PSYCHOLOGY
LEARNING & TEACHING

Teaching Research in Principle and in Practice: What Do Psychology Instructors Think of Research Projects in Their Courses?

Jordan R. Wagge (iD)
Avila University, USA

Michelle A. Hurst (iD)
University of Chicago, USA

Mark J. Brandt
Michigan State University, USA

Ljiljana B. Lazarevic (iD)
University of Belgrade, Serbia

Nicole Legate
Illinois Institute of Technology, USA

Jon E. Grahe (iD)
Pacific Lutheran University, USA

Abstract

Psychology majors typically conduct at least one research project during their undergraduate studies, yet these projects rarely make a scientific contribution beyond the classroom. In this study, we explored one potential reason for this—that student projects may not be aligned with best practices in the field. In other words, we wondered if there was a mismatch between what instructors teach *in principle* and what student projects are *in practice*. To answer this, we asked psychology instructors ($n = 111$) who regularly teach courses involving research projects questions about

Corresponding author:
Jordan R. Wagge, School of Psychology and Cognitive Science, Avila University, Kansas City, Missouri, USA.
Email: jordan.wagge@avila.edu

these projects. Instructors endorsed many of the commonly assumed pitfalls of student projects, such as not using rigorous methodology. Notably, the characteristics of these typical student projects did not align with the qualities instructors reported as being important in research practice. We highlight opportunities to align these qualities by employing resources such as crowdsourced projects specifically developed for student researchers.

Student research projects are an essential part of undergraduate education in psychology, providing hands-on experience with the research process and more direct experience working with the instructor (American Psychological Association, 2012). For some students, these projects may provide their only opportunity to engage in research (Banger & Brownell, 2014). However, the finished products vary in their rigor and scientific contribution. In the current study, we explore what these research projects typically look like and whether they align with instructors' views of best practices when conducting psychological research. We do this to provide insight and suggestions for how student projects can be better utilized for both pedagogy and scientific advancement.

Dissemination of Student Research

According to Perlman and McCann (2005), only 10% of data from undergraduate research projects are presented outside of class in any format, representing a missed opportunity for data collection from one of the most massive data collection machines in the world—undergraduate social science students. The number of students that spend time and energy on research projects is considerable, and although we assume the primary purpose of this data collection is pedagogical and that not all projects should be made publicly available by default, is it possible for students to achieve the same or better learning outcomes while also contributing this time and energy to advance the field?

Understanding the range of quality in student projects may be essential for supporting the potential contributions to the field by students. Informally, we (the authors), as academics and professors of psychology, sometimes hear certain assumptions about student research projects: that student hypotheses can be uninteresting and not grounded in theory, that student-developed methods are not sound, that student data collection is underpowered, and that data from student projects ends up in the file drawer (which is not necessarily a bad place for the data to end up if most of these other intuitions are true). These intuitions about student projects are echoed in the literature. As Frank and Saxe (2012, p. 601) say, "...the resulting experiments are often silly, poorly designed, and unlikely to connect to current issues in psychological science." Standing and colleagues (2014) agree, stating: "Proposed projects may show little theoretical understanding, knowledge of the literature, or scientific awareness. Too often, the experimental hypothesis is unclear or has been created by the teacher, with methodological flaws sometimes evident, and samples are usually so small that statistical power is low" (p. 98). Grahe and colleagues (2012) go a step further and highlight that this data is then "'lost' to the discipline" (p. 605).

If these observations of student research projects in psychology are correct, then two things hold. First, students might benefit from having options for projects that will be disseminated and contribute to the field. Meeting student learning outcomes is the ultimate instructional goal, but as we

discuss below, it may be possible to meet these outcomes *through* projects that are *also* more likely to be rigorous and disseminated. This may be where instructors diverge based on pedagogy: if we want students in these courses to be a part of the scientific method, which parts of the scientific method are prioritized? Reporting data publically is a key step in the process, so the opportunity to contribute to something that could be disseminated would elevate that part of the process for students.

Second, we may be adopting a "do as I say, not as I do" approach to teaching research skills. For example, we might say good studies must be well-powered but then oversee student projects that are not. This misalignment could be caused by a variety of factors, many of which are structural, such as lack of time, class size, or limited access to participants. Course, departmental, or university requirements could limit what can reasonably be done. Additionally, there is often so much to accomplish in research courses that instructors simply have to prioritize which parts to emphasize (such as having students test their own hypotheses) and which parts to heavily annotate with caveats (such as the limits of small sample sizes). Regardless of the cause, the misalignment between what we teach *in principle* and what we allow student projects to be *in practice* may result in students not adequately engaging with or learning about the things we *say* are important.

Recently, scholars have developed models for supporting student research to take advantage of these missed opportunities and improve the student research experience. For example, the Collaborative Replications and Education Project (CREP; Wagge et al., 2019a) crowdsources students' direct replications of recent, high-interest work (see Ghelfi et al., 2020; Hall et al., 2018; Leighton et al., 2018; Wagge et al., 2019b for examples of CREP replication projects). Similarly, Emerging Adulthood Measured at Multiple Institutions 2 (EAMMi2; Grahe et al. 2018) produced a large dataset with many measures related to early adulthood and the 2016 election for students to explore. However, to ensure that these models meet students' needs in a way that aligns with instructors' preferences and beliefs about what is important, we need to understand what instructors are currently doing, what they value, and whether there is indeed a mismatch between the two.

