reproducible surveys. *PLoS One*, 20(8):e0331002, August 2025.

[10] Ankit Kariryaa, Gian-Luca Savino, Carolin Stellmacher, and Johannes Schöning. Understanding users' knowledge about the privacy and security of browser extensions. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, pages 99–118. USENIX Association, August 2021.

[11] D. J Leiner. SoSci Survey & Onlinebefragung, DSGVO-konform, deutsches Unternehmen — soscisurvey.de. https://www.soscisurvey.de/, 2024. [Accessed 24-05-2026].

[12] Sebastiaan Mathôt and Jennifer March. Conducting linguistic experiments online with OpenSesame and OSWeb. *Lang. Learn.*, 72(4):1017–1048, December 2022.

[13] Jonathan Peirce, Jeremy R Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. PsychoPy2: Experiments in behavior made easy. *Behav. Res. Methods*, 51(1):195–203, February 2019.