





Merits and Limits of Preregistration for Visualization Research


Lonni Besançon 
Linköping University

Brian A. Nosek [†]
Center for Open Science, Inc. & University of Virginia

Tamarinde Haven [‡]
Tilburg University

Miriah Meyer [§]
Linköping University

Cody Dunne [¶]
Northeastern University

Mohammad Ghoniem ^{||}
Luxembourg Institute of Science and Technology

ABSTRACT

The replication crisis has spawned a revolution in scientific methods, aimed at increasing the transparency, robustness, and reliability of scientific outcomes. In particular, the practice of preregistering study designs has shown important advantages. Preregistration can help limit questionable research practices, as well as increase the success rate of study replications. Many fields have now adopted preregistration as a default expectation for published studies. In 2022, we set up a panel “Merits and Limits of User Study Preregistration” with the overall goal of explaining the concept of preregistration to a wide VIS audience and discussing its suitability for visualization research. We report on the arguments and discussion of this panel in the hope that it can benefit the visualization community at large. All materials and a copy of this paper are available on our OSF repository at <https://osf.io/wes57/>.

Index Terms: Preregistration, Open Science.

1 INTRODUCTION

The replication crisis is a methodological crisis that stemmed from the lack of reproducibility of many scientific studies [60]. While often associated with specific fields such as medicine or psychology, the replication crisis is undoubtedly affecting other areas of research as well [3]. Researchers in empirical computer science often think that a replication crisis in their field is unlikely. However, recent findings have eroded such beliefs [2, 7, 17], identifying several serious threats to the replication of empirical computer science research, namely, the predominance of dichotomous interpretations of statistical results [7]; the lack of adoption of Open Science principles [17]; and few replications being attempted or published [2].

The replication crisis has led to methodological reform in many disciplines. One of the proposed reforms is the adoption of preregistration. Preregistration consists of storing on a public repository an immutable time-stamped record of the research outcomes, hypotheses, methods, and/or the planned analyses—ideally, before a study is conducted [37]. In practice, preregistration helps alleviate issues around questionable research practices such as outcome switching or HARK-ing [17, 18], selective reporting [1, 16], or p-hacking [31], and is also seen as a way to reduce the amount of duplicated research efforts [8]. In general, however, principles to increase the transparency of research outputs are rarely considered by visualization and human-computer interaction researchers [56]

such that preregistration is still far from being the norm in visualization manuscripts [26], despite all its advantages.

While preregistration may play a role in increasing transparency, some researchers consider it redundant at best [52]. Preregistration alone does not replace, nor provide a clear path to the development of good theories [55]. As flawed or low-quality research could also be preregistered, considering preregistration as a silver bullet is potentially harmful. Understanding how to improve a theory depends on debate and thoughtful discussion beyond preregistration alone.

Recent research has highlighted the tremendous importance of wider adoption of transparency in research ventures [8]. Empirical computer science—and visualization research in particular—is not an exception. The adoption of preregistration would already constitute a tremendous step forward, increasing robustness and transparency in our field. When questioned about their possible adoption of registration in a recent survey [9], visualization researchers highlighted several concerns about how suitable preregistrations are for their research methods. This paper aims to highlight and answer these concerns and begin conversations on how and whether preregistrations could be adapted to a variety of visualization research methods—in particular in the case of qualitative research [33].

It was also in this mindset of discussing preregistration and its potential benefits for the visualization community that we organized a Panel at IEEE VIS 2022 entitled “Merits and Limits of User Study Preregistration.” We defined the following goals:

1. explain the concept of preregistration to a wide visualization audience,
2. refute common misconceptions about the preregistration process,
3. provide insights about the merits and limits of preregistration from various fields, and
4. address the suitability of preregistration for a variety of types of visualization research.

In this position paper, we summarize the 2022 panel’s discussions and arguments for the wider visualization and human-computer interaction community, point to useful resources, and discuss implications along with any needed community-driven efforts. We first demonstrate the potentially wide application of preregistration through a preregistered LLM-based analysis of our IEEE VIS 2022 panel. We then manually extracted the salient points of discussions and the views of the panelists, which we believe are important to highlight. Finally, we propose practical advice for preregistration and their review to the VIS community.

Our discussion is nuanced with many caveats, but based on it, we offer some overly-simplistic guidance for visualization researchers:

1. You can modify your study design after preregistration if you are transparent about and justify your changes.
2. Preregistration is likely valuable for controlled experiments, technique evaluations, system or algorithm performance benchmarks, and systematic surveys; its value is less clear for design studies, other design and applied visualization research, and theory papers.

*e-mail: lonni.besancon@gmail.com

[†]e-mail: nosek@virginia.edu

[‡]e-mail: t.l.haven@tilburguniversity.edu

[§]e-mail: miriah.meyer@liu.se

[¶]e-mail: c.dunne@northeastern.edu

^{||}e-mail: mohammad.ghoniem@list.lu

3. Preregistration is only one part of communicating the analytic provenance behind research—detail and explanation in the paper is also needed.
4. Use the Open Science Framework (OSF) and standard templates¹ for preregistration.
5. Reviewers should evaluate the correctness of a paper's claims in the context of its preregistered studies.
6. Journals and conferences should allow the submission of registered reports.
7. Critically reflecting on our practices can improve guidance for researchers and reviewers.

2 METHODOLOGY AND ORGANIZATION

The video of the panel is available online on YouTube². As a playful demonstration that preregistration can be adapted to multiple methodologies and goals, we provide a first summary of the panel using a large-language-model-driven analysis. The exact methodology for this part of our analysis is preregistered on OSF³ and available on our OSF repository.⁴ The preregistration states:

For this purpose, we extract the audio track of the YouTube video of our panel.

The audio track is then given to Whisper-diarization with the following parameters `language = "en", num_speaker = ""`, `prompt = Lonni Besançon, Brian Nosek, Tamarinde Haven, and Miriah Meyer talking about the value of preregistration for visualization research.` The software used [43] is available here: <https://replicate.com/thomasmol/whisper-diarization>.

