

MIT Open Access Articles

Preregistration and registered reports

The MIT Faculty has made this article openly available. ***Please share*** how this access benefits you. Your story matters.

Citation: Reich, Justin. 2021. "Preregistration and registered reports." *Educational Psychologist*, 56 (2).

As Published: 10.1080/00461520.2021.1900851

Publisher: Informa UK Limited

Persistent URL: <https://hdl.handle.net/1721.1/144437>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike



PREREGISTRATION AND REGISTERED REPORTS

Abstract

Preregistration and registered reports are two promising open science practices for increasing transparency in the scientific process. In particular, they create transparency around one of the most consequential distinctions in research design: the data analytics decisions made before data collection and post-hoc decisions made afterwards. Preregistration involves publishing a time-stamped record of a study design before data collection or analysis. Registered reports are a publishing approach that facilitates the evaluation of research without regard for the direction or magnitude of findings. In this paper, I evaluate opportunities and challenges for these open science methods, offer initial guidelines for their use, explore relevant tensions around new practices, and illustrate examples from educational psychology and social science.

Keywords: open science, methods, metascience, preregistration, registered reports

Preregistration and Registered Reports

When educational psychologists first begin to analyze data from a study, they cross an important threshold in the research process. Before that moment, hypotheses are predictions, positionality is shaped by prior work rather than new findings, and data analytic decisions are made without knowledge of exactly how those decisions will affect the possible set of findings to present to a public audience. Once the data analysis process begins, research decisions have different kinds of consequences. After looking at one's own data, subsequently developed hypotheses are no longer predictions; researchers' initial positionality and stance is reshaped by their encounter with data, and researchers begin to recognize how data analytic decisions will shape the presentation—and the publishability—of findings from a specific study.

Because the consequences of data analytic decisions are different before and after data analysis begins, critical consumers of educational psychology research need to understand the timing of research design decisions across this important threshold. As educational researchers, summarizing and reporting how we conduct our investigations is among our most important professional responsibilities. Our mission is to provide practitioners, policymakers, and other researchers with data, theory, and explanations that illuminate educational processes and improve the work of teaching and learning. Education stakeholders need to be able to judge the quality and contextual relevance of research, and that judgment depends greatly on how researchers share the methods and decisions behind their work. In this article, I advocate for the broader use of two promising approaches for increasing transparency and replicability in educational psychology: preregistration and registered reports. These two open science approaches, which are focused on the research process, complement other aspects of open

PREREGISTRATION AND REGISTERED REPORTS

education science, including open data and open code and methods (Cook et al., 2018; van der Zee & Reich, 2018).

Preregistration is the process of documenting the hypotheses and analytic plan for a study and then creating a time-stamped record of that document, usually through submission to a study registry (Nosek et al., 2018). By recording hypotheses and methods before conducting any analysis of data, researchers and readers can clearly delineate between hypotheses and decisions that were made a priori to data collection and those that were made post hoc, after viewing data. Preregistrations can be written in different forms and with different levels of granularity. The study registry AsPredicted.org uses a short, nine-question template that can be answered in less than a page (<https://aspredicted.org/nfj4s.pdf>). The Open Science Foundation preregistration template includes 25 main questions, which each have one to five sub-questions, in a ten page template. My own preregistrations have ranged from a short narrative (<https://osf.io/6rp5w/>) for a study on planning prompts to improve persistence in MOOCs, to a complete package of predictions and plans that included the exact analytic code that we planned to run after data collection was complete in a study testing a variety of behavioral interventions in MOOCs (<https://osf.io/5kvqf/>). What all preregistrations share in common is a desire to clearly delineate between prediction and “postdiction” (Nosek et al., 2018), between analytic plans generated ex ante to data analysis and those created post hoc.

Registered reports are a new format of journal publication that closely integrate preregistration with peer review (Chambers et al., 2015). The peer review process is formalized into two phases. Authors submit a Phase I manuscript that looks very similar to the first half of a typical journal article, with sections for the introduction, background and context, and method. The method section delineates the author’s hypotheses or claims to be assessed and the pre-

PREREGISTRATION AND REGISTERED REPORTS

specified analytic plans for testing those hypotheses or assessing those claims. This Phase I manuscript is then sent out for peer review, and reviewers evaluate the importance of the research questions and the appropriateness of the data and methods to address those questions. In sum, authors submit a research proposal and study design that goes through the peer review process; much like how dissertation proposals go through review and acceptance before a secondary review at the final submission stage. In contrast to a traditional peer review process, reviewers can provide suggestions early enough to meaningfully influence the structure and design of the study.

