

Additional Key Words and Phrases: data exploration, computational notebooks, mixed-methods analysis

ACM Reference Format:

Jiawen Stefanie Zhu, Zibo Zhang, and Jian Zhao. 2024. Facilitating Mixed-Methods Analysis with Computational Notebooks. In *CHI '24: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/1122445.1122456>

1 INTRODUCTION AND BACKGROUND

In Human-Computer Interaction (HCI) research, researchers often collect and analyze a variety of data, including quantitative ones like Likert-scale items or task completion times, and qualitative ones like interview responses or diaries. This act of combining quantitative and qualitative analysis is called mixed-methods analysis, and is beneficial as it provides triangulation and allows research questions to be studied from different perspectives [4]. One crucial step in mixed-methods analysis is Data Exploration, or the act of trying to make sense of the data without a clear, predefined goal [7]. The data exploration process is inherently messy and challenging due to its vague and exploratory nature [3], and it requires researchers to go back-and-forth from different portions of the data, when trying to extract holistic patterns and connections.

Existing research has explored how to support the exploration of quantitative data. In particular, computational notebooks such as Jupyter Notebook and RStudio are deemed the most widely-used and useful tools due to their ability to integrate code, outputs, and documentation together [5, 7, 10]. However, computational notebooks are not optimized for mixed-methods analysis because they do not provide support for analyzing qualitative data. On the other hand, while commercial softwares such as NVivo and Google Docs [1, 2] do provide excellent support for qualitative analysis, they cannot be used for quantitative analysis and is not sufficient for mixed-methods research either. Thus, currently there exist few tools that adequately integrate both quantitative and qualitative analysis, and researchers are forced to switch between different tools during mixed-methods research, suffering from constant context-switching.

In this work, we aim to propose a tool that streamlines data exploration in mixed-methods research, by integrating both capabilities for quantitative and qualitative analysis in the same environment. We first propose a design concept that consists of *Data Extraction & Preparation* (Figure 1a), *Data Summary & Aggregation* (Figure 1b), *Data Integration & Interpretation* (Figure 1c) components, guided by Picolli and Card's sensemaking process [9]. Our design is also inspired by works in computational notebook, such as fluid move interactions to facilitate direct manipulation [6] and visual dashboards to enhance interactivity [11]. The proposed design allows users to examine and manipulate raw data (*Data Extraction & Preparation*), produce and study overviews such as visualizations (*Data Summary & Aggregation*) and investigate the high-level semantics of the data (*Data Integration & Interpretation*), in an augmented computational notebook environment. We then conducted a scenario study where we walk 3 HCI researchers through our interactive mock-up, to gain feedback on our prototype and to understand their experience with mixed-methods tools.

In summary, we will make two main contributions in this project: 1) a design for a **integrated notebook-based environment** that integrates quantitative and qualitative data exploration into the same environment, and 2) a **walkthrough** of our mock-up with 3 HCI researchers providing empirical understanding of the current workflows and pain points of mixed-methods researchers during data analysis, and the tools they may want in the future.

2 DESIGN CONCEPT

We design a design concept (Figure 1) that integrates support for both quantitative and qualitative analysis to streamline mixed-methods data exploration. The design concept consists of three widgets: *Data Extraction & Preparation* (Figure

1a), *Data Summary & Aggregation* (Figure 1b), and *Data Integration & Interpretation* (Figure 1c). Our design was guided by Pirolli and Card’s sensemaking process [9], since the main goal of data exploration is to understand, or *make sense* of the data. *Data Extraction & Preparation* is related to the foraging loop, where users search, filter, and extract the raw data to garner initial insights and evidence. In *Data Summary & Aggregation*, users examine overviews and aggregated results such as data visualizations, further working towards schematization. Finally, *Data Integration & Interpretation* roughly supports the sensemaking loop, where users evaluate the data and analysis findings to derive hypotheses and conclusions. Each component is implemented as a widget occupying one tab. The components appear side-by-side in the environment and share the same variable space.

2.1 Data Extraction & Preparation (Figure 1a)

The *Data Extraction & Preparation* component allows researchers to examine raw data. This widget can be seen as file viewer that supports two file types: text for qualitative and table for quantitative data. Users can freely switch between different files to examine different data. Textual data are displayed in a scrollable document reader as shown in Figure 1a right. Users can annotate by selecting the relevant text and adding notes. This is to enable coding of the data, which is crucial in qualitative analysis methods such as content analysis. When annotating, users can either quickly select from existing annotations (i.e. codes) or create new ones. Tabular data can also be viewed, as Figure 1a left shows. In the tabular view, users can filter or sort tables depending on their immediate goals, for example by choosing to only view the responses of one participant.

2.2 Data Summary & Aggregation (Figure 1b)

In addition to exploring raw data, users can also look at higher-level summaries and overviews through the *Data Summary & Aggregation* component. Overviews sit at a higher abstraction level than raw data, and helps users uncover deeper patterns and develop a more profound understanding of the data. This widget is implemented as a computational notebook, due to their versatility and integration of code, outputs and documentation [7]. Since both qualitative and quantitative data are stored in the same space in our environment, users can aggregate both types of data. For example, users can use this widget to generate visualizations such as word clouds or scatterplots, run statistical tests, or conduct frequency analysis, using code cells in the computational notebook.