The obtained JSON file is sent to ChatGPT-4o with the following prompt: "Can you clean this JSON file into a proper text file that has the continuous talk of everyone without all the unnecessary inferences/-data? Please separate the arguments from each individual speaker into different paragraphs for clarity." The result is then downloaded before it is sent again with the prompt: "Now, using only their own words, as given in the text file above, not adding any text to this at all, summarize the viewpoint of each speaker."

The results of this analysis will be reported in a section titled "Preregistered LLM Analysis" of the paper, and deviations, if any, will be clearly reported.

Input from the speaker Lonni Besançon that would be redundant with the introduction of this paper will be removed from the report.

Outside of this preregistered analysis and for the remainder of this manuscript, we summarize the panel based on the main messages from each panelist and the discussions that took place their presentations. We paid particular attention to try and reflect upon different methodologies and epistemologies and their meaning and need for preregistration. In doing so, we hope to make this paper about preregistration and the need for increased transparency encompassing of the VIS community at large.

3 PREREGISTERED LLM ANALYSIS

We now detail and discuss the deviations from our preregistration and our analysis results.

3.1 Deviations and results

The only deviation from our preregistered analysis is that the initial prompt "Now, using only their own words, as given in the text file above, not adding any text to this at all, summarize the viewpoint of each speaker." was appended with "Always write the Speaker's name at the beginning of each paragraph" when parsing the JSON file given by whisper-diarization. We then edited the produced text file, replacing generic speaker numbers with their real names.

¹<https://help.osf.io/article/229-select-a-registration-template>

²https://www.youtube.com/watch?v=77YeUe_Bkco

³<https://osf.io/j4hqh>

⁴<https://osf.io/wes57/>

The following subsections detail the output of the final command sent to ChatGPT4o, broken down by speaker and panel moments.

3.1.1 Brian Nosek

Brian provides context for preregistration, explaining its role in improving research credibility by addressing publication bias and distinguishing between planned and unplanned research. He highlights that preregistration can help correct publication bias by making all conducted studies discoverable, not just those that are published. Brian also emphasizes that preregistration helps clarify the distinction between confirmatory (planned) and exploratory (unplanned) research, allowing for more credible statistical inferences. He acknowledges the evolving exploration of preregistration's boundaries across different methodologies and stresses the importance of transparency in the research process.

3.1.2 Tamarinde Haven

Tamarinde focuses on preregistration in qualitative research, emphasizing its potential benefits for boosting credibility and transparency. They explain that qualitative research often involves an emergent design, making preregistration seem counter-intuitive. However, Tamarinde argues that preregistration can clarify initial beliefs and expectations, even in exploratory research. Along with collaborators, they conducted a Delphi study⁵ involving the qualitative research community to develop a pre-registration form tailored for qualitative studies. Tamarinde highlights that preregistration is a plan, not a prison, and should be used voluntarily to provide a systematic starting point for research while allowing for necessary updates as the study evolves.

3.1.3 Miriah Meyer

Miriah presents a dissenting view on preregistration, particularly for design-oriented visualization research. They argue that the VIS community is diverse and not just limited to quantitative and qualitative research. Miriah emphasizes that design-oriented work, characterized by its dynamic and iterative nature, may not align well with preregistration. They express concerns that preregistration could be harmful by imposing positivist values on other research approaches. Miriah advocates for diverse research values and guidelines that accommodate the community's methodological diversity and highlights ongoing work on improving the transparency of design processes through traceability and technology.

3.1.4 Discussion points

The panelists engage in a discussion on the applicability and limitations of preregistration across different research paradigms. They acknowledge the challenges of implementing preregistration in qualitative and design-oriented research, and stress the importance of transparency and flexibility. The panelists discuss alternative ways to achieve transparency, such as detailed reporting and using technology to capture the research process. They also consider the broader goal of improving research practices and the need for inclusive guidelines that respect different research methodologies.

3.2 Discussion

This analysis of the content of the panel was mostly made to highlight that preregistration is not solely relevant to quantitative research, but that it can be useful for a multitude of data collection and analysis types. In this case, the preregistration forced us to be transparent about what we tried and how we have deviated in the analysis, confirming what is reported by almost all panelists in the subsections above. However, this preregistered analysis is

⁵A Delphi study is a methodology involving a questionnaire and multiple iterations designed to develop a consensus of opinion concerning a specific with experts of the field [5].

clearly insufficient. Some of the speakers' arguments are not transcribed despite their importance. Thus, we wish to provide a more constructed, comprehensive, and substantiated argument about the value and limits of preregistration for visualization research.

4 THE VALUE, LIMITS, AND FUTURE OF PREREGISTRATION

The panel discussion helped clearly identify the benefits and limits of preregistration, as well as when it should be adopted or how it could be improved upon. We discuss these aspects below.

4.1 A plan, not a prison

Preregistration is often seen as a procedure that constrains research, both within [6] and outside [19, 36] the visualization community. However, this idea is based on a misconception of preregistration itself: its goal is not to forbid any deviation from the initial plan, especially if the deviation would help to contribute more robust research results, but rather to make sure that the deviations from the initial plan are transparently reported [36]. In fact, the panelists and authors of this paper have often deviated from the preregistrations that they have made. Consequently, they had to report these deviations in their papers, therefore increasing the transparency of their methodology and allowing readers to judge for themselves the appropriateness of the deviations. In essence, preregistration helps to provide a better and more systematic distinction between planned research (often confirmatory to validate hypotheses) and unplanned research (often exploratory to generate hypotheses) without valuing one over the other [35, 41]. Statisticians and methodologists [4, 41] have also argued that, in the case of inferential results, forcing researchers to report all of their statistical tests would enable more accurate and credible statistical inferences by, e.g., avoiding the inflation of false positives. In addition, preregistrations help combat the publication bias towards positive results [37, 54]. Although preregistered null results are not more likely to be published (unless they are part of registered reports), they are at least discoverable through their preregistration.

