Reviewer’s report

Title: Integrating and extending cohort studies: lessons from the Extending Treatments, Education and Networks in Depression (xTEND) study

Version: 3 Date: 24 May 2013

Reviewer: Scott Hofer

Reviewer’s report:

The scientific discovery determinants and risk factors associated with individual differences in developmental and aging-related outcomes is a major research priority. Results can often differ across studies of different population samples, design characteristics, measurements, and statistical analyses and often leads to seemingly conflicting results across studies. This paper provides a very useful overview of the issues and approaches for parallel and pooled data analysis of independent studies and details a number of challenges and solutions to problems typically encountered in multiple-study analysis. The paper draws from current literature on measurement harmonization and invariance and research synthesis and highlights many of the major issues for approaching multiple-study replication and pooled-data analysis. Overall, I found this paper to be an exceptional overview of the importance and issues for replication and generalizability of longitudinal research and insights into the issues that arise from simultaneous analysis of multiple study data.

I do think that the utility of individual studies remains high for answering many questions and that replication across samples, designs, and measurements is an important outcome. Pooled data analysis is difficult because of non-equivalent measurements, making coordinated or parallel analysis a much more feasible approach (e.g., Hofer & Piccinin, 2009) and this might be considered further in the paper. Considerations such as birth cohort and historical differences are raised and can present an additional confound in pooled-data analysis. The complications of pooled-data analysis also increase with international studies that differ in language.

A further distinction might be made between multiple-group analysis (with tests of invariance of parameters) and pooled-data analysis. In particular, while dummy codes can be included to evaluate study as an effect modifier within a pooled-data analysis, variance terms and other parameters are held constant by default across study samples. This is not the case in multiple-group analyses where all parameter differences can be tested for equivalency and generally provides more opportunities for cross-study comparison of measurement and model equivalence.

Below are a few additional recent papers that are relevant to pooled-data analysis and measurement harmonization that might also be considered here.


**Level of interest:** An exceptional article

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests:

I declare that I have no financial or non-financial competing interests associated with this article.