A few years ago I made changes to the syllabus for my graduate course on research methods in social psychology in response to what is widely referred to as the “replicability crisis”. The biggest change was to remove the typical “research proposal” requirement that students completed individually and replace it with a hands-on group replication project (you can view the syllabus of the course here: https://osf.io/nxytf/). In 13 weeks (the length of the course), the group replication project proceeds as follows:

- Week 1: Groups of up to 4 students are established. Their first task is to identify social psychological research published within the past few years that (a) collected data online (e.g., via a University subject pool, or Amazon’s Mechanical Turk) to minimize time on data collection, and (b) is of interest to group members. Students typically share their research interests when introducing themselves to the class, and form groups with students that share similar interests (e.g., close relationships, stereotypes and prejudice, judgement and decision making).

- Week 2: as a class, we review the research identified by each group, paying close attention to study details (e.g., is it possible to reproduce the study methods and procedures from the methods section alone?; do descriptions of statistical models and results make sense?). Following this discussion, we select one published study for each group to closely replicate.

- Week 3: Each group writes a “replication report” that attempts to reproduce as best as possible the exact methods, procedures, and statistical models used in the study to be replicated. Groups make note of important study information that is not available in the published manuscript (e.g., instructions for participant recruitment, items for scales created by the authors, how covariates were coded and entered into statistical models). Next, groups draft an email to be sent to the corresponding author discussing the project. After I review and edit the email, the replication report is sent along with this email as well as questions asking for additional information as needed (example of email sent to corresponding authors: https://osf.io/q7ps8/; example of replication report: https://osf.io/nrkej/).

- Week 4: Groups edit the replication report as needed based on feedback from the corresponding author. For the 5 groups, across two years, that have worked on these projects the corresponding authors have responded with very helpful feedback within days in every instance so far! Groups then submit a research ethics application, and wait (example ethics application: https://osf.io/bncq4/). It typically takes 4-6 weeks to obtain ethics approval.

- While awaiting ethics approval, groups prepare all study materials and procedures on Qualtrics. Groups are also encouraged to prepare data analytic code for the primary analyses. Shortly after receiving final ethics approval, groups pre-register study hypotheses on the Open Science Framework (OSF; e.g., https://osf.io/exnqu/). This typically occurs around week 10.
Following pre-registration, data collection begins via Amazon’s Mechanical Turk (other options are available, but we have used Mturk to this point). Data is typically collected with 48 hours. To date I have funded data collection for each project.

Each group analyzes their data, conducts follow-up analyses as needed, and prepares a written report to be submitted at the end of the course (week 13).

After all reports are graded, and the course is officially over, I meet with each group to discuss the next steps for their projects. If needed we contact the corresponding author of the original study to ask for additional study information. In all instances we collect at least one additional large sample of participants to provide a more thorough test of hypotheses (in some instances we have collected three or four additional samples, adding additional conditions or sampling from a different population). When appropriate the results of the replication samples and original study are meta-analyzed (e.g., https://osf.io/y86zp/).

I work with each group to prepare a manuscript for publication in a peer reviewed journal. Presently one manuscript is in press at the Journal of Experimental Social Psychology (see pre-print here: http://ir.lib.uwo.ca/psychologypub/102/; two replication samples with ~500 participants total). Four other manuscripts will be submitted very soon; these four manuscripts attempt to replicate four different published studies, and contain data from 12 replication samples with over 2000 participants.

Lessons Learned by the Students:

1. It is not possible to completely reproduce the methods and procedures of the published studies selected to date without requiring additional information from corresponding authors. This surprises the students, and allows for in class discussions of how to make their own research more reproducible by others (e.g., posting study materials and procedures on the OSF).
2. The data analytic approach used in the original research is also often not described in enough detail to reproduce without asking questions of the corresponding author.
3. The students take great care to work with original authors in a respectful manner, and to reproduce the study design as closely as possible. This approach has resulted in the original researchers being very helpful and responsive.
4. When the results from their replication attempts (with large samples) are not similar to the published results the students are genuinely surprised.* This allows for a class discussion on the possible reasons for the inability to replicate published results. It also highlights to them the importance of replicating their own results, when feasible, prior to submitting manuscripts for publication.
5. The students learn that the methods and results sections of published papers are very, very valuable. For example, they now focus more attention on statistical power, if the paper includes a direct/close replication or not, and the effect size of key findings. They also pay more attention to the reproducibility of the methods and procedures as discussed in the paper. If a given study was low powered, not replicated, and yielded a large effect size, and furthermore if the methods and procedures as discussed are not reproducible, they are now more likely to question the conclusions of the authors as put forward in the discussion section.

Overall, the students develop what I feel is a healthy skepticism of results from one study (someone else’s study or their own). They learn the importance of statistical power first hand, and that published effect sizes may overestimate the true effect size. They also learn to value the importance of sharing study materials and procedures to increase the reproducibility of their own research. In my opinion, lectures and readings alone are not as effective at teaching these lessons compared to running an actual close replication of published research.

* Of the five replication projects completed in my class to date (containing a total of 13 large independent samples), evidence consistent with the original published research has been obtained in 1 project (20%)