Cross Study
Operating Guidelines

Hussong 07/20/03, updated 819/2016

Table of Contents

Cross Study Staffing

Grant Aims

The project focuses on the development and evaluation of psychometric techniques to overcome measurement challenges with integrative data analysis (IDA; i.e., the simultaneous analysis of data pooled across multiple sources). The goals of the project are as follows:

Aim 1: To refine and evaluate psychometric models for harmonizing continuously-scaled substance use outcomes (e.g., symptom severity and quantity/frequency of drug use). To better accommodate the unique features of substance use measures in IDA, we recently proposed a moderated nonlinear factor analysis model. Additional research is needed to further refine this model and to evaluate its performance under a wide array of potential differences in measurement across studies.

Aim 2: To develop and evaluate psychometric models for harmonizing discretely-scaled substance use outcomes (e.g. diagnostic status). We seek to develop and empirically evaluate novel latent class models suited to this purpose.

Aims 3 and 4: To develop and evaluate psychometric models for harmonizing continuous measures of substance use (Aim 3) and diagnostic measures of substance use disorder and related disorders (Aim 4) when data are obtained from different sources across studies. Building on Aims 1 and 2, we will extend the application of novel psychometric modeling techniques for harmonizing continuous measures of substance use quantity/frequency and symptom severity as well as categorical measures of disorder to allow for variation in sources of assessment (e.g., self, peer, parent, clinician, bioassay, or archival records) across studies.

A Common Language

We will be dealing with a lot of different mini-projects within this study and as we work on publications some consistency in terms will be useful. To ease communication, I propose we use the following language to provide a consistent reference point.

REAL-U Study

Refers to all data collected with our original cohort of n=854 to address Aims 1 and 2.

MFS Study

Refers to all data collected with our original cohort of n=854 to address Aim 3 and 4.

Aims 1 and 2 Simulation Study

Refers to all data generated to address Aim 1.

Aims 3 and 4 Simulation Study

Refers to all data generated to address Aim 2.

Cross Study Server File Structure (X:\)

We will use the server space to archive all programs related to manuscripts or data management for the project. Because data are constantly being cleaned and updated (for the collected data), please do not access the data without first speaking with the management team and having permission to use the data for a given project.

NOTE: You may not use the data for a project (class presentation, conference, manuscript, example in a methods manuscript) without having an approved concept paper on file.

Administrative Files

Contains files pertaining to the accounting and grants management functions of the project as well as other correspondence with the studies.

Project Folders

Data are organized around the mini-studies within the overall grant on the server. For a filemap of specific projects, see core documentation for the study.

Staff Folders

Each staff member may have a folder on the server for storing working files, sample code, etc.. You should not store anything in this file that belongs elsewhere - so nothing in the staff folders should be a key data file or file of code that is needed to process data for a manuscript. In other words, no archiving of code or data files should happen in the staff folders.

Data Management Issues

The data management team is responsible for:

- Tracking, obtaining, storing and processing all incoming data files from the collected data studies.
- Updating documentation (in electronic and hard copy form) as received from the studies.
- Processing original data files to create master system files that pool data for a given data collection project. This does not include harmonization of variable names across scenarios within study.
- Creating code to rename and recode variables, score existing unharmonized scales, and fix errors in the data.
- Miscellaneous analyses may be needed along the way to help verify the accuracy of the data, the structure of measurement, etc..

Some guidelines for operating..

- Use the common language and in writing the variables needed, the data structure and the form of the file (text, SAS system, etc..). Each study has a naming convention and it would be useful to keep that throughout. They are detailed in the core documentation for each study.
- EVERYONE should be checking the data. This is easiest when done in conjunction with a codebook.
- Please avoid using very vague names for files that are likely to be redundant with a lot of projects (i.e., cross study, data, a person's name only) and use a date to label all created data files.
- It is fine to use working files as you develop code for master system and sub-master system files. However, when you have a final product, clean up your code, use comments often, and delete old files. If you want to keep code as a sample of programming, move it to your staff folder.

Authorship Issues

We have a mechanism for initiating a manuscript that involves the consultation of the study PI's. To facilitate data sharing, we want all to know what particular manuscripts we are working on and for each PI to have the opportunity to choose to be an author on any paper given their role in conceptualizing and designing the study. Here is the process.

1. All papers should have a coordinating study PI (Dan, Andrea or Patrick).
2. Anyone on the project may initiate a concept paper for a manuscript. Part of the concept paper is indicating all potential authors on the manuscript and defining their roles. These decisions should be made in collaboration with the coordinating study PI. Non-cross study members whom you wish to include as authors will need to be reviewed by the coordinating PI as well.
3. You should only agree to be on a concept paper if you will devote the time to doing so. As is consistent with APA guidelines, RA duties are unlikely to be sufficient for authorship, and work on manuscripts for authorship may extend beyond the length of the RA. Data management and analyses may be insufficient for authorship, but these issues should be addressed at the time work is assigned and concept papers are formed.
4. An initial order of authorship will be proposed on the concept paper. However, this order may change throughout the work of the project, with the final order reflecting contributions to the final manuscript. All study PIs have the option of authorship on manuscripts. In addition, side manuscripts that originate from a given concept paper should become their own concept paper after conceptualization. Authorship on these side manuscripts should reflect work throughout the project that contributes substantively to the final manuscript.
5. Data may not be used for any class presentation, talk or poster presentation, or manuscript or written communication without approval from the project. ALL DATA are confidential. Please do not store data on non-project computers.
6. For each proposed manuscript, Andrea will circulate concept papers she receives to the PI's that lists the hypotheses, needed data and variables, etc.. She will ask each of the PIs to review it and to let her know about what level of involvement they might be interested in having in the manuscript (see attached concept paper). To keep things moving along, we would ask that PIs reply to our concept papers within a two week period.
7. If PIs have concerns about any of the ideas in the concept paper or want further clarification, simply contact Andrea and we'll talk about it on a case by case basis.
8. We will keep in regular contact with authors depending on their level of contribution during the manuscript development process (via email and phone conferences).
9. We will distribute drafts of manuscripts to all contributing authors prior to submission. Because we may end up with a number of authors on some of these papers, I'd again ask that manuscript review occur within a three week period.
10. In addition to acknowledging the cross-study grant in publications, we will also acknowledge the supporting grants from which the contributing datasets are drawn.
11. If progress on a manuscript is delayed beyond the proposed deadline, then the lead author should seek an extension. The extension may or may not be granted, depending on a number of factors, so delays should be avoided and the PIs should be kept in the loop.

CONCEPT PAPER RESPONSE FORM

A. To be completed by the proposing author:

Provisional Paper Title _____

Proposed First Author _____

Other UNC co-authors _____

_____

Intended Submission Date _____

Studies to be used ☐ REAL-U ☐ MFS ☐SIM I ☐SIM II

B. To be completed by PIs:

☐ Let's discuss, I have concerns

Comments: _____
_____
_____
_____

Please check your contribution(s) for authorship:

_____ Conceptualizing and designing the study

_____ Conceptualizing and designing the specific paper project

_____ Statistical analysis

_____ Writing

_____ Reviewing updates to analysis and paper development prior to a first draft

_____ Reviewing manuscript drafts

_____ Acknowledgement only, I will not be a co-author

Signature _____

Concept Paper Template

Provisional Paper Title _____

Proposing Author _____

Objective of the Study:

Variables needed: