

This document is now locked. The content below are copied to the respective pages in the draft chapters

[Links to all chapters and their current status](#)

Input for the Transparent Statistics Guidelines

Purpose: We seek your input for the [Transparent Statistics Guidelines](#), which encourage transparency in the practice of statistics in HCI. Your input could be pointers to research articles, books, website, blog, or other materials that discuss practices or opinion on the following topics:

- [Statistical inference in general](#)
- [Bayesian Inference](#)
- [Likert-scale data](#)
- [Experiment and analysis planning](#)
- [p-values](#)
- [Data transformation](#)

How to contribute:

Example:

[Wasserstein2016 - American Statistical Association's statement on p-value](#): Section 3 provides principles on how to interpret and report p-values. (Chat Wacharamanotham, 05.06.19)

- [McShane2019 - Abandon Statistical Significance](#): argues to remove thresholding and treat p-value as continuous (Chat Wacharamanotham, 05.06.19)
 - [Anrhein2019](#) might better be broadly known (Exemplum Nomen, 05.06.19)

Instruction for contributing material:

- Add pointers to the materials in the respective section below.
 - If the material is open-access, provide a link to the material. (Select the text, right-click, and choose "Link")
 - Otherwise, upload the file [here](#), get the link to the file (right-click on the uploaded file → Copy Link Address), and add the link to your text. (See the previous point.)
- For each pointer, write a sentence or two to argue why it should be considered in the guideline
- If appropriate, please specify page/section/paragraph for us to look for.
- Add your name and date so that we can acknowledge your contribution.
- If your pointer is relevant to any other entries, consider adding it as a bullet point.
- If your pointer is applicable in multiple sections, add separate entry in each section. (Each topic will be processed a group of people. They may not look at materials from other topics.)

Instruction for contributing comments

- If your real name is not apparent on the comment, add your name in the comment itself so that we can acknowledge your contribution.
- Should you contribute as a comment or a bullet point? Rule of thumb: if you need to add a link, use a bullet point.

How will your contribution be used: We will process your input in a workshop on 26–28 June, 2019. We aim for the outcome as a draft version of some chapters to be refined and released in 2019. Your

contribution will be acknowledged in the respective chapter of the guideline such as at the top of [this page](#).

Inferential statistics (general)

Google doc draft

Bayesian Inference

Google doc draft

- The Visualization in Bayesian Workflow paper is a good tour of diagnostics for Bayesian models and could be inspiration for some reporting standards (at least for model validation): <https://arxiv.org/abs/1709.01449>
- The latest APA guidelines has some more detail on what should be in a Bayesian write-up: <https://psychology.concordia.ca/fac/kline/library/a18.pdf>
- I (Matt) believe Kruschke's book also has recommendations for reporting, I can check
- Reporting $\hat{r}$ is typical, though recent work has suggested some better convergence diagnostics: <https://arxiv.org/abs/1903.08008>

Likert-scale data

Google doc draft

- This is a recent (and fairly comprehensive) look at different types of ordinal regression models (audience is psychologists, so should be at least fairly HCI-applicable):
<https://psyarxiv.com/x8swp/>
- There is a debate about how Likert-scale data should be analyzed, in particular, about whether it is generally appropriate to analyze individual items separately (handling them as ordinal data). A representative highly-cited paper that strongly argues in favor of analyzing aggregated scores only (as ratio data) is the following: <https://thescipub.com/pdf/10.3844/jssp.2007.106.116>
However, such strict view does not seem to be shared by everyone. For example, the following authors argue among others *“that single items are appropriate when the referenced concept is singular, concrete, and understandable to the respondent.”*
<http://journalofruralsocialsciences.org/pages/Articles/JRSS%202016%2031/3/JRSS%202016%2031%203%20126-139.pdf>
-

Experiment and analysis planning

Google doc draft

Determining sample size: Caine 2016

p-values

Google doc draft

Wasserstein2016 - American Statistical Association's statement on p-value: Section 3 provides principles on how to interpret and report p-values. (Chat Wacharamanotham, 05.06.19)

- McShane2019 - Abandon Statistical Significance: argues to remove thresholding and treat p-value as continuous (Chat Wacharamanotham, 05.06.19)
- Anrhein2019 might better be broadly known (Exemplum Nomen, 05.06.19)

Reporting adequate details for figuring out the actual statistical procedure for reviewers to figure out what exact procedure is used (Chat Wacharamanotham 12.06.19)

Statcheck package for checking that p-values are reported are consistent with the specified test statistics and the degrees of freedom

At some point I (Matt) made this chart based on a memory of someone else's visualization as an alternative for p-value reporting, but I can't now remember the original source (it might have been Deborah Mayo?): <https://mjskay.github.io/uncertainty-vis-book/scratch-space.html> --- basically allows you to see p value as a function of the null hypothesis

Data transformation

Google doc draft