Open Science Education

The recent emphasis on replications (e.g., Open Science Collaboration, 2015), the flexibility of many researcher decisions (e.g., Simmons et al., 2011), and the frequency with which researchers engage in questionable research practices (e.g., John et al., 2012; Sacco et al., 2018) has further highlighted the importance of including scientific best practices in undergraduate education (e.g., Grahe et al., 2020; Morling & Calin-Jageman, 2020). Despite the advantages of recent methodological reforms, some of these practices, including open data, open materials, and high power, may not yet be as emphasized in student research projects. In the current study, therefore, we were particularly interested in attitudes and beliefs about scientific best practices for reproducible research, a topic of increasing importance and discussion in psychology.

The Current Study

In this project, we had three main goals: First, to describe what student research projects generally look like in terms of both the type of projects and the kinds of research activities in which students engage. Second, to investigate instructors' intuitions about student-led projects, both in terms of their agreement with the commonly assumed pitfalls of student research described above and

their opinions about what is important. Third, to investigate whether instructors' feelings about what is important in the field align with what students in their classrooms do.

Method

Participants

We recruited psychology instructors by word-of-mouth on social media (Twitter, Facebook, and the Listserv for APA's Division 2, the Society for the Teaching of Psychology). We had 118 initial respondents. Early in the survey, participants reported whether they regularly (at least once every two years) teach a psychology course at the undergraduate college level that includes a research project as part of the coursework. Participants who answered "no" ($n = 4$) or did not respond ($n = 3$) were excluded from further analysis for a final sample of 111. We did not collect data related to region or country, however, the pattern of responses and spontaneous write-in responses suggest that most respondents were likely from US institutions with some from the UK and Europe as well. The sample consisted of 34 men, 64 women, and 13 with missing gender data. The instructors' average age ($n = 98$ reported) was 44.70 years old ($SD = 11.50$). Self-reported ethnicities were 75% White, 3% Latino/Hispanic, 5% other reported, and 18% not reported. Most respondents were from small liberal arts colleges ($n = 48$; 43%), followed by research institutions ($n = 28$; 25%), community colleges ($n = 6$; 5%), and "other" ($n = 20$; 18%), with 9 participants not responding. On average, instructors who responded to a question regarding time spent teaching ($n = 102$) had been teaching for 15.1 years ($SD = 10.11$). Lastly, participants ($n = 107$ responded) taught an average of 3.87 ($SD = 3.96$) sections that included a research project in the last two years and those sections included an average of 36.69 students ($SD = 92.41$, $Median = 20.00$).

A closing date governed our stopping rule for data collection; we collected all data between July 30 and August 18 of 2018. Because this was exploratory research, we had no specific *a priori* sample size in mind but had been hoping to collect data from at least 100 participants, which we were able to meet.

Materials, Data, and Analysis code are available on the Open Science Framework [https://osf.io/s62kx/?view_only=336910f887c24354bca6c0ec3da6ab83]

Procedure and Materials

The [Avila University] Institutional Review Board approved the study. Data were collected using the SurveyMonkey platform. After reading and agreeing to the informed consent, participants answered several groups of questions about their attitudes toward scientific best practices, characteristics of projects in their classes, their beliefs about those projects, and demographic questions. The list of items used in the study is presented in Table 1.

Results

All data were analyzed in R version 4.0.2 (R Core Team, 2020) with RStudio (R Studio Team, 2016) and the following R packages: *tidyr* v1.1.1 (Wickham & Henry, 2018), *dplyr* v1.0.1 (Wickham et al., 2018a), *readr* v1.3.1 (Wickham et al., 2018b), *readxl* v1.3.1 (Wickham & Bryan, 2019), *forcats* v0.5.0 (Wickham, 2020), rstatix v0.6.0 (Kassambara, 2020), *ggridges* v0.5.2 and *cowplot* v1.0.0 (Wilke, 2019, 2020).

Table 1. Survey items used.

Domain	Item	Response scale
Instruction: Please estimate the degree to which you feel each of the following is important/worthwhile in relation to the field of psychology in general:		
Opinions About Scientific Best Practices	Direct replication Conceptual replication Power analysis Open methods Open data Presentation/publication of null findings Avoiding questionable research practices	Not important at all; Slightly important; Moderately important; Very important; Not sure
Course Characteristics	Do you regularly (at least once every two years) teach a course in psychology at the undergraduate college level that includes a research project as part of the coursework?	Yes; No
	How many sections of undergraduate courses have you taught in the past two years in which there is a research project involved? (do not include lab sections if you do not teach the lab section)	Insert number
	On average, how many students were in each section?	Insert number
Project Characteristics	Think about the most recent undergraduate course you taught with a research project included. What type of project did it include?	Individual project; Group (subset of class) projects; Entire class works on the same project; Other (please specify)
	Please select the option that best characterizes the research questions addressed by the projects:	Student-generated; Instructor-generated; Replication; Other (please specify)
	Please describe the participants for the research project (check all that apply):	Classmates; Subject pool participants; Social media contacts; Nonhuman participants; Students generate fake data; Other (please specify)
Instruction: *Please estimate the percentage of students in the course who completed each of the following research components:*		
Project Characteristics	IRB Application Data collection Data analysis Writing (includes any component of a research report) Power analysis Preregistration of hypotheses Presenting in class	Percentage between 0% and 100%

Table 1. Continued.