2.3 Data Integration & Interpretation (Figure 1c)

The *Data Integration & Interpretation* component is central to the sensemaking loop [9]. In this component, information at all abstraction levels co-exist to assist users in piecing together a comprehensive picture of the data, and to synthesize holistic findings.

This widget provides a canvas-like area that helps users organize analysis results. Users can drag-and-drop content from *Data Extraction & Preparation* and *Data Summary & Aggregation* to become *blocks* in this widget (Figure 1.1). *Blocks* can be representative raw data, such as key quotes, or insightful aggregations, such as useful visualizations. Users can freely move and arrange these blocks on the canvas. For example, the user may choose to group the blocks based on the research question they are related, by putting related ones in proximity to each other.

In addition to organization, this widget also supports the linking of blocks to reveal the relationship between various parts of the data and the analysis. Users can link connected blocks (Figure 1.2) or chains of blocks (Figure 1.3) together. If applicable, the links can also act on a part of the block (e.g. portions of a quote, elements of a visualization), rather than the entire thing. For example, in Figure 1.2, the user links a relevant quote to the outlier on the scatter plot to

contextualize that data point. As another example, in Figure 1.3, the user links a chain of connected blocks. Each bar in the median bar plot represents the median participant response for one question in one condition. They first link a bar in the median bar plot to a histogram that displays the distribution of the data points contributing to that median. This allows the user to more accurately interpret the median by going one level down the abstraction ladder and providing details about how the data points led to the observed median. Then, the user further links a bar in the histogram to a quote. The quote goes down one more abstraction level and allows users to see the underlying reason why participants picked that particular response. In essence, this type of chaining process unwinds aggregations level-by-level, helping users to navigate between low-level raw data and high-level overviews.

3 SCENARIO STUDY

As a part of the iterative design process, we conducted walk-through evaluations and semi-structured interviews to collect feedback on our design concept and extract insights for possible future improvements.

3.1 Participants

We recruited 3 participants (2 women, 1 man; aged 29.3 ± 4.0) through convenience sampling. All participants are HCI researchers with experience in mixed-methods research, who have used computational notebooks for data analysis. They were able to participate the study in two ways: first as HCI researchers with expertise in designing user-friendly systems, they evaluated our design based on several heuristics criteria; and second as potential target users of this system, they provided feedback on the strengths and limitations of the system from a user's perspective. The study was approved by University of Waterloo's ethics review board.

3.2 Study Design and Procedure

The study was conducted in-person or online, depending on the participant's preference, with an approximate duration of 1 hour. The study starts with a demographic survey where participants report their previous experience and confidence level in quantitative and qualitative analysis, and proficiency with analysis tools including computational notebooks. Following the survey, participants were asked several warm-up interview questions regarding their experience with mixed-methods analysis, such as their current workflow, the tools they use, and challenges they faced. In particular, participants were asked to recall one research project in which they analyzed both qualitative and quantitative data to guide their responses. This puts them in the context of an authentic mixed-methods analysis project and helps elicit more concrete feedback.

Next, we presented the design concept to the participants as a Figma mock-up prototype. We introduced our design goals, the main widgets and features supported by each widget, and the connection between the widgets. Based on this prototype walk-through, the participants were asked about their feedback. We first asked them to evaluate the prototype's usability based on an adapted version of Nielsen's usability heuristics [8], where items on error prevention and recovery, and documentation are removed since they are hard to gauge with a non-working design concept prototype and less relevant to the goals of our current design phase. The participants were told to evaluate the design as HCI researchers and to leverage their expertise in design to provide feedback on the heuristics criteria. In addition to providing Likert-scale ratings, they had the option of either writing their feedback in the text box provided in an online form or verbally communicating their feedback with us. After the heuristics evaluation, participants were instructed to examine the prototype's utility as potential users. Participants evaluated whether the prototype sufficiently supports each of qualitative and quantitative analysis, whether it meaningfully integrates both, and whether it helps them make

sense of the data as a whole. Participants were also encouraged to imagine using this system in an authentic data analysis project that they may encounter in their research, and provide feedback on how they may use our prototype to address the challenges they encounter. They were asked to fill out a questionnaire with Likert-scale items to collect their ratings, which was followed by a semi-structured interview for more elaborate responses.

3.3 Results

From the walk-through, we were able to gather feedback from expert users on their perception of the current system design. In general, the participants believe the system can positively help the mixed-methods analysis process. Based on their experience as HCI researchers and as users of qualitative and quantitative analysis tools, they also provided feedback to support future iterations of our system design.

3.3.1 Support Various Granularities. Participants commented on the granularity of the data that could be segmented and presented in the *Data Integration & Interpretation* widget. To support flexibility in presenting qualitative data, the system should support different sizes of qualitative data to be selected from the dataset, from as small as one specific quotes from one participant, to general overviews like word clouds. This flexibility is also needed for quantitative data. Participants believed that the system should allow users to visualize and manipulate single data points, in addition to high-level visualizations.