4.2 It is not just for quantitative analyses

In the previous section, we highlighted that preregistration is an essential tool to separate confirmatory work from exploratory work. Yet, there are still many misconceptions about the targeted research methodologies of preregistration. Many researchers still believe that it is only useful for quantitative analyses. This misconception likely stems from the fact that preregistration was initially designed for fields that rely strongly on frequentist statistical analyses [14].

In the case of qualitative work, preregistration inherently seems counter-intuitive since quantitative research commonly operates on exploratory logic [33]. First of all, since some qualitative research is not only exploratory but also has a clear confirmatory goal, preregistration is helpful to clarify the planned testing procedure and highlight deviations, if any. In the case of qualitative research that is exploratory, scholars may bring prior knowledge and preconceptions to a particular phenomenon they study, and the preregistration may help clarify which of the findings really drew on the initial beliefs and expectations and existing knowledge and which diverged from those. Not only is this going to increase the transparency of the manuscript with respect to the authors' previous knowledge and potential positionality, but it would also help better distinguish between expected and unexpected findings. Furthermore, preregistration of exploratory work can be considered as systematically logging design iterations [21]. Indeed, the log of differences between the initial preregistration and the final conducted investigation could serve as an "open laboratory notebook" [49]. This log can then be used by the authors and readers to expose all the design choices that were considered, as well as justify the final decisions [21]. It is, in this case, apparent that the ability given to

exploratory researchers to adapt and adjust as their exploration advances is not lost with the preregistration, but rather that the adjustments are all systematically highlighted. However, for qualitative and exploratory work, the preregistration templates are bound to be different and are, like templates for quantitative preregistration, still a topic of research [33]. The current templates have not been developed with visualization research in mind and, to the best of our knowledge, rarely used in visualization (examples are available, e.g., in [45, study 2], [59]). It may, therefore, be important for visualization researchers to study and use these templates and propose new ones that would better suit the needs of exploratory visualization research. In such cases, an aspect of preregistration that is not often considered outside of clinical trials is the possibility of modifying the plan and justifying any modifications.

4.3 It is, however, all about epistemology

The examples highlighted above are all within the philosophical school of *positivism*, that is, the set of methods to verify a priori hypotheses [40]. However, some contributions visualization researchers provide do not fit this framework and would be better characterized within other epistemologies such as *critical theory* or *constructivism* [25]. Particularly, the design work often conducted by visualization and human-computer interaction researchers does not fit the positivist lens, and tools like preregistration seem ill-fitted for such research [33]. Because different epistemologies emphasize different criteria for what eventually constitutes good research, the values that preregistration helps to maintain are unlikely to be useful for studies outside the positivist lens. In other words, as preregistration was initially designed to increase the reproducibility of research results, it is difficult to apply it to research that is bound to be non-reproducible. Design work, design studies, and applied visualization research—which are more likely to fall within the epistemology of *interpretivism* [34, 50]—are good examples of this. For those, transparency of the design process can be an important criterion for rigor [34], but the strong focus of preregistration on data analysis lessens its potential utility.

Other examples outside the positivist epistemology include system papers, some review papers, and theoretical papers. Preregistration would only be useful for system papers if the systems' performances are tested. For instance, one may argue that the performance benchmarks of D3 [13] and the example interactive visualizations of VegaLite [48] could have been preregistered. Similarly, when developing domain-specific visualization software, a performance evaluation could be preregistered. For instance, in the analysis of molecular Dynamics provided by VIAMD [51], the authors could have preregistered their performance and scalability tests. In the case of review papers, the usefulness of preregistration largely depends on the type of review conducted (comprehensive lists of review paper types have been identified in the literature, see, e.g., [24]). Review papers that aim to provide exhaustive and comprehensive searching (see [24, Table 1]) of the literature would directly benefit from preregistering their research protocols to ensure that deviations are clearly reported (e.g., meta-analyses, scoping reviews, state-of-the-art reviews). Review papers that aim to provide a critical appraisal of the reviewed literature, either within the text itself or for inclusion/exclusion criteria (see [24, Table 1]) would also benefit from stipulating their appraisal criteria. Conversely, review papers that aim to provide a narrative presentation of the work conducted in a field would likely not benefit from posting a preregistration. Similarly, it is quite unlikely that a theory paper would benefit from a preregistration.

4.4 The place for traceability and the role of HCI to shape better tools for it

Some of the above discussions have highlighted that the need for transparency remains apparent across the very diverse set of

methodologies and epistemologies used in visualization research, although it is clear that preregistration is likely to be the wrong tool to achieve the desired transparency. However, the idea behind preregistration, and in particular preregistration of qualitative and exploratory work—such as constructing a lab notebook to justify decisions and provide rationales for approaches taken—is an interesting topic that we believe deserves more attention. Increasing the traceability of design decisions back to the acquired observation, documenting the evolution of the investigation or development process, or providing dynamic links between published papers and supplemental materials are all interesting outcomes for works that would not otherwise benefit from preregistration.

If one considers that the ultimate goal of preregistration is to increase transparency and clarity of what was planned versus what was iteratively refined or exploratory, then preregistration alone is perhaps insufficient even for the use case of quantitative positivist research. Indeed, preregistering and sharing research materials without properly detailing or explaining how they correlate to findings in the research manuscript may only increase confusion, in particular for design studies where thousands of artifacts are created or collected. The idea behind OSF and the tools it offers is to allow researchers outside the original scientific team to “replay the movie” of the research conducted. While this seems easy, the implementation is complex. One has to carefully consider what, when, and how to record the research “movie” so that it makes sense, a topic that our community has labeled *analytic provenance* [39] and for which research results could help foster more clarity and transparency of research manuscripts. This is where we, as visualization and human-computer interaction researchers, can make an impactful contribution to our own community and to science as a whole.