Domain	Item	Response scale
	Presenting outside of class (e.g. conference presentation)	
	Publication in a journal	
	Sharing data in a public repository	

Instruction: *Please rank-order how important you believe it is for students to participate in each of the following components:*

Domain	Item	Response scale
Project Characteristics	Hypothesis generation	Rank order
	Data collection	
	Writing	
	Presenting	
	Data analysis	
	Data entry	
	IRB application	

Instruction: *Think about research projects that have been completed in the undergraduate courses you teach. Please rate the extent to which you agree or disagree with the following statements.*

Domain	Item	Response scale
Instructor Beliefs About Student Projects	Student-generated hypotheses are interesting.	Strongly disagree;
	Student research questions are typically grounded in theory.	Disagree;
	Students adequately review the literature prior to forming a research question.	Neither agree nor disagree; Agree;
	Student-developed methodology is scientifically sound and rigorous.	Strongly agree
	Enough data is collected to make appropriate statistical conclusions.	
	Data is collected from a diverse population.	
	Data contribute to knowledge beyond the classroom	
	Data contribute to knowledge within the classroom.	
	In your experience, what percentage of student research projects result in statistical significance?	Percentage between 0% and 100%
Demographics		
Age	What is your age in years?	Insert number
Gender	What is your gender?	Text box
Race/ethnicity	What is your race/ethnicity?	Text box
Teaching experience	How many years have you been teaching college courses?	Insert number
Institution size	How would you characterize your institution?	Small, liberal-arts college; Research institution (R1, R2, R3); Community college; Other (please specify)

Table 1. Continued.

Domain	Item	Response scale
Instructor role at institution	How would you characterize your role in your institution? Please check all that apply	Tenure-track; Non tenure-track; Assistant professor; Associate professor; Department chair/head; Adjunct instructor; Lecturer; Visiting assistant professor; Other (please specify)

Characteristics of Student Projects

The first goal of our project was to describe student research projects in general. In terms of the format of their most recent research project course ($n = 106$ responded; 5 missing), most respondents reported using either individual ($n = 39$) or small group ($n = 43$) projects, but some reported having the whole class work on the same project ($n = 15$). Those who responded with other ($n = 9$) typically included a mix of both individual and small group formats or another hybrid variation. In terms of who generated the research question ($n = 105$ responded; 6 missing), most reported that they were student-generated ($n = 65$); a smaller number reported these being instructor-generated ($n = 13$) or replication work ($n = 13$). The remaining instructors selected "other" ($n = 14$), and provided explanations that typically involved a mix of these options, either across different projects or within the same projects (e.g., students suggest sub-hypotheses or moderators within the instructor-generated larger question). Lastly, in terms of the typical participants in these research projects ($n = 104$ responded; 7 missing), the projects most often involved data from classmates ($n = 60$), a university or departmental subject pool ($n = 44$), or via social media ($n = 39$), with very few instructors selecting simulated or fake data ($n = 2$), non-human data ($n = 3$), or no data ($n = 4$). A reasonable number of participants selected "other" ($n = 33$), and provided a range of other responses including friends and family or students in other classes.

Next, we analyzed the percentage of students involved in various aspects of the research process (summary statistics are reported in Table 2 and density plots are shown in Figure 1). Most students engaged in writing, data collection, data analysis, and class presentations. The distribution of estimated percentages for writing IRB applications was bimodal, although there were more scores at the higher end. In contrast, most respondents reported low levels of engagement with hypothesis generation and presenting outside of class, and even lower engagement with power analyses, sharing data, or journal publication.

Instructors' Beliefs About Student Projects

Our project's second goal was to investigate instructors' beliefs about student projects, both in terms of what they believe to be important and their agreement with the commonly assumed pitfalls of student research described in the introduction.

First, we report instructors' rankings of the research project components that they believed were most important for students in these courses (see full rankings and summary statistics in Table 3, ordered from most to least important). Overall, both writing and hypothesis generation were highly ranked, each with

Table 2. Descriptive statistics for percentages of projects that include specific components.

Component	*n*	*M*	*SD*
Writing research report	104	97.27	12.47
Data collection	102	92.64	20.25
Data analysis	103	84.24	29.94
Presenting in class	101	78.82	36.62
IRB Application	98	59.15	46.81
Hypothesis preregistration	98	40.17	45.26
Presenting outside of class	96	30.08	38.29
Power analysis	97	23.32	34.93
Sharing data	93	8.08	23.90
Journal publication	92	4.58	15.62

Note: Because participants skipped questions, *n* is reported as the number of participants who completed that question.

Table 3. Median, mode, and number of respondents at each rank for each activity.

Activity	Median Rank	Modal Rank	Number of Respondents at Each Rank							
			1st	2nd	3rd	4th	5th	6th	7th	NA
Writing	2nd	1st	36	19	12	14	9	9	3	9
Hypothesis generation	2nd	1st	28	24	15	12	8	10	3	11
Data analysis	3rd	3rd	9	20	26	19	16	6	2	13
Data collection	4th	5th	13	17	15	18	22	14	0	12
Presenting	5th	5th	3	12	17	16	22	20	11	10
IRB application	6th	7th	8	5	7	13	13	18	35	12
Data entry	6th	7th	3	4	8	7	9	22	46	12

the top-ranked position as the modal response. The medium-ranked activities included data analysis, data collection, and presenting, followed by IRB applications and data entry at the lowest positions. Indeed, most participants placed these last two components in the bottom two positions.