If reviewers and editors find merit in the Phase I manuscript, the article is then granted “in-principle” acceptance. This means that as long as authors conduct the study as agreed, while making reasonable and well-documented modifications, the study will be accepted without prejudice as to the direction or magnitude of the findings. Articles can still be rejected at this later stage for a wide variety of reasons, including implementation flaws or other serious research problems, but whether the results are null, negative, positive, or surprising—as long as the research is well-conducted and well-documented, the editors and journal agree to publish the findings. With Hunter Gehlbach and Caspar Albers, I recently (2020) edited a special topic for AERA Open that published some of the first registered reports in educational psychology, including findings on student-teacher relationships (Robinson et al., 2019), pre-service teacher perceptions of researcher trustworthiness (Merk & Rosman, 2019), text framings to improve college matriculation (Kramer, 2020), and identification policies for gifted education assignment (Peters et al., 2019). Registered reports offer a promising new format and process to address some of the current misaligned incentives in scholarly publishing.

Addressing Problems in Scholarly Publishing with Open Science

PREREGISTRATION AND REGISTERED REPORTS

In the introduction to this special issue, Gehlbach and Robinson describe in detail a variety of negative consequences that can follow from insufficient transparency in academic publishing. To summarize: when scholarly journals do not demand transparency about the timing of research design decisions, authors have the freedom to elide the distinctions between ex ante and post hoc research design decisions; authors are more likely to achieve “publishable” findings through post hoc data analytic decisions or questionable research practices (Makel et al, 2019); prestigious journals prioritize “significant findings”; null results are hidden in file drawers; the literature becomes over populated with illusory results (Gehlbach & Robinson, 2018); and the field finds itself on a sandy, unsteady foundation. In psychology, medicine, and other fields, preregistration and registered reports have begun to show evidence that they have the potential to address some of these serious problems (Allen & Mehler, 2019; Dechartres et al., 2016; Hardwicke & Ionnandis, 2018).

Educational psychology should adopt preregistration and registered reports through both top-down and bottom-up approaches. From the field’s leadership, journal editors and conference organizers should encourage and incentivize preregistration and create new submission pathways for registered reports. Meanwhile, individual researchers should begin adopting these practices in their own research programs to increase the transparency of their publications while modeling and encouraging wider adoption of these practices. Researchers conducting quantitative, experimental, laboratory based studies will find many examples from psychology and neuroscience to inspire these shifts in practices. There is less guidance available to qualitative researchers (Haven & van Grootel, 2019), design researchers, and those who must negotiate access to schools and other field sites. But open science practices exist on a spectrum rather than as a binary, and researchers need not adopt whole cloth the approaches of experimentalists to

PREREGISTRATION AND REGISTERED REPORTS

find specific ways of increasing transparency in research design that work in particular subfields and methodological approaches. Moreover, educational psychologists, with a long tradition of methodological diversity and interdisciplinarity, might be particularly well positioned to help develop new norms and best practices as open science approaches continue to evolve (Gehlbach and Robinson (2021/this issue).

Objections and Responses to the Expansion of Preregistration and Registered Reports

There are at least three common objections to this course of action: a) we don't know if these new methods will actually address the field's problems, b) the amount of time required to make research more transparent is too costly, and c) these approaches to greater transparency emerged primarily among scientists conducting randomized trials in laboratory studies, and they will not translate well to field studies, qualitative research, design research, or other diverse settings (Chambers et al., 2014; van 't Veer & Giner-Sorolla, 2016). These are legitimate concerns, and it is fair to say that I am calling for the broader use of open science practices without definitive evidence that they solve the problems we face and without perfected guidance for how these open science methods can be implemented across the methodological diversity of educational research.