We can leverage the diversity of our methods and epistemologies to add to reflections around the diversity of templates needed or propose new interactive systems in the vein of *trrace* [44], an auditable website collection of research artifacts gathered by the authors of a design study paper to increase the transparency of their methodology. Thus, we argue that our community should consider research transparency as a research topic to which we can contribute using our knowledge and experience, particularly our expertise in collaborating with researchers, actors, and stakeholders from different domains. One such group of actors and collaborators could be the researchers and employees from OSF, whose goals seem to align with ours if we are to judge from this panel alone. We, therefore, not only echo but extend past calls on the subject (e.g., “Transparent Quantitative Research as a User Interface Problem (Dagstuhl Seminar 22392)” [57]).

5 PRACTICAL ADVICE

We set out to summarize for the visualization community our panel’s discussions, explore implications for the visualization and HCI communities, and provide useful resources for preregistrations. We have, hopefully, managed the first two goals throughout the previous section. We now hope that readers will arrive at this section with a clear understanding of whether they should preregister. If so, there are several points we want them to consider.

5.1 Hosting platforms and templates

One of the first choices that researchers need to make is the platform to use for their preregistration. A good overview of these platforms and their benefits and limits is available in Haroz’s comparison of preregistration platforms [27]. The comparison includes various factors, including formatting options, the presence of timestamps, persistency, the possibility of tracked changes, and more.

The choice that comes right after the platform is one of the templates to use, if any. A close look at some templates and their degrees of producibility is available in van den Aker et al.’s study [53].

As readers may have devised, the present paper relied on an “Open-Ended Registration,” but many other templates exist for qualitative research, for replications, for registered reports (see e.g., [9]), and for many others. The specific choice of a template will depend on the research goal, methodology, and even potentially on the institution of the researchers. OSF, for instance, offers detailed documentation about each type of preregistration template that they host.⁶

5.2 Preregistration in practice

As a preregistration cannot be deleted after its submission, it is often recommended to iterate over the study design and analysis plan before submitting the preregistration. Visualization researchers may find it helpful to test their protocols in a pilot study or on artificial data, for instance, to ensure the quality and appropriateness of the method. The data gathered in a pilot study could be used to refine the analysis plan before preregistering. It is also interesting to note that OSF allows researchers to submit a copy of their codebooks or code for their quantitative data analysis, which would be subsequently frozen along with the preregistration when it is submitted. The VIS and HCI communities have already used such features in the past (see, e.g., [11, 22]).

This is a default feature of the OSF repository: files in the repository when the registration is submitted are included within the preregistration. Should such files evolve later, researchers could use visual diff tools to highlight any changes by comparing the new version of the files with the frozen preregistered ones. Note that authors should be careful not to include private or deanonymizing information in the repository when the preregistration is submitted.

Registrations also offer the possibility to specify how results will be reported in the paper such as where they would be described and what the visualizations will be. Registering the analysis and visualization code makes the latter explicit. It is generally recommended that the results of preregistered analysis be reported in a specific section and the results of non-registered analysis, along with any deviations, in a separate section. While one may, of course, deviate from this reporting plan, it is recommended to follow such a structure and indicate it clearly in the preregistration, as well as provide the name of the section that will contain the non-registered analysis.

Finally, while preregistrations are frozen and cannot be modified directly, it is possible to transparently update them such that the changes made are highlighted to readers and reviewers alike. This is a basic feature in clinical research and a developing feature in OSF.io⁷ that visualization researchers may benefit from and contribute to with our HCI expertise.

5.3 Reviewing preregistrations

An important part that is rarely discussed in our community is for reviewers to value the presence of preregistration, even if deviations from the preregistered plan happened (see subsection 4.1). Guidelines on the reviewing of preregistrations are scarce, and the lack of a standard way to report deviations can be problematic [61]. We can only recommend that reviewers refrain from penalizing reasonable deviations from a preregistered study design and value the authors’ increased transparency. Reviewers should, however, ensure that the authors do not present exploratory results as confirmatory ones and make use of language that correctly matches the strength of evidence that they obtained.

5.4 Going further: Registered Reports

Although, as we have highlighted above, preregistrations can increase the transparency of research processes, they are unlikely to

⁶<https://help.osf.io/article/229-select-a-registration-template>

⁷<https://help.osf.io/article/382-updating-registrations>

help make them more robust or valid [8]. This is where *Registered Reports* can benefit us [15, 23, 38, 58]. With a *Registered Report* researchers must also commit to a complete data collection and analysis plan before they run their study. However, a *Registered Report* is not just an analysis plan that reviewers can later check for potential deviations. Instead, it takes the form of a paper with partial content that is submitted for peer review instead of only being uploaded to an online registry. This is called a *Stage 1 Registered Report*, and it usually contains an introduction, related work as well as the methodology, including analysis plan. Reviewers can make comments about the study plan and ask for changes before the study is conducted and the participants are recruited. As such, potential omissions or mistakes in the proposed methodology can be corrected before any substantial resources are used, thereby limiting scientific waste [8, 15]. Although their potential for visualization research has already been discussed by Besançon et al [6] at alt.VIS 2021, it remains that their implementation with our conference cycle would be difficult [6], but other visualization venues such as *The Journal of Visualization and Interaction (JoVI)* [10] currently accept registered reports.

5.5 A community call to action

Preregistration is, as we discussed, very often associated with the notion of confirmatory or positivist research contributions and the notion of replication and robustness. While there are many guidelines and inspiring examples for the former, we believe the latter still lacks useful templates or guidelines. For contributions that focus on confirmatory quantitative experiments, several papers within the visualization (e.g., [12, 30, 47]) and HCI (e.g., [32, 46]) communities can serve as a good starting point for researchers wanting to be preregistering. For qualitative research contributions, the groundwork laid by Haven et al. [33] and an early example within visualization for expert feedback collection [59] may be inspiring to scholars of our community. However, we argued that preregistration may have value outside of these methodologies and epistemologies although guidelines and examples are scarce.

As we have discussed, the variety of methodologies and epistemologies makes it difficult to give concrete recommendations that would apply to all the varied research contributions and members of our community. However, we have argued here and in the past [6] that this variety should not bar us from trying to define community guidelines. Looking back at our (the authors) own research, we speculate there were projects that could have benefited from being preregistered. We reflect here on two of those projects while also pointing at what we believe to be good examples of preregistrations. We put forward these examples to offer initial guidance of why, when, and how to use preregistration in visualization research, and call on the community to engage in a discussion how to develop more comprehensive guidelines.