Next, we report instructors' perceptions or beliefs about student projects (descriptive statistics for these statements are reported in Table 4). Overall, and as shown in Figure 2, many of the instructors reported the same intuitions as reported in prior work. For example, while many respondents felt that data contribute to knowledge within the classroom (41 and 49 out of 102 said "Strongly Agree" and "Agree," respectively), fewer believe that student-generated hypotheses are interesting (9 "Strongly Agree" and 44 "Agree" out of 102). Even fewer believe that student-developed methodology is scientifically sound and rigorous (1 "Strongly Agree" and 20 "Agree" out of 102). Similarly, instructors mostly disagree that student projects are adequately powered ("Enough data is collected to make appropriate statistical conclusions") or that research questions are grounded in theory (3 "Strongly Disagree" and 16 "Disagree" out of 102).

Opinions About Best Scientific Practices

Our project's third and final goal was to investigate whether the characteristics of student research projects actually align with instructors' reported beliefs about what is important for the field
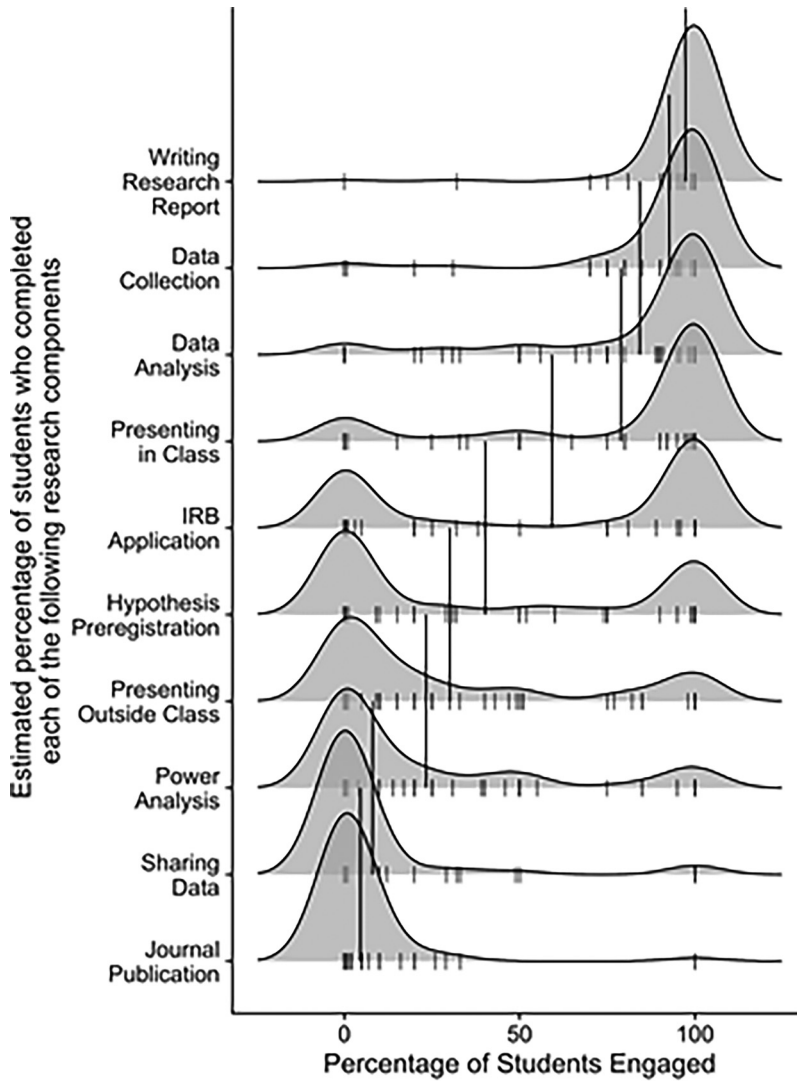
Figure 1. Density plots of the estimated percentages of students involved in each listed activity. The activities are ordered from the highest to lowest mean estimate. Individual data is displayed using the small tick marks along the bottom of the density plots and the mean percentage is indicated using the longer vertical line segment.

(Figure 3). In general, our respondents had a favorable view of open science practices, with most respondents answering "very important" to every prompt. Although not displayed in the figure, some participants also selected "not sure" (Avoiding QRPS = 2, Conceptual Replications = 0 (although 1 participant did not respond to this prompt), Open Methods = 11, Presenting Null Findings = 2, Direct Replications = 1, Open Data = 7, Power Analysis = 0)

As is evident from Figure 3, all the listed practices were rated as highly important. However, Figure 1 demonstrates that many student projects do not actually use these practices. For

Table 4. Descriptive statistics for instructor belief statements.

Statement	Median	M	SD
Data contribute to knowledge within the classroom	4 (Agree)	4.23	0.83
Student-generated hypotheses are interesting	4 (Agree)	3.42	0.94
Data contribute to knowledge beyond the classroom	3 (Neutral)	3.00	1.12
Enough data is collected to make appropriate statistical conclusions	3 (Neutral)	2.93	1.08
Student-developed methodology is scientifically sound and rigorous	3 (Neutral)	2.83	0.89
Student research questions are typically grounded in theory	3 (Neutral)	2.64	0.95
Students adequately review the literature prior to forming a research question	2 (Disagree)	2.49	0.96
Data is collected from a diverse population	2 (Disagree)	2.42	0.91

