

Special Interest Group on Transparent Statistics Guidelines

Chat Wacharamanatham, Shion Guha, Matthew Kay, Pierre Dragicevic, Steve
Haroz

► **To cite this version:**

Chat Wacharamanatham, Shion Guha, Matthew Kay, Pierre Dragicevic, Steve Haroz. Special Interest Group on Transparent Statistics Guidelines. CHI 2018 - Conference on Human Factors in Computing Systems, Apr 2018, Montreal QC, France. ACM Press, CHI EA '18 Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, pp.1-4, <10.1145/3170427.3185374>. <hal-01818141>

HAL Id: hal-01818141

<https://hal.inria.fr/hal-01818141>

Submitted on 18 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Special Interest Group on Transparent Statistics Guidelines

Chat Wacharamanotham
University of Zurich
Zurich, Switzerland
chat@ifi.uzh.ch

Shion Guha
Marquette University
Milwaukee, WI, USA
shion.guha@marquette.edu

Matthew Kay
University of Michigan
Ann Arbor, MI, USA
mjskay@umich.edu

Pierre Dragicevic
Inria
Orsay, France
pierre.dragicevic@inria.fr

Steve Haroz
Sorbonne University
Paris, France
stats@steveharoz.com

Abstract

Transparent statistics is a philosophy of statistical reporting whose purpose is scientific advancement rather than persuasion. At our CHI 2017 workshop, *Moving Transparent Statistics Forward*, we identified that an important first step is to develop detailed guidelines for authors and reviewers in order to help them practice and promote transparent statistics. We propose a SIG to solicit feedback from the CHI community on a first working draft of *Transparent Statistics Guidelines* and engage potential contributors to push the transparent statistics movement forward.

Author Keywords

Statistics; methodology; user studies.

Background

HCI is a large, multidisciplinary field drawing on a variety of statistical approaches. However, many of our existing practices have drawn increasing criticism, such as our overreliance on null-hypothesis significance testing (NHST), our lack of replications and meta-analyses, and our infrequent sharing of data or study materials. In our statistical practice, there has been too much focus on persuasion and not enough on transparency. Practices going against transparency have been criticized within both HCI [2, 3, 5, 7, 10] and related fields [4, 11, 12].

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. *CHI'18 Extended Abstracts*, April 21–26, 2018, Montréal, QC, Canada. ACM ISBN 978-1-4503-5621-3/18/04. © 2018 Copyright is held by the owner/author(s). <https://doi.org/10.1145/3170427.3185374>

Such practices are thought to have greatly contributed to the *replication crisis* in social sciences and medicine, a crisis that has received popular press coverage and may harm the credibility of science.

In order to promote positive change, the authors of this SIG proposal started a “transparent statistics @ CHI” movement in 2016 (transparentstatistics.org). Two CHI events were organized: the CHI 2016 Special Interest Group on Transparent Statistics [8] and the CHI 2017 workshop on Moving Transparent Statistics Forward [9]. At these two events, a series of incremental changes were proposed for the CHI community, such as more specific author and reviewer guidelines and badges for transparent statistical practices. This new SIG aims to build upon the enthusiasm generated by these two events and move forward by furthering the development of *transparent statistics guidelines* for HCI.

The organizers and attendees of the CHI 2017 workshop began drafting guidelines for authors and reviewers in order to help improve the transparency and quality of statistical reports in the field. After the workshop, the authors of this new SIG proposal worked together to consolidate a first working draft of transparent statistics guidelines. The purpose of this new SIG is to present the working draft, solicit feedback from the CHI community, and engage existing and new contributors in further developing the guidelines. The SIG meeting format will engage a wider audience of CHI participants who are interested, but who may not necessarily have a strong background or motivation enough to participate in a workshop style as in 2017 (workshops also have fewer participants by necessity). This is especially important because we wish the guidelines to be usable by researchers who do not necessarily have strong statistical background

or interest, but still wish to adopt a transparent statistics approach. We also wish the guidelines to reflect the interests of the broader CHI community.

Guiding Principles for the Transparent Statistics Guidelines

The working draft of the transparent statistics guidelines is available at transparentstats.github.io/guidelines. The document starts by laying out general guiding principles for the guidelines themselves. The purpose of the transparent statistics guidelines is not to admonish an entire field of researchers for their existing practices or to attempt to shame them to do better. The multifaceted nature of HCI means practices will always vary, and a fixed set of DOs and DON'Ts would be both too brittle to change over time and too restrictive in the face of the various ways of generating knowledge in our field.

Instead, the guidelines aim to advance a vision of *transparent statistical communication* for the field. Whatever the methods used, we can at least provide guidance that makes the communication of those methods more transparent, that makes reproduction and replication of work easier, and that makes evaluation of work (e.g., by peer reviewers) easier and more fair. *Transparent statistics* is a philosophy of statistical reporting whose purpose is to advance scientific knowledge rather than to persuade. This quote from Ronald Fisher captures the essence of transparent statistics:

“we have the duty of [...] communicating our conclusions in intelligible form, in recognition of the right of other free minds to utilize them in making their own decisions.” [6].

The working draft lays out nine high-level guiding

principles for transparent statistics: (1) faithfulness to the data and phenomena studied, (2) robustness to departures from statistical assumptions, (3) resilience to statistical noise, (4) full explanations of analysis processes and reporting strategies, (5) clarity and accessibility of study reports, (6) preference for simplest analysis procedures, (7) avoidance of analysis decisions that are contingent on data, (8) pursuit of statistical power and precision, and (9) making study material available.