Throughout this article, I will try to address these concerns in two ways. One through line of my arguments is technical. By trial and iteration, we can adapt these new publishing mechanisms to the various settings of educational psychology. We can be judicious in using these approaches, for instance by expecting alignment with fewer open science practices in new, exploratory lines of research and expecting greater levels of transparency in research designed to make a stronger policy case. These concerns appear to be surmountable; similar objections were raised when pharmaceutical researchers and experimentalists began adopting these practices, and

PREREGISTRATION AND REGISTERED REPORTS

yet new open science practices seem to be working to address problems of the scientific literature in medicine, psychology, and elsewhere (Allen & Mehler, 2019; Dechartres et al., 2016; Hardwicke & Ionnandis, 2018). Recent metascience results concerning registered reports are promising: in a comparison of registered reports and non-registered reports published in psychology and neuroscience, a group of 353 peer reviewers found that registered reports outperformed typical papers on all nineteen criteria evaluated, such as methodological quality and overall paper quality (Soderberg et al, 2020). Common sense suggests that too much emphasis on open science practices will lead to diminishing returns, but these early findings suggest that fields that were early adopters of open science have yet to reach these limits.

My second appeal is that early reports from educational researchers who adopt these practices suggest that authors find them useful and enjoyable (Reich et al., 2020). When researchers create a preregistration, they are glad to have a more complete record of thinking and designs. An emphasis on early writing means that when data analysis is complete, a preregistration or Phase I manuscript means that a final published is already one-third finished. Adopters of open science practices enjoy being able to share methods and approaches more completely with researchers and colleagues. Writers of registered reports appreciate getting peer feedback on research designs before data collection and analysis choices are locked in—when they can still be changed. Registered report authors report that they find that reviewers are more constructive and more curious about research design decisions when they, too, do not know what the findings will be (Reich et al., 2020). Of course, these statements will not be true for every researcher in every study context, but the growth of open science practices is not solely fueled by utilitarian considerations; they grow because individuals and institutions believe in transparency

PREREGISTRATION AND REGISTERED REPORTS

as a core value of science, and preregistrations and registered reports are ways of acting on those deeply held beliefs.

In the sections that follow, I advocate in greater depth for the wider use of these two practices—preregistration and registered reports. To emphasize the importance of preregistration, I examine one of the core concerns with contemporary statistical research in the social sciences: insufficient consideration of the importance of ex ante prediction. I then describe a case study from my own research, where a regimen of preregistration helped avoid the enthusiastic reporting of spurious findings from a large study in online learning. Next, I provide suggested guidelines for how individual researchers and institutional leaders (such as conference organizers and journal editors) can contribute to the growth and development of preregistration in educational psychology. I then turn to registered reports and describe how they offer an institutional structure for incentivizing preregistration, and I share my own experience editing the first special topic for *AERA Open*. I conclude with final thoughts about how individuals can participate in the movement for open science in educational research and how future research can better inform these approaches.

The Importance of Transparency in Confirmatory Claims: Prediction, Postdiction, and Preregistration

In the introduction to this special issue, Gehlbach and Robinson provide a useful overview of some of the shortcomings of the scholarly process that open science seeks to address. To underscore the importance of preregistration, I focus in depth on one particular place in the research process where greater transparency can improve the critical consumption of educational research: the timing of data analytic decisions. In the preregistration process, researchers create a time-stamped record of their hypotheses, positionality, and data analytic

PREREGISTRATION AND REGISTERED REPORTS

plans. None of this strictly binds future researcher decisions, but it can play a crucial role in helping readers understand the important issue of distinguishing prediction—hypothesis generation ex ante to data collection—from postdiction—hypothesis generation that occurs after data collection.

Many studies in educational psychology use statistical toolkits built on null-hypothesis significance testing, “a method of statistical inference by which an experimental factor is tested against a hypothesis of no effect or no relationship based on a given observation” (Pernet, 2017, p. 3), and these statistical methods require researchers to make a theoretically informed prediction (Nosek et al., 2018). Researchers must make predictions in order to define hypotheses and null hypotheses to test, and readers must be able to discern which researcher statements were predictions and which were post hoc interpretations. Nosek et al. (2018) adopted the term *postdiction* to describe hypotheses that are defined after beginning the process of analyzing data. It is a clumsy neologism, and the awkwardness of its sound as a word is a good match for its problematic use in quantitative research.

The importance of distinguishing prediction from postdiction has been a subject of scholarly attention at least since de Groot wrote about the topic in 1956 (1956/2014). At the heart of frequentist statistics, which still dominates the quantitative education sciences, is the concept of long-term error control. While false positives will unavoidably be reported in individual studies, the frequency of this type of error is controlled and will, in the long run, not exceed the alpha value—commonly set at 5%. Relative frequencies depend on a denominator: the total quantity of tests that have been (or even could have been) performed. If the hypotheses and analyses are not predesignated, this denominator becomes unspecified and undefinable.