Note: 102 participants provided responses for these questions, except the item about the literature review, where there were 101 respondents. Responses of "Strongly disagree" were given a value of 1 and responses of "Strongly agree" were given a score of 5. Statements are presented in order of agreement.
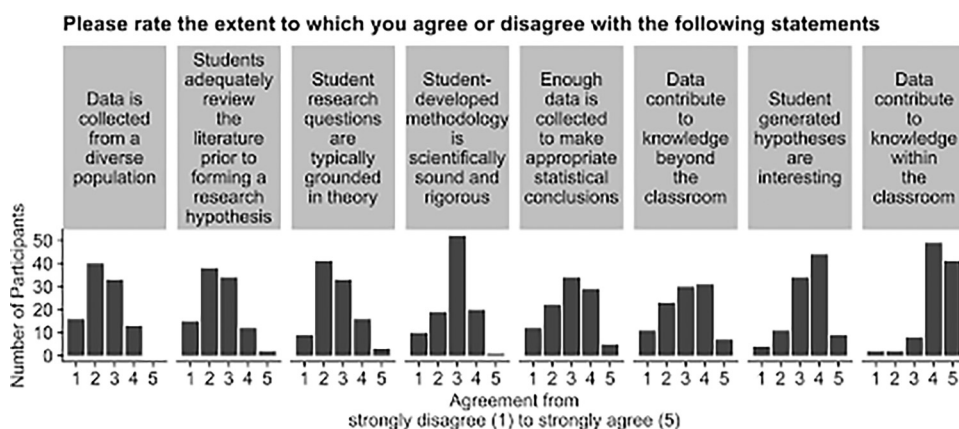

Figure 2. Bar plots of the number of participants (y-axis) who selected each agreement level (with 1 = "strongly disagree" and 5 = "strongly agree") on the x-axis, separated for each statement. The statements are ordered left to right from least to most average agreement.

example, although avoiding questionable research practices, power analysis, and open data were rated as being highly important for the field, only 40%, 23%, and 8% of students (on average, respectively) were reported to actually engage in these or related practices (e.g., preregistration to reduce questionable research practices). Moreover, respondents agree that conceptual and direct replications are important, yet only 12% reported using replications in their student research projects. Lastly, only 35% of students (on average) were reported as reporting their findings outside of the classroom (in journals, 5%, or presentations, 30%), despite the rated importance of sharing null results and that participants reported that most student projects resulted in non-significance ($n = 100$, $M = 41\%$, $SD = 23\%$). It is worth noting, though, that if these student projects are typically underpowered and poorly designed (as discussed earlier), then it may not be appropriate to share them, even as null results.

Taking these results together, the disparity between the ratings of open science practices (Figure 3) and what students actually do in their projects (Figure 1) indicates that there is ample room for student projects to align more closely with instructors' beliefs about what is important in the field.
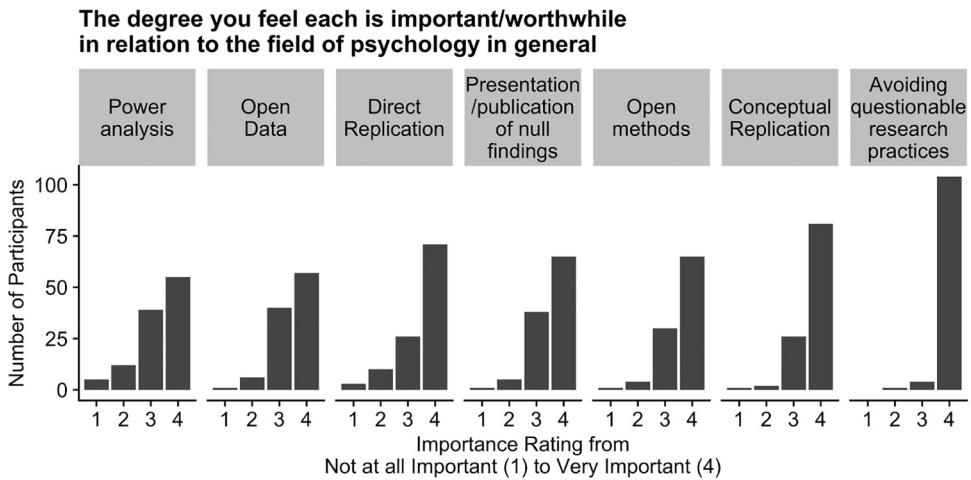
The degree you feel each is important/worthwhile in relation to the field of psychology in general

Figure 3. Bar plots of the number of participants (y-axis) who selected each importance rating from "Not at all important" (1) to "very important" (4) on the x-axis, separated for each practice. The practices are ordered left to right from least to most important based on the mean rating.

Exploratory Analyses

At the request of one of our reviewers, we conducted several exploratory analyses to better understand potential nuances or patterns in the data. The SPSS syntax used to generate all exploratory analyses can be found at https://osf.io/yqac6/?view_only=336910f887c24354bca6c0ec3da6ab83.

First, we were interested in further exploring associations between how highly instructors ranked "hypothesis generation" as an important task and responses to specific questions related to best practices (Table 1, "Opinions About Scientific Best Practices") and student-generated hypotheses (Table 1, "Instructor Beliefs About Student Projects"). We found no association between instructor responses to the statement "Student generated hypotheses are interesting" and how highly instructors ranked the hypothesis ($\tau_b = -.11$, $p = .185$). We similarly found no correlation between the ranking of student-generated hypotheses and responses to the statements "Data contribute to knowledge beyond the classroom" ($\tau_b = -.03$, $p = .734$) and "Data contribute to knowledge within the classroom" ($\tau_b = -.07$, $p = .424$). There were also no associations between hypothesis rank and the extent to which instructors say they value direct replications ($\tau_b = -.087$, $p = .370$) or conceptual replications ($\tau_b = .087$, $p = .387$), suggesting that instructors who value replications do not vary in how they value the process of hypothesis formation.