A first example we want to consider is our systematic review methodology presented by Di Bartolomeo et al. [20, esp. Section 3 & Figure 3]. This survey methodology was not preregistered, so it is unclear to the reader whether any modifications to the methodology were performed during collection that could have affected results. Although these modifications did not occur, it is not apparent from the existing paper and materials. Survey papers have different aims, and their findings directly result from the surveying methodology and the different criteria for including and excluding papers. One methodology for systematic reviews is to use the PRISMA guidelines.⁸ Preregistering the survey methodology would make the process more rigorous and transparent, including inclusion and exclusion criteria, as well as any critical assessment criteria. This specific kind of contribution may, however, call for its own tailored template that would help preregistering state-of-the-art papers.

⁸PRISMA: Transparent Reporting of Systematic Reviews and Meta-Analyses <http://www.prisma-statement.org/>

Another example from our own research is an application-oriented contribution that iteratively involved domain experts to improve the design of the application. In our autopsy-focused application paper, Pooryousef et al. [42] conducted multiple iterations of prototype development and testing with pathologist experts. Reflecting on the process, we speculate that preregistering each step of the process before they unfolded could help make them clearer. The results are unlikely to have been different, but the reporting process would have been simplified by being able to go back to each preregistration with our assumptions and goals and compare them to the lessons learned after each round of expert feedback. Ultimately, we believe that the set of methods, insights, and challenges in this situation all offer rich learning points for the community that preregistration may help highlight, although it remains unclear how all of it could be reported within a single paper. As shown by Yang et al.'s recent paper [62] exposing the prequel of their work published the previous year on election forecast [63], a single paper presenting final results to a research question may not offer the whole community the necessary insights the authors gained throughout their investigation. This is, perhaps, an opportunity for us to also reflect on our page-limited contributions and their limitations.

Beyond hypothesis-driven research, we speculate that the processes and tools for preregistration could inspire new approaches to recording and reporting on design-oriented research. In particular, design-oriented research relies on reflection throughout the research process in order to make decisions on next steps, to refine goals and outcomes, and ultimately to distill learning into insightful and new knowledge contributions [34]. Preregistration process and tools, perhaps recast as *prereflection*, could offer ideas for how to more thoroughly and transparently record and report on reflective research practices.

We do not claim that these examples represent a complete or exhaustive list of contribution types in visualization research. Rather, we selected them to highlight that preregistration may be beneficial in cases that are not as straightforward as quantitative confirmatory research for which there are already so many examples within and outside our community. Our hope is that we can start discussing the value of preregistration as a *tool to increase transparency* and *improve reporting* rather than just a tool to make research reproducible and limit questionable research practices. We acknowledge that visualization research has many flavors, and papers often combine multiple methodologies or epistemologies. While we would have liked to be able to offer guidelines and examples for all relevant research contributions and methodologies, we realize that this is an effort that would go beyond the knowledge of our authoring team. As such, we hope to be able to leverage the community's input during and after the workshop to continuously craft guidelines, reflect on the limitations of preregistration and its current set of templates, and collect examples of what we believe to be good preregistrations. Ultimately, we may even manage to find a visualization mantra that could help us, as a very heterogeneous community, to reflect and ponder what eventually impacts our findings and/or the contents of our publications. To achieve this goal, we would like to offer the possibility for the visualization community to weigh in and edit our Google Doc which is linked on our OSF project's wiki at <https://osf.io/wes57/>.

6 CONCLUSION

The visualization community incorporates a diverse set of methodologies and epistemologies, all of which benefit from research transparency. We organized a panel to discuss the merits and limits of one way to promote transparency—preregistering our study designs. Preregistering the study design before collecting data can help alleviate issues around questionable reporting practices and increase research robustness, including by combating publication biases towards positive results. Preregistration intentionally limits

the degrees of freedom researchers have when collecting data and analyzing the results, but changes to the plan are allowed—if they are communicated transparently. This can help reduce HARK-ing, selective reporting, and p-hacking. Moreover, preregistration encourages researchers to think through their study design in advance and write it down clearly and completely for readers to understand and future researchers to build upon.

Preregistration is most useful for confirmatory studies, including quantitative analyses, but it is also relevant for qualitative research that has a clear confirmatory goal. Moreover, preregistering exploratory research plans can help clarify the distinction between findings based on existing knowledge and any new divergences. It can also be used to systematically log design iterations to expose and justify design decisions. However, preregistration is not the right tool for achieving research transparency in all cases. Preregistration is less useful for design studies, applied visualization research, system development, non-exhaustive reviews, and theoretical work. It is also not a silver bullet for conducting high-quality research, as study designs can still be flawed or poorly specified.

We argue that research transparency should be a research topic in our community in its own right. The topic of research transparency is of the utmost importance to solidify the credibility of research findings as a whole and encompasses several aspects that would go beyond the scope of this paper (e.g., Open Data, Open Source, Open Reviews, positionality statements, ... see [8, 28]). Our diverse knowledge and experience in visualization research can help us develop more appropriate preregistration templates for the varied types of work we do, including qualitative and exploratory work. Our community can also contribute interactive systems and tools (e.g., [29]) to assist in making all our research transparent, extending our prior work on analytic provenance.

ACKNOWLEDGMENTS

This work is supported in part by Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation, the Knut and Alice Wallenberg Foundation (KAW 2019.0024), the Marcus and Amalia Wallenberg Foundation (MAW 2023.0130), Luxembourg's National Research Fund (FNR) grant #INTER/ANR/20/14956115, and the U.S. National Science Foundation (IIS-2145382).