3.3.2 Indicating Origin of Data. Participants also cared about the ability to see the source of each element presented in *Data Integration & Interpretation*. In mixed-methods research, data can be collected from a variety of different sources, such as interviews, focus groups, or controlled studies. Participants believed that awareness of the origin of the different data is helpful as it establishes context for properly interpreting the data. For example, participants thought that adding small textual descriptions or icons to indicate the source of the data can assist them in the sensemaking process.

3.3.3 Synergy across Components. Participants also wondered the extent to which the components are synchronized. In particular, they were curious about whether blocks in *Data Integration & Interpretation* changed as their original analysis in *Data Extraction & Preparation* or *Data Summary & Aggregation* updated. Real-time syncing would ensure consistency and that *Data Integration & Interpretation* would always be an up-to-date summary view of the current analysis. However, history would be lost in this case, which may be undesirable as users could want to refer back to previous analysis results. Whether and how to implement synchronization remains an important design choice to investigate for future iterations.

3.3.4 Scalability. Ability to adapt to large datasets and analyses was important to the participants. If datasets get large and complex, the canvas-like view of *Data Integration & Interpretation* could become cluttered and hard to navigate. One of the solutions could be to segment the canvas to create smaller chunks for smaller sub-problems. Another solution could be to use design to encourage users to only include contents necessary to their sensemaking process on the canvas, and to update, and clean the canvas frequently to ensure important information is always present.

3.3.5 Generalizability. Participants wondered about the flexibility and adaptability of the proposed tool. Qualitative and quantitative data might be very different across different scenarios and data analysis tasks, and the user needs for analyzing those data also vary as a result. If the system cannot exhaustively consider all potential use cases, it should allow users to customize the system based on their own preferences, and to expand the system features leveraging notebook's flexible environment.

4 CONCLUSION

Overall, we proposed a design concept that integrates quantitative and qualitative analysis in the same computational notebook-based environment. Based on a preliminary user walkthrough with 3 expert users, participants seemed to face the problem of context switching and liked the idea of an unified tool. However, they also saw and provided feedback for future design iterations.

Currently, the design concept we proposed is not yet implemented. For future work, we plan on implementing a working system, and conducting additional user studies to fully understand the implications and potential utility of our proposed design. It may also be interesting to investigate how we could make the proposed tool more intelligent, by automating the workflow and providing suggested actions where possible. For example, instead of forcing users to manually drag everything to the *Data Integration & Interpretation* widget, we could instead automatically add certain elements based on the context. Similarly, automatically suggesting the next relevant block when chaining can also be useful.

REFERENCES

- [1] [n. d.]. <https://help-nv.qsrinternational.com/20/win/Content/about-nvivo/about-nvivo.htm>
- [2] [n. d.]. <https://www.facebook.com/GoogleDocs/>
- [3] Leilani Battle and Jeffrey Heer. 2019. Characterizing Exploratory Visual Analysis: A Literature Review and Evaluation of Analytic Provenance in Tableau. *Computer Graphics Forum* 38, 3 (2019), 145–159. <https://doi.org/10.1111/cgf.13678>
- [4] R. Burke Johnson, Anthony J. Onwuegbuzie, and Lisa A. Turner. 2007. Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research* 1, 2 (April 2007), 112–133. <https://doi.org/10.1177/1558689806298224>
- [5] Mary Beth Kery, Marissa Radensky, Mahima Arya, Bonnie E. John, and Brad A. Myers. 2018. The Story in the Notebook: Exploratory Data Science using a Literate Programming Tool. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–11. <https://doi.org/10.1145/3173574.3173748>
- [6] Mary Beth Kery, Donghao Ren, Fred Hohman, Dominik Moritz, Kanit Wongsuphasawat, and Kayur Patel. 2020. mage: Fluid Moves Between Code and Graphical Work in Computational Notebooks. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20)*. Association for Computing Machinery, New York, NY, USA, 140–151. <https://doi.org/10.1145/3379337.3415842>
- [7] Xingjun Li, Yizhi Zhang, Justin Leung, Chengnian Sun, and Jian Zhao. 2023. EDAssistant: Supporting Exploratory Data Analysis in Computational Notebooks with In Situ Code Search and Recommendation. *ACM Transactions on Interactive Intelligent Systems* 13, 1 (March 2023), 1:1–1:27. <https://doi.org/10.1145/3545995>
- [8] Jakob Nielsen. 1994. *Heuristic evaluation*. John Wiley & Sons, Inc., USA, 25–62.
- [9] Peter Pirolli and Stuart Card. 2005. *The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis*.
- [10] Adam Rule, Aurélien Tabard, and James D. Hollan. 2018. Exploration and Explanation in Computational Notebooks. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3173574.3173606>
- [11] Yifan Wu, Joseph M. Hellerstein, and Arvind Satyanarayan. 2020. B2: Bridging Code and Interactive Visualization in Computational Notebooks. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20)*. Association for Computing Machinery, New York, NY, USA, 152–165. <https://doi.org/10.1145/3379337.3415851>