Second, we created two scales: one using the seven questions related to best science practices ($\alpha = .81$) and another using the eight questions related to student project characteristics ($\alpha = .72$). We found that these two scales were not correlated, $r(81) = -.06$, $p = .58$, suggesting that perceived quality of student projects did not vary with attitudes toward best practices in science. In other words, respondents who strongly endorse best practices in science, such as preregistration and power analysis, do not vary in their assessments of student research quality.

Third, we explored whether beliefs about student research and best practices (our two scales we created for this section) were associated with age ($M = 44.40$, $SD = 11.93$) or years spent teaching college courses ($M = 15.06$, $SD = 10.11$), and found no association between these variables (student research and age: $r(94) = .07$, $p = .51$; student research and years teaching: $r(97)$

= .02, *p* = .84; best practices and age: $r(79) = .05$, $p = .64$; best practices and years teaching: $r(83) = .07$, $p = .55$).

In the spirit of exploration, we also conducted multiple one-way ANOVAs to determine whether group type (individual, group, entire class, or "other") or project type (student-generated, replication, instructor-generated, or "other") were associated with differences in mean scores on the two scales we created for best practices and instructor beliefs. We found no differences between these groups, all *p*s > .05. In sum, there were no differences across types of projects, age, or experience teaching in opinions on best practices and beliefs about student projects, and surprisingly, these two factors were independent of one another.

Discussion

This study aimed to describe the characteristics of typical research projects in undergraduate psychology courses, instructors' attitudes about those projects, and the alignment or misalignment between instructors' beliefs about scientific best practices and the practices used in student research. Within our sample, we found that instructors acknowledge the importance of open science and related practices in psychology, but identify limitations of student research in their courses related to these practices.

Our findings align with prior research suggesting that very few student projects are shared beyond the classroom (Hauhart & Grahe, 2010; Perlman & McCann, 2005). Further, as we anticipated and in line with anecdotal intuitions and existing arguments in the literature, many instructions reported that student projects were not perceived as sufficiently rigorous (e.g., not well-powered, poorly designed) or interesting, which may be a contributing factor in why student projects are often not disseminated. Although it is not surprising that instructors prioritize specific experiences over others (e.g., high priority to writing and hypothesis generation, but low priority to data entry and writing IRB applications), an important next step for future research is to better understand what leads to these priorities and how we can make the barriers to implementing other aspects of the research process easier. For example, does hypothesis generation need to happen in the context of a research project, or can it be a smaller activity within the course? Could a replication-with-extension, whereby students generate and justify a novel hypothesis within the replication study, meet instructors' and departments' requirements for their learning outcomes?

Lastly, and most importantly, we also find that instructors' intuitions about what is important for the field as a whole were misaligned with the characteristics of their student projects. These findings suggest that there was a mismatch between what instructors teach in principle and what student projects look like in practice. In this study, we did not attempt to determine potential causes for the misalignment; it is important to acknowledge the potential role of structural limitations that instructors face, such as limited time, resources, or support.

One way to address these potential structural limitations would be to take a look at the existing structures and build solutions that either take advantage of them or work around them. There are a growing number of projects, for example, that can provide opportunities to align priorities with possibilities, including the Collaborative Replications and Education Project (CREP, Wagge et al., 2019a, 2019b), Emerging Adulthood Measured at Multiple Institutions 2, (EAMMi2, Grahe et al., 2018), or Psi Chi's Network for International Collaborative Exchange (NICE; Cuccolo et al., 2021). Projects like CREP—at least compared to a novel study—require relatively little prior knowledge for either the student or the instructor and have a high likelihood of contributing to the peer-reviewed literature.

Limitations & Future Directions

Although the current study provides important insight into the student research experience, there are some limitations that must be considered. The primary limitation of the present study is sampling

bias. First, we acknowledge that people already likely to "buy in" to open science practices were more likely to participate in this study because of our sampling methods. Second, even without this bias, instructors who feel strongly about student research projects could have been more likely to click on the link and complete the survey than instructors who do not. However, it is worth noting that even with these limitations in mind, many of the instructors still reported that student projects are not likely to use the scientific best practices that the instructors endorsed. Another limitation is that much of the survey was US-centric in wording and in the assumptions about project structure and institution types, but the country or region of the respondents was not collected. Thus these results can only generalize to institutional practices that are similar to those in the US and might not be relevant to approaches outside of the US.

Another limitation is that we did not include some of the early aspects of the research process in the survey; namely, the project characteristics we measured did include hypothesis generation but not idea formation, literature search, or literature review. These are important aspects of the research process. We want to be clear in this paper that we do not think that the models we propose should present instructors with an either/or dilemma in which they have to choose *whether* to teach some parts of the process at the cost of others, but rather we propose that these models can be used creatively in conjunction with others. In a traditional 16-week semester, for example, it would be possible for students to complete a CREP replication and then propose an additional study based on the findings and the literature, or familiarize themselves with the literature related to a CREP study and propose an extension hypothesis prior to data collection.

It is possible that whereas instructors recognize best practices, they do not implement the best practices in their classroom because the instructors themselves do not have the requisite experience or knowledge. This may be especially the case for teaching newer methods and open science practices that have only recently become part of the research culture in psychology. Although we do not have data on this issue, if this is the case it suggests that professional development and research opportunities for teachers without such experience would be valuable for implementing best practices in the classroom. Similarly, implementing the best practices in textbooks and off-the-shelf course materials may help inject best practices into the psychology curriculum.