REFERENCES

- [1] D. G. Altman, D. Moher, and K. F. Schulz. Harms of outcome switching in reports of randomised trials: Consort perspective. *Bmj*, 356, 2017. doi: 10.1136/bmj.j396 1
- [2] V. Amrhein, D. Trafimow, and S. Greenland. Inferential statistics as descriptive statistics: There is no replication crisis if we don't expect replication. *The American Statistician*, 73(sup1):262–270, 2019. doi: 10.1080/00031305.2018.1543137 1
- [3] M. Baker. 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604), 2016. doi: 10.1038/533452a 1
- [4] M. Bakker, C. L. S. Veldkamp, M. A. L. M. van Assen, E. A. V. Cromptvoets, H. H. Ong, B. A. Nosek, C. K. Soderberg, D. Mellor, and J. M. Wicherts. Ensuring the quality and specificity of preregistrations. *PLOS Biology*, 18(12):1–18, 12 2020. doi: 10.1371/journal.pbio.3000937 3
- [5] E. Beattie and K. Mackway-Jones. A delphi study to identify performance indicators for emergency medicine. *Emergency Medicine Journal*, 21(1):47–50, 2004. doi: 10.1136/emj.2003.001123 2
- [6] L. Besançon, A. Bezerianos, P. Dragicevic, P. Isenberg, and Y. Jansen. Publishing visualization studies as registered reports: Expected benefits and researchers' attitudes. working paper or preprint, July 2021. doi: 10.31219/osf.io/3z7kx 3, 5
- [7] L. Besançon and P. Dragicevic. The continued prevalence of dichotomous inferences at CHI. In *CHI '19 - Proceedings of CHI Conference on Human Factors in Computing Systems Extended Abstracts*, 11 pages, p. 1–11, May 2019. doi: 10.1145/3290607.3310432 1
- [8] L. Besançon, N. Peiffer-Smadja, C. Segalas, H. Jiang, P. Masuzzo, C. Smout, E. Billy, M. Deforet, and C. Leyrat. Open science saves lives: lessons from the COVID-19 pandemic. *BMC Medical Research Methodology*, 21(1):1–18, 2021. doi: 10.1186/s12874-021-01304-y 1, 5, 6
- [9] L. Besançon, A. Bezerianos, P. Dragicevic, P. Isenberg, and Y. Jansen. Publishing visualization studies as registered reports: Expected benefits and researchers' attitudes, Jul 2021. Stage 1 available under. doi: 10.31219/osf.io/3z7kx 1, 4
- [10] L. Besançon, F. Ehtler, M. Kay, and C. Wacharamanotham. The journal of visualization and interaction: A diamond open-access journal for the InfoVis and HCI communities. *Journal of Visualization and Interaction*, 1(1):3, Apr. 2023. doi: 10.54337/jovi.v1i1.7782 5
- [11] L. Besançon, A. Semmo, D. Biau, B. Frachet, V. Pineau, E. H. Sariali, M. Soubeyrand, R. Taouachi, T. Isenberg, and P. Dragicevic. Reducing affective responses to surgical images and videos through stylization. *Computer Graphics Forum*, 39(1):462–483, 2020. Preregistered data and code at <https://osf.io/34vzj/files/osfstorage>. doi: 10.1111/cgf.13886 4
- [12] L. Besançon, A. Semmo, D. Biau, B. Frachet, V. Pineau, E. H. Sariali, M. Soubeyrand, R. Taouachi, T. Isenberg, and P. Dragicevic. Reducing affective responses to surgical images and videos through stylization. *Computer Graphics Forum*, 39(1):462–483, 2020. doi: 10.1111/cgf.13886 5
- [13] M. Bostock, V. Ogievetsky, and J. Heer. D³: Data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, 2011. doi: 10.1109/TVCG.2011.185 3
- [14] B. Bowers. Pre-registering qualitative research: Benefits, unintended consequences, and unanswered questions. *Innovation in Aging*, 3(Suppl 1):S400, 2019. doi: 10.1093%2Fgeron%2F38.1483 3
- [15] C. D. Chambers, E. Feredoes, S. D. Muthukumaraswamy, and P. Etchells. Instead of “playing the game” it is time to change the rules: Registered reports at aims neuroscience and beyond. *AIMS Neuroscience*, 1(1):4–17, 2014. doi: 10.3934/Neuroscience2014.1.4 5
- [16] K. Chiu, Q. Grundy, and L. Bero. ‘spin’ in published biomedical literature: a methodological systematic review. *PLoS Biology*, 15(9):e2002173, 2017. doi: 10.1371/journal.pbio.2002173 1
- [17] A. Cockburn, P. Dragicevic, L. Besançon, and C. Gutwin. Threats of a replication crisis in empirical computer science. *Communications of the ACM*, 63(8):70–79, 10 pages, July 2020. doi: 10.1145/3360311 1
- [18] A. Cockburn, C. Gutwin, and A. Dix. HARK no more: On the pre-registration of CHI experiments. In *Proc. 2018 CHI Conference on Human Factors in Computing Systems*, 12 pages, p. 1–12, 2018. doi: 10.1145/3173574.3173715 1
- [19] A. DeHaven. Preregistration: A plan, not a prison, May 2017. 3
- [20] S. Di Bartolomeo, T. Crnovrsanin, D. Saffo, E. Puerta, C. Wilson, and C. Dunne. Evaluating graph layout algorithms: A systematic review of methods and best practices. *Computer Graphics Forum*, p. e15073, 2024. doi: 10.1111/cgf.15073 5
- [21] U. Dirnagl. Preregistration of exploratory research: Learning from the golden age of discovery. *PLOS Biology*, 18(3):1–6, 03 2020. doi: 10.1371/journal.pbio.3000690 3
- [22] W. Gatterbauer and C. Dunne. On the reasonable effectiveness of Relational Diagrams: explaining relational query patterns and the pattern expressiveness of relational languages. *Proc. ACM on Management of Data—SIGMOD/PACMOD*, 2(1), article no. 61, 2024. Preregistered data and code at <https://osf.io/4zpsk/>. doi: 10.1145/3639316 4
- [23] J. A. Grand, S. G. Rogelberg, G. C. Banks, R. S. Landis, and S. Tonidandel. From outcome to process focus: Fostering a more robust psychological science through registered reports and results-blind reviewing. *Perspectives on Psychological Science*, 13(4):448–456, 2018. PMID: 29961411. doi: 10.1177/1745691618767883 5
- [24] M. J. Grant and A. Booth. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2):91–108, 2009. doi: 10.1111/j.1471-1842.2009.00848.x 3
- [25] E. G. Guba, Y. S. Lincoln, et al. Competing paradigms in qualitative research. *Handbook of qualitative research*, 2(163-194):105, 1994. 3
- [26] S. Haroz. Open access VIS. Visited Jun. 6 2022. 1
- [27] S. Haroz. Comparison of preregistration platforms. *OSF Preprint*, 2022. doi: 10.31222/osf.io/zry2u 4

