

**Final Report: Assessing the impact of integrating open and transparent research practices
in the curriculum of Psychology research methods**

Alexia Galati & Doug Markant

Department of Psychological Science

A. Summary of project activities and what we've learned: The aim of this project is to develop and validate the *Open Science Concept Inventory (OSCI)*, an instrument that will be used to assess conceptual understanding of open, transparent, and reproducible research practices. New course content in this domain will be introduced into the curriculum of *PSYC2103: Research Methods II (RM2)* in AY 2019-2020, in the form of an open science course module and lab materials. Assessment of learning gains in this domain will take place across two RM2 sessions in Fall 2019 and Spring 2020 by administering the OSCI prior to and after introduction of this new curriculum. While supported by this Scholarship of Assessment grant this summer, we focused on two main aspects of the project: (1) stage 1 of the development of the OSCI (Study 1), and (2) the development of the new curriculum for RM2.

Design of Study 1. We began by identifying a list of target concepts related to open science and questionable research practices (e.g., publication bias, p-hacking, HARKing). We then created 41 vignettes describing hypothetical scenarios faced by a researcher, science consumer, or other stakeholder (see attachment OSCI_Study1.docx). For example, a scenario might depict a researcher engaging in a questionable research practice in order to increase their likelihood of obtaining a significant effect during statistical analysis. Each vignette was followed by a question prompting respondents to express their opinion about what that person should do in the situation.

During this design phase, we had two in-person meetings with Drs. Elise Demeter and Karen Singer-Freeman, during which we discussed how to improve the readability of this vignettes and minimize ambiguity in their interpretation. Following these discussions, we revised the survey in a number of ways. Specifically, we provided a set of definitions of key terms (e.g., hypothesis, sample size, statistical significance, pre-registration) at the beginning of the survey, in order to provide necessary background while offloading the verbiage in individual vignettes. Moreover, we ensured that the prompts following the vignettes had a similar syntactic and logical structure (namely, "*Do you agree with how X [...] ? / Would you advise X to [...] ? Why or why not?*"). Toward this end, in some cases, we unpacked a scenario into

two (or more) decision dilemmas faced by the researcher, with separate prompts corresponding to each decision point. Finally, we piloted the survey with our research assistants to assess its duration and feasibility. Since according to our RAs the survey took up to 2 hours to complete, we divided the items into two lists (of 20 and 21 items) to shorten the duration of the study and reduce the burden on our participants. The two lists were reasonably matched in terms of the topics covered, the complexity of the prompts, and the incidence of multiple prompts associated with the scenario (see p. 14 in attachment OSCI_Study1.docx). Within each list, the presentation order of the items is randomized (with the constraint that multi-part vignettes would be presented in sequence) to mitigate fatigue and other order effects.

Deployment of Study 1. We launched Study 1 in late June and have collected data from 43 participants recruited from the SONA subject pool and from the UNC Charlotte undergraduate student body. We are now aiming for 60 participants (30 per list). Based on our qualitative observations and initial coding of participants' open-ended responses, a sample size of 60 should be adequate for generating the multiple-choice options of the OSCI. Already, even with 43 participants, responses fall consistently into categories, permitting the creation of target and distractor responses in a fairly straightforward manner.

Coding of Study 1 responses. We began preliminary coding of the 43 participants' open-ended responses in order to establish the protocol for generating multiple choice options for the OSCI. We have developed the following procedure for generating MC options for a given vignette:

1. Group participants' responses under consistent themes; these are identified by reading the entire set of open-ended responses. Summarize each theme as a statement capturing the common thread of this group of responses (e.g., "I would advise David to not abandon the project because non-significant findings are informative and should be published"). Whenever possible, use similar language to that generated by the respondents. If a given response addresses more than one theme, cross-list it with relevant phrases highlighted under each classification. If a given response does not fit into any of the emerging themes, classify it as "other/unclassified".

2. Identify the correct response category for this vignette among the generated options. If needed, modify the wording to align the correct response better with our desired student learning objectives regarding this theme.
3. Exclude from the selection process those response categories that may be broadly correct but don't capture the target response for this prompt (e.g., "I would advise David to not abandon the project because additional replications are needed to establish if there is a true effect."). The rationale of this decision is that we want to have MC options that clearly distinguish between correct and incorrect responses.
4. Include in the selection process those response categories that are clearly incorrect (e.g., "I would advise David to not abandon the project because he should change his methodology to see if he finds a significant effect"). If needed, modify the wording to minimize ambiguity in their interpretation.
5. If there are fewer than 3 response categories that could serve as distractors, generate additional ones to yield a total of 3-4 distractor options.

Next step: Study 2. Once we adapt these items into multiple choice questions, each involving a prompt followed by 4-5 response options, we will administer this questionnaire to 250 Psychology undergraduates (Study 2). Again, the items will be presented in the two lists used in Study 1 (with 125 participants per list) with the same randomization constraints. We will use Item Response Theory (IRT) analysis (Veilleux & Chapman, 2018) to select items that vary in difficulty and are high in discrimination. These will constitute the final items to be included in OSCI.

B. Integration into the course curriculum. In parallel to running Study 1, we have drafted new lecture material (including slides and script) for the new Open Science modules for RM2. The lectures are structured around three student learning objectives: (1) Recognize problems caused by current incentives and norms in science (Lecture 1); (2) Identify questionable research practices that undermine the robustness of the scientific literature (Lecture 1); and (3), Identify and select open science solutions to address these problems (Lecture 2). We plan to make video recordings for both lectures early in the Fall 2019 semester.

The new material will be implemented in two sections of RM2 (approximately N=50 students) in Fall 2019 (Study 3) and in Spring 2020 (Study 4), and evaluated using the OSCI and the Attitudes Toward Research Questionnaire (ATRQ). The ATRQ is adapted from Kardash (2000) and Chopik et al. (2018) to measure students' self-efficacy and attitudes toward research. As we have proposed, the OSCI will be used to evaluate learning gains at two time points: A pre-test during the first 3 weeks of the semester, and a post-test following completion of the open science modules, approximately 2-3 weeks prior to the end of the semester. We will analyze within-subjects learning gains following the introduction of the new curriculum (Post-test - Pre-test). Based on the findings of Study 3, we will revise the materials and evaluation plan for the second implementation round in Spring 2020 (Study 4).

C. Impact on students or anticipated impact. We expect that students in the 4 RM2 sections across Studies 3-4 will experience learning gains in their conceptual understanding of robust and reproducible research practices, as evidenced in the change in their OSCI scores. We also anticipate that students will report increased self-efficacy in terms of their perceived ability to complete certain research activities (e.g., generating a hypothesis, conducting statistical analysis, etc.) and increased confidence in psychology as a field that follows the scientific method; these improved attitudes will be evidenced in their ATRQ scores. The new materials and assessments developed during this project will also be made openly available for future RM2 sections, facilitating the inclusion of open science practices in the research methods curriculum.

D. What we plan to do differently in the future. We have learned that the development of assessment tools can be very time consuming: generating the vignettes for Study 1 proved to take a lot of time and inspiration, as well as multiple rounds of revision. In light of that, we will adjust our timelines for the development of similar tools in the future. In addition, we anticipate that the OSCI assessment may require additional revision as we fine-tune our course materials and observe students' performance during the first implementation round. We intend to continue to solicit feedback from faculty and staff to ensure the validity of the OSCI and to increase the likelihood that these materials will be adopted by other instructors.