Related, we can learn from existing projects encouraging high-quality student research opportunities to better serve the goals of undergraduate research methods courses. For instance, the CREP or NICE (or further iterations of the EAMMi) could collect information from instructors and students who complete studies as part of the project maintenance about how the project went, what could be improved, and other aspects about the course (e.g., course size, classroom resources). These data could help inform the ways these projects augment or interfere with course goals, and provide guidance to help instructors choose the project that best fits their course. For example, EAMMi might be better suited for instructors without access to participant pools (as CREP projects typically require samples of 100 participants), and the CREP might be better suited for instructors who want students to engage in subject recruitment and data collection procedures. Recognizing the different constraints of each classroom in terms of course size and available resources, our goal is to make it easier for instructors to infuse high-impact research experiences into their research methods courses.

Conclusions

Teaching research methods and related courses is difficult and time-consuming, as anyone who incorporates research experiences into their classes knows. Our goal is to uplift the pedagogical efforts of psychology instructors, by helping them consider possible mechanisms to help align learning objectives

and course outcomes with the goals and values of psychology instructors themselves. Instructors may have many valid reasons *not* to align these, such as the philosophy that students need to experience mistakes in research, or that novel projects provide them the opportunity to learn how difficult it is to design an effective study or recruit an adequate number of participants. In college courses, the learning outcomes matter most, but there may be ways we can meet learning outcomes while also meeting other goals, such as service to the field or the possibility for students to publish their work.

Perhaps small adaptations can be made that minimize increases in instructor load while maintaining (or even elevating) pedagogy and increasing the broader scientific impact of student research projects. Because no one class could do everything—or even lots of things—the instructor-prioritized experiences should be centered within the student experience of research projects. We believe that some crowdsourced science models are well-suited to elevate student learning and incorporate research experiences *while also* advancing the field.

Author Statement

Declaration of Conflicting Interests

Funding

ORCID iDs

Jordan R. Wagge https://orcid.org/0000-0002-5105-2084
Michelle A. Hurst https://orcid.org/0000-0001-6076-334X
Ljiljana B. Lazarevic https://orcid.org/0000-0003-1629-3699
Jon E. Grahe https://orcid.org/0000-0001-6970-0863

References

American Psychological Association (2012). APA Guidelines for the undergraduate psychology Major, guidelines 2.0. Retrieved From: https://www.apa.org/ed/precollege/about/psymajor-guidelines.pdf

Banger, G., & Brownell, S. E. (2014). Course-based undergraduate research experiences can make scientific research more inclusive. *CBE Life Sciences Education*, *13*(4), 602–606. https://doi.org/10.1187%2Fcbe.14-06-0099

Cuccolo, K., Irgens, M. S., Zlokovich, M. S., Grahe, J., & Edlund, J. E. (2021). What crowdsourcing can offer to cross-cultural psychological science. *Cross-Cultural Research*, *55*(1), 3–28. https://doi.org/10.1177/1069397120950628

Frank, M. C., & Saxe, R. (2012). Teaching replication. *Perspectives on Psychological Science*, *7*(6), 600–604. https://doi.org/10.1177/1745691612460686

Ghelfi, E., Christopherson, C. D., Urry, H. L., Lenne, R. L., Legate, N., Ann Fischer, M., Wagemans, F. M. A., Wiggins, B., Barrett, T., Bornstein, M., de Haan, B., Guberman, J., Issa, N., Kim, J., Na, E., O'Brien, J., Paulk, A., Peck, T., & Sashihara, M., … Sullivan, D. (2020). Reexamining the effect of gustatory disgust on moral judgment: A multilab direct replication of Eskine, Kacinik, and Prinz (2011). *Advances in Methods and Practices in Psychological Science*, *3*(1), 3–23. https://doi.org/10.1177/2515245919881152

Grahe, J. E., Chalk, H. M., Alvarez, L. D., Faas, C. S., Hermann, A. D., & McFall, J. P. (2018). Emerging adulthood measured at multiple institutions 2: the data. *Journal of Open Psychology Data*, *6*(1), 4. https://doi.org/10.5334/jopd.38.

Grahe, J. E., Cuccolo, K., Leighton, D. C., & Cramblet Alvarez, L. D. (2020). Open science promotes diverse, just, and sustainable research and educational outcomes. *Psychology Learning & Teaching, 19*(1), 5–20.

Grahe, J. E., Reifman, A., Hermann, A. D., Walker, M., Oleson, K. C., Nario-Redmond, M., & Wiebe, R. P. (2012). Harnessing the undiscovered resource of student research projects. *Perspectives on Psychological Science*, *7*(6), 605–607. https://doi.org/10.1177/1745691612459057

Hall, B. F., Wagge, J. R., Pfuhl, G., Stieger, S., Vergauwe, E., IJzerman, H., Musser, E. D., Schmidt, K., Weissgerber, S. C., Buchanan, E. M., Chen, S., Werner, K. M., Field, A., Andreychik, M., Storage, D., Voracek, M., Tran, U., Hartanto, A., Schild, C., ... Lewis, S. (2018, October 15). Accelerated CREP - RRR: Turri, Buckwalter, & Blouw (2015). https://doi.org/10.31234/osf.io/zeux9