- [28] Y. Jansen, J. B. Vornhagen, O. Iarygina, K. S. Niksirat, L. Besançon, P. Dragicevic, J. Gori, and C. Wacharamanotham. The many ways of being transparent in human-computer interaction research. Technical report, OSF preprint, 2024. doi: [10.31234/osf.io/3crhg/](https://doi.org/10.31234/osf.io/3crhg/)
- [29] E. Jun, A. Seo, J. Heer, and R. Just. Tisane: Authoring statistical models via formal reasoning from conceptual and data relationships. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, article no. 490, 16 pages. Association for Computing Machinery, New York, 2022. doi: [10.1145/3491102.3501888](https://doi.org/10.1145/3491102.3501888)
- [30] A. Kale, M. Kay, and J. Hullman. Visual reasoning strategies for effect size judgments and decisions. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):272–282, 2021. doi: [10.1109/TVCG.2020.3030335](https://doi.org/10.1109/TVCG.2020.3030335)
- [31] R. M. Kaplan and V. L. Irvin. Likelihood of null effects of large NHLBI clinical trials has increased over time. *PLoS One*, 10(8):e0132382, 2015. doi: [10.1371/journal.pone.0132382](https://doi.org/10.1371/journal.pone.0132382)
- [32] M. Koval and Y. Jansen. Do you see what you mean? using predictive visualizations to reduce optimism in duration estimates. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, article no. 30, 19 pages. Association for Computing Machinery, New York, NY, USA, 2022. doi: [10.1145/3491102.3502010](https://doi.org/10.1145/3491102.3502010)
- [33] T. L. Haven and L. Van Grootel. Preregistering qualitative research. *Accountability in Research*, 26(3):229–244, 2019. doi: [10.1080/08989621.2019.1580147](https://doi.org/10.1080/08989621.2019.1580147)
- [34] M. Meyer and J. Dykes. Criteria for rigor in visualization design study. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):87–97, 2020. doi: [10.1109/TVCG.2019.2934539](https://doi.org/10.1109/TVCG.2019.2934539)
- [35] E. B. Nilsen, D. E. Bowler, and J. D. C. Linnell. Exploratory and confirmatory research in the open science era. *Journal of Applied Ecology*, 57(4):842–847, 2020. doi: [10.1111/1365-2664.13571](https://doi.org/10.1111/1365-2664.13571)
- [36] B. A. Nosek, E. D. Beck, L. Campbell, J. K. Flake, T. E. Hardwicke, D. T. Mellor, A. E. van't Veer, and S. Vazire. Preregistration is hard, and worthwhile. *Trends in cognitive sciences*, 23(10):815–818, 2019. doi: [10.1016/j.tics.2019.07.009](https://doi.org/10.1016/j.tics.2019.07.009)
- [37] B. A. Nosek, C. R. Ebersole, A. C. DeHaven, and D. T. Mellor. The preregistration revolution. *Proc. National Academy of Sciences*, 115(11):2600–2606, 2018. doi: [10.1073/pnas.1708274114](https://doi.org/10.1073/pnas.1708274114)
- [38] B. A. Nosek and D. Lakens. Registered reports. *Social Psychology*, 45(3):137–141, 2014. doi: [10.1027/1864-9335/a000192](https://doi.org/10.1027/1864-9335/a000192)
- [39] D. Park, M. Suhail, M. Zheng, C. Dunne, E. Ragan, and N. Elmqvist. StoryFacets: A design study on storytelling with visualizations for collaborative data analysis. *Information Visualization*, 21(0):3–16, 2022. Preprint: <https://osf.io/48esj>. doi: [10.1177/14738716211032653](https://doi.org/10.1177/14738716211032653)
- [40] Y. S. Park, L. Konge, and A. R. Artino Jr. The positivism paradigm of research. *Academic medicine*, 95(5):690–694, 2020. doi: [10.1097/ACM.0000000000003093](https://doi.org/10.1097/ACM.0000000000003093)
- [41] A. Peikert, M. S. Ernst, and A. M. Brandmaier. Why does preregistration increase the persuasiveness of evidence? A bayesian rationalization. *PsyArXiv*, 2023. doi: [10.31234/osf.io/cs8wb](https://doi.org/10.31234/osf.io/cs8wb)
- [42] V. Pooryousef, M. Cordeil, L. Besançon, C. Hurter, T. Dwyer, and R. Basset. Working with forensic practitioners to understand the opportunities and challenges for mixed-reality digital autopsy. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, article no. 843, 15 pages. Association for Computing Machinery, New York, NY, USA, 2023. doi: [10.1145/3544548.3580768](https://doi.org/10.1145/3544548.3580768)
- [43] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever. Robust speech recognition via large-scale weak supervision. In *International Conference on Machine Learning*, pp. 28492–28518. PMLR, 2023.
- [44] J. Rogers, A. H. Patton, L. Harmon, A. Lex, and M. Meyer. Insights from experiments with rigor in an EvoBio design study. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1106–1116, 2021. doi: [10.1109/TVCG.2020.3030405](https://doi.org/10.1109/TVCG.2020.3030405)
- [45] D. Saffo, S. Di Bartolomeo, C. Yildirim, and C. Dunne. Remote and collaborative virtual reality experiments via social VR platforms. In *Proc. CHI Conference on Human Factors in Computing Systems*, CHI, 15 pages, 2021. Preprint & supplemental material: <https://osf.io/3crhg/>. doi: [10.1145/3411764.3445426](https://doi.org/10.1145/3411764.3445426)
