
Abstract
XH: Experience and/or Heuristic
Studies reported here of <quantitative empirical software engineering research> supported by <statistical analysis> generate a number of findings concerning <proper study design, analysis, and discussion>, including <the importance of precisely defining, documenting, and justifying (as necessary) the context, purpose, and method comprising the study>. They indicate that <scientific rigor> is met by <applying these guidelines>.

Question - [Method/means of development]
What is a better way to do quantitative empirical software engineering research?

Results - [Report]
The authors list a number of guidelines, as listed on the next page (because they are so worth repeating). The most interesting thing is how often the words “define” or “specify” appear in the guidelines, showing that the main thrust of their guidelines is to be precise. This enhances believability as well as repeatability.

Validation – [Experience]
The authors base their guidelines on personal experience (both as practitioners, and what they found to constitute credible research in the papers that they reviewed). In writing these guidelines, they largely adapted comparable guidelines in medical research, as reflected in a number of professional research association documents and survey papers.
The guidelines:

- Be sure to specify as much of the industrial context as possible. In particular, clearly define the entities, attributes, and measures that are capturing the contextual information.
- If a specific hypothesis is being tested, state it clearly prior to performing the study and discuss the theory from which it is derived, so that its implications are apparent.
- If the research is exploratory, state clearly and, prior to data analysis, what questions the investigation is intended to address and how it will address them.
- Describe research that is similar to, or has a bearing on, the current research and how current work relates to it.
- Identify the population from which the subjects and objects are drawn.
- Define the process by which the subjects and objects were selected.
- Define the process by which subjects and objects are assigned to treatments.
- Restrict yourself to simple study designs or, at least, to designs that are fully analyzed in the statistical literature. If you are not using a well-documented design and analysis method, you should consult a statistician to see whether yours is the most effective design for what you want to accomplish.
- Define the experimental unit.
- For formal experiments, perform a preexperiment or precalculation to identify or estimate the minimum required sample size.
- Use appropriate levels of blinding.
- If you cannot avoid evaluating your own work, then make explicit any vested interests (including your sources of support) and report what you have done to minimize bias.
- Avoid the use of controls unless you are sure the control situation can be unambiguously defined.
- Fully define all treatments (interventions).
- Justify the choice of outcome measures in terms of their relevance to the objectives of the empirical study.
- Define all software measures fully, including the entity, attribute, unit and counting rules.
- For subjective measures, present a measure of interrater agreement, such as the kappa statistic or the intraclass correlation coefficient for continuous measures.
- Describe any quality control method used to ensure completeness and accuracy of data collection.
- For surveys, monitor and report the response rate and discuss the representativeness of the responses and the impact of nonresponse.
- For observational studies and experiments, record data about subjects who drop out from the studies.
- For observational studies and experiments, record data about other performance measures that may be affected by the treatment, even if they are not the main focus of the study.
- Specify any procedures used to control for multiple testing.
- Consider using blind analysis.
- Perform sensitivity analyses.
- Ensure that the data do not violate the assumptions of the tests used on them.
- Apply appropriate quality control procedures to verify your results.
- Describe or cite a reference for all statistical procedures used.
- Report the statistical package used.
- Present quantitative results as well as significance levels. Quantitative results should show the magnitude of effects and the confidence limits.
- Present the raw data whenever possible. Otherwise, confirm that they are available for confidential review by the reviewers and independent auditors.
- Provide appropriate descriptive statistics.
- Make appropriate use of graphics.
- Define the type of study.
- Specify any limitations of the study.