Hauhart, R. C., & Grahe, J. E. (2010). The undergraduate capstone course in the social sciences: results from a regional survey. *Teaching Sociology*, *38*(1), 4–17. https://doi.org/10.1177/0092055X09353884

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, *23*(5), 524–532. https://doi.org/10.1177/0956797611430953

Kassambara, A. (2020). Rstatix: pipe-friendly framework for basic statistical tests (0.4.0) [computer software]. https://CRAN.R-project.org/package=rstatix

Leighton, D. C., Legate, N., LePine, S., Anderson, S. F., & Grahe, J. (2018). Self-esteem, self-disclosure, self-expression, and connection on Facebook: A collaborative replication meta-analysis. *Psi Chi Journal of Psychological Research*, *23*(2), 98–109. https://doi.org/10.24839/2325-7342.JN23.2.94

Morling, B., & Calin-Jageman, R. J. (2020). What psychology teachers should know about open science and the new statistics. *Teaching of Psychology*, *47*(2), 169–179. https://doi.org/10.1177/0098628320901372

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. Science, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Perlman, B., & McCann, L. I. (2005). Undergraduate research experiences in psychology: A national study of courses and curricula. *Teaching of Psychology*, *32*(1), 5–14. https://doi.org/10.1207/s15328023top3201_2

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

R Studio Team. (2016). *RStudio: integrated development for R*. RStudio Inc. http://www.rstudio.com/

Sacco, D. F., Bruton, S. V., & Brown, M. (2018). In defense of the questionable: defining the basis of research scientists' engagement in questionable research practices. *Journal of Empirical Research on Human Research Ethics*, *13*(1), 101–110. https://doi.org/10.1177/1556264617743834

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Standing, L. G., Grenier, M., Lane, E. A., Roberts, M. S., & Sykes, S. J. (2014). Using replication projects in teaching research methods. *Psychology Teaching Review*, *20*(1), 96–104.

Wagge, J. R., Baciu, C., Banas, K., Nadler, J. T., Schwarz, S., Weisberg, Y., IJzerman, H., Legate, N., & Grahe, J. E. (2019a). A demonstration of the collaborative replications and education project: replication attempts of the red-romance effect. *Collabra: Psychology*, *5*(1), 5. https://doi.org/10.1525/collabra.177

Wagge, J. R., Brandt, M. J., Lazarevic, L. B., Legate, N., Christopherson, C., Wiggins, B., & Grahe, J. E. (2019b). Publishing research with undergraduate students via replication work: the collaborative replications and education project. *Frontiers in Psychology*, *10*, 247. https://doi.org/10.3389/fpsyg.2019.00247

Wickham, H. (2020). Forcats: tools for working with categorical variables (factors) (0.5.0) [R package]. https://CRAN.R-project.org/package=forcats

Wickham, H., & Bryan, J. (2019). Readxl: read excel files (1.3.1) [computer software]. https://CRAN.R-project.org/package=readxl

Wickham, H., Francois, R., Henry, L., & Muller, K. (2018a). Dplyr: A grammar of data manipulation (0.7.7) [computer software]. https://CRAN.R-project.org/package=dplyr

Wickham, H., & Henry, L. (2018). Tidyr: easily tidy data with "spread()" and "gather()" functions (0.8.1) [computer software]. https://CRAN.R-project.org/package=tidyr

Wickham, H., Hester, J., & Francois, R. (2018b). Readr: read rectangular text data (1.3.1) [computer software]. https://CRAN.R-project.org/package=readr

Wilke, C. O. (2019). Cowplot: streamlined plot theme and plot annotations for "ggplot2" (1.0.0) [R program]. https://CRAN.R-project.org/package=cowplot

Wilke, C. O. (2020). Ggridges: ridgeline plots in "ggplot2" (0.5.2) [R package]. https://CRAN.R-project.org/package=ggridges

Author Biographies

Jordan R. Wagge is a Professor of Psychology and Cognitive Science at Avila University in Kansas City, Missouri, and the current Executive Director for the Collaborative Replications and Education Project, an organization that organizes and supports the completion of replication projects by students. Her current research interest focuses on the pedagogy of research methods and statistics in psychology courses.

Michelle A. Hurst is a Postdoctoral Scholar in the Department of Psychology at the University of Chicago. She studies cognitive development and is particularly interested in the cognitive and developmental processes that underlie proportional reasoning, how proportional reasoning differs across contexts, and how to best support proportional reasoning in formal math education and decision making.

Mark J. Brandt is an Assistant Professor of Psychology at Michigan State University and contributor to the Collaborative Replications and Education Project, an organization that organizes and supports the completion of replication projects by students. His current research interest focuses on moral and political attitudes and decision-making.

Ljiljana B. Lazarevic is a personality psychologist at the Faculty of Philosophy, University of Belgrade, Serbia, and the current Assistant Director of the Collaborative Replications and Education Project. She is interested in basic personality structure and personality measurement, specifically in using ambulatory assessment techniques in the measurement of basic personality traits.

Nicole Legate is an Associate Professor of Psychology at Illinois Institute of Technology and regularly supervises replication projects in her undergraduate research methods course. Her current research focuses on how to motivate sustained behavior change including motivating personal commitment to diversity, equity, and inclusion.

Jon E. Grahe is a Professor and Department Chair of Psychology at Pacific Lutheran University in Tacoma, Washington. He has been the President of Psi Chi and the Western Psychological Association, as well as a founder of and first Executive Director for the Collaborative Replications and Education Project. His current research interests include interpersonal perception, undergraduate research, capstone pedagogy, and crowdsourced science.