- [46] K. Salehzadeh Niksirat, L. Goswami, P. S. B. Rao, J. Tyler, A. Silacci, S. Aliyu, A. Aebli, C. Wacharamanotham, and M. Cherubini. Changes in research ethics, openness, and transparency in empirical studies between chi 2017 and chi 2022. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, article no. 505, 23 pages. Association for Computing Machinery, New York, NY, USA, 2023. doi: [10.1145/3544548.3580848](https://doi.org/10.1145/3544548.3580848)
- [47] A. Sarma, S. Guo, J. Hoffswell, R. Rossi, F. Du, E. Koh, and M. Kay. Evaluating the use of uncertainty visualisations for imputations of data missing at random in scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):602–612, 2023. doi: [10.1109/TVCG.2022.3209348](https://doi.org/10.1109/TVCG.2022.3209348)
- [48] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer. Vega-lite: A grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):341–350, 2017. doi: [10.1109/TVCG.2016.2599030](https://doi.org/10.1109/TVCG.2016.2599030)
- [49] M. Schapira, R. J. Harding, O. L. N. Consortium, et al. Open laboratory notebooks: good for science, good for society, good for scientists. *FI1000Research*, 8, 2019.
- [50] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012. doi: [10.1109/TVCG.2012.213](https://doi.org/10.1109/TVCG.2012.213)
- [51] R. Skånberg, I. Hotz, A. Ynnerman, and M. Linares. VIAMD: a software for visual interactive analysis of molecular dynamics. *Journal of Chemical Information and Modeling*, 63(23):7382–7391, 2023. doi: [10.1021/acs.jcim.3c01033](https://doi.org/10.1021/acs.jcim.3c01033)
- [52] A. Szollosi, D. Kellen, D. J. Navarro, R. Shiffrin, I. van Rooij, T. Van Zandt, and C. Donkin. Is preregistration worthwhile? *Trends in Cognitive Sciences*, 24(2):94–95, 2020. doi: [10.1016/j.tics.2019.11.009](https://doi.org/10.1016/j.tics.2019.11.009)
- [53] O. van den Akker, M. Bakker, M. A. van Assen, C. R. Pennington, L. Verweij, M. Elsherif, A. Claesen, S. D. M. Gaillard, S. K. Yeung, J.-L. Frankenberger, et al. The effectiveness of preregistration in psychology: Assessing preregistration strictness and preregistration-study consistency. *MetaArXiv*, 2023. doi: [10.31222/osf.io/h8xjw](https://doi.org/10.31222/osf.io/h8xjw)
- [54] M. van der Naald, S. Wenker, P. A. Doevendans, K. E. Wever, and S. A. Chamuleau. Publication rate in preclinical research: a plea for preregistration. *BMJ open science*, 4(1), 2020. doi: [10.1136/bmjos-2019-100051](https://doi.org/10.1136/bmjos-2019-100051)
- [55] I. van Rooij. Psychological science needs theory development before preregistration. <https://featuredcontent.psychonomic.org/psychological-science-needs-theory-development-before-preregistration/>, 2019. Online; accessed 2024-07-05.
- [56] C. Wacharamanotham, L. Eisenring, S. Haroz, and F. Ehtler. Transparency of CHI research artifacts: Results of a self-reported survey. In *Proc. 2020 CHI Conference on Human Factors in Computing Systems*, 14 pages, p. 1–14, 2020. doi: [10.1145/3313831.3376448](https://doi.org/10.1145/3313831.3376448)
- [57] C. Wacharamanotham, Y. Jansen, A. A. McNamara, K. Hornbæk, J. Robertson, and L. Goswami. Transparent quantitative research as a user interface problem (Dagstuhl Seminar 22392). In *Dagstuhl Reports*, vol. 12. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2023. <https://www.dagstuhl.de/22392>
- [58] E.-J. Wagenmakers, G. Dutilh, and A. Sarafoglou. The creativity-verification cycle in psychological science: New methods to combat old idols. *Perspectives on Psychological Science*, 13(4):418–427, 2018. PMID: 29961413. doi: [10.1177/1745691618771357](https://doi.org/10.1177/1745691618771357)
- [59] X. Wang, D. Rousseau, L. Besançon, M. Sereno, M. Ammi, and T. Isenberg. Towards an understanding of augmented reality extensions for existing 3D data analysis tools. In *CHI '20 - 38th SIGCHI conference on Human Factors in computing systems*. Honolulu, United States, Apr. 2020. doi: [10.1145/3313831.3376657](https://doi.org/10.1145/3313831.3376657)
- [60] Wikipedia contributors. Replication crisis — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Replication_crisis&oldid=1092093214, 2022. [Online; accessed 2022-06-09].
- [61] E. C. Willroth and O. E. Atherton. Best laid plans: A guide to reporting preregistration deviations. *Advances in Methods and Practices in*

Psychological Science, 7(1):25152459231213802, 2024. doi: [10.1177/25152459231213802](https://doi.org/10.1177/25152459231213802) 4

- [62] F. Yang, M. Cai, C. Mortenson, H. Fakhari, A. D. Lokmanoglu, N. Diakopoulos, E. C. Nisbet, and M. Kay. The backstory to “swaying the public”: A design chronicle of election forecast visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 1, To appear. doi: [10.31219/osf.io/927vy](https://doi.org/10.31219/osf.io/927vy) 5

- [63] F. Yang, M. Cai, C. Mortenson, H. Fakhari, A. D. Lokmanoglu, J. Hullman, S. Franconeri, N. Diakopoulos, E. C. Nisbet, and M. Kay. Swaying the public? impacts of election forecast visualizations on emotion, trust, and intention in the 2022 u.s. midterms. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):23–33, 2024. doi: [10.1109/TVCG.2023.3327356](https://doi.org/10.1109/TVCG.2023.3327356) 5