Guidelines on Specific Topics

In addition to the general guiding principles, the transparent statistics guidelines are designed around particular topics in statistics that are relevant to both authors and reviewers. At the CHI 2017 workshop, we developed a list of topics (such as experiment and analysis planning, effect sizes, p-values, Likert-scale data, data transformation, and Bayesian inference). At and after the workshop, we focused on developing a single topic—effect sizes—in order to work out issues of how to effectively structure guidelines for a particular topic. After working out the bugs and format with that first topic, we will develop chapters on the remaining topics.

The current draft of the guidelines is available at transparentstats.github.io/guidelines. For each topic, we settled on a structure that consists of a general *FAQ* (frequently asked questions) followed by a set of *exemplars*, which demonstrate analyses on hypothetical datasets and include analysis code where appropriate. The choices made in the exemplars are grounded on the FAQ and on the high-level guiding principles. The combination of a FAQ and exemplars constitute a *chapter* providing guidance on a specific topic (e.g., effect size). While these guidelines are meant to provide a path for HCI researchers to better conduct and report statistics, we acknowledge

that there are still unaddressed issues (e.g., local standards for sample sizes in HCI [1]), as well as purposefully excluded issues that only loosely couple with statistics and are therefore beyond the scope of our effort (e.g., finding good research questions, UI design, etc.).

At the SIG, we will briefly present the state of the guidelines and obtain feedback on them from participants. We will also solicit contributors interested in creating guidelines on topics other than effect size in order to expand the guidelines. To that end, we will also describe how to contribute to the guidelines.

Contributing to the Guidelines

The transparent statistics guidelines are meant to be a collaborative artefact that enjoys the consensus support of the CHI community. As such, it welcomes contributions and comments from anyone within the community. The guidelines are written in RMarkdown and hosted on Github. To lower the entry barrier and allow anyone to easily comment and contribute, the initial version of each chapter is developed on Google Docs, allowing writing and commenting directly with a WYSIWYG interface. This first stage typically consists of developing the FAQ part, and delineating an initial structure for exemplars. After the content is adequately stable, it is ported to RMarkdown and version-tracked on Github. At this stage, major contributions can be done by forking and creating pull requests, as in typical open source software development. Contributors with less technical background can also comment on the content by creating Github issues. This stage mostly focuses in developing exemplars, although the FAQ can keep being refined.

Before the release of a chapter, two reviewers with subject matter expertise from the CHI community will review the

chapter through Github's in-line reviewing/commenting interface. Then, to ensure usability, the chapter will be tested with other two members without subject matter expertise. Once a chapter is released, any member of the CHI community will have the possibility of endorsing it, even if they did not contribute. Releases are never final: the transparent statistics guidelines are meant to evolve over time and reflect evolutions in the methodological debate within and outside CHI.

A working draft of the guide for contributors is available at github.com/transparentstats/guidelines/wiki. The guidelines are developed under the CC-BY 4.0 license (for the text) and the MIT license (for the code). These licenses are flexible enough to allow any researcher to, e.g., use exemplar code to analyze their data and release their code as supplementary material.

Conclusion

We propose this SIG as an effort to continue our nascent *transparent statistics* movement and foster active development from interested researchers in the HCI community. Specifically, we will present a working draft of the transparent statistics guidelines and a working draft of the development process, both developed based on the outcomes of the CHI 2017 workshop. We will solicit members of the HCI community for feedback on the guidelines and on the processes, we will recruit volunteers for developing the guidelines further, and we will lay out a working plan for the future.

References

- [1] Caine, K. Local standards for sample size at chi. In *Proc. of CHI 2016*, CHI '16, ACM (New York, NY, USA, 2016), 981–992.
- [2] Cairns, P. Hci... not as it should be: inferential statistics in hci research. In *People and Computers: HCI... but not as we know it*, vol. 1 (2007), 195–201.
- [3] Cockburn, A., Gutwin, K., and Dix, A. Hark no more: on the preregistration of chi experiments. In *Proc. of CHI 2018*, ACM (2018).
- [4] Cumming, G. The new statistics why and how. *Psychological science* (2013).
- [5] Dragicevic, P. Fair Statistical Communication in HCI. In *Modern Statistical Methods for HCI*, J. Robertson and M. C. Kaptein, Eds. Springer, 2016.
- [6] Fisher, R. Statistical methods and scientific induction. *Journal of the Royal Statistical Society. Series B (Methodological)* (1955), 69–78.
- [7] Kaptein, M., and Robertson, J. Rethinking statistical analysis methods for chi. In *Proc. of CHI 2012* (2012).
- [8] Kay, M., Haroz, S., Guha, S., and Dragicevic, P. Special interest group on transparent statistics in hci. In *Proc. of CHI EA 2016*, ACM (New York, NY, USA, 2016), 1081–1084.
- [9] Kay, M., Haroz, S., Guha, S., Dragicevic, P., and Wacharamanotham, C. Moving transparent statistics forward at chi. In *Proc. of CHI 2017*, CHI EA '17, ACM (New York, NY, USA, 2017), 534–541.
- [10] Kay, M., Nelson, G. L., and Hekler, E. B. Researcher-Centered Design of Statistics: Why Bayesian Statistics Better Fit the Culture and Incentives of HCI. In *Proc. of CHI 2016* (2016), 1–12.
- [11] Kline, R. B., Association, A. P., et al. Beyond significance testing: Reforming data analysis methods in behavioral research.
- [12] Open Science Collaboration and others. Estimating the reproducibility of psychological science. *Science* 349, 6251 (2015), aac4716.