PREREGISTRATION AND REGISTERED REPORTS

Effectively, it makes null-hypothesis tests lose their informative value and decisive nature. As de Groot wrote almost 70 years ago:

If the processing of empirically obtained material has in any way an “exploratory character,” i.e. if the attempts to let the material speak leads to ad hoc decision in terms of processing . . . then this precludes the exact interpretability of possible outcomes of statistical tests. . . . One “is allowed” to apply statistical tests in exploratory research, just as long as one realizes that they do not have evidential impact. (de Groot, 1956/2014, p. 193)

Tukey made this point forcefully again in 1980, arguing that confirmatory research required two elements: “1) RANDOMIZE! RANDOMIZE! RANDOMIZE! 2) Preplan THE main analysis” (Tukey, 1980, p. 24; the emphasis is Tukey’s own). The first point has been widely adopted in educational psychology and related fields; the second much less so.

A 2003 report from the Institute of Education Sciences describing key elements of well-designed causal studies puts a great deal of emphasis on properly implemented randomization and makes no mention of preplanning. As important as randomization (or some other approach to causal assignment) is, the rigor of a confirmatory study also depends on researchers’ ensuring that analytic decisions are *not dependent* on the data (McBee & Field, 2017).

Gelman and Loken (2013) developed the metaphor of the garden of forked paths to illustrate this same problem. Whenever choices are made based on the data instead of being predesignated, there are so many possible ways to analyze the data that at least one trail in this garden of forking paths will likely lead to a statistically significant result. While this problem holds for studies of any size, it becomes more problematic with an increasing number of variables and/or samples (van der Sluis et al., 2017). Interpretable null-hypothesis testing

PREREGISTRATION AND REGISTERED REPORTS

depends on preregistration of hypotheses and all other decisions that affect the kind and number of statistical tests that might be run and/or reported.

In his 1956 paper, de Groot used an invented illustration from parapsychology to illustrate the potential problems of ignoring this advice. A mentalist claims to be able to concentrate on the roll of a six-sided die and make one side come up more frequently. Unfortunately, the mentalist's concentration wavers, so the power does not work consistently. A researcher studying these powers throws a die 300 hundred times and finds that by clustering throws into six groups of 50, there are certain series where six indeed comes up significantly more often than chance would predict. By examining these data, the researcher has "discovered" the series during which the mentalist's concentration was at its peak! Of course, the example does not prove psychokinesis but rather the incredible power of post hoc data analysis and storytelling for composing analytic designs that conform to publishing expectations. As de Groot observed:

After some practice [the research] will also not find it difficult to show for any die that it (or the person who throws it) behaves "significantly" exceptionally "somewhere". This claim can always be maintained, the "proof" can always be provided, as long as one does not need to specify in advance, where exactly "somewhere" is located. (1956/2014, p. 192)

Preregistration helps assure that critical readers of educational research can determine what predictions have been specified in advance and therefore have an evidentiary weight that post hoc explanations lack.

Preregistration: A Case Study

PREREGISTRATION AND REGISTERED REPORTS

My own recent version of the mentalist's experiment came in the midst of conducting one of the largest experiments to date in online learning and higher education (Kizilcec et al., 2020). Colleagues and I implemented a series of experimental learning-support interventions in nearly every massive open online course published by Harvard, MIT, and Stanford over the course of three years, with over 250,000 learners randomly assigned to a treatment in over 250 courses.

Our interventions—including a planning prompt (Yeomans & Reich, 2017), a value relevance affirmation (Kizilcec et al., 2017), a mental contrasting with implementation intentions (Kizilcec & Cohen, 2017), and a social accountability intervention—had shown moderate to large positive effects on persistence and completion in a set of prior studies, each of which included thousands of participants. From these prior studies, we were quite optimistic that our low-cost interventions would prove effective at even greater scale.

When we began our study, we wrote a detailed preregistration of our research plans to increase transparency and strengthen confirmatory evidence for our findings. Since data collection is well defined and constrained in online learning, and since we were at a late stage in a well-developed research agenda, we were able to compose a very thorough preregistration, which is posted at <https://osf.io/hjesw>. We wrote a narrative description of our hypotheses and analytic techniques, and we also included a copy of the exact analytic code that we planned to run.

We initially gathered data from only a single semester's worth of courses, and when we conducted our first analysis, we were surprised to find that our interventions had no significant effect on our targeted population. However, we also observed that not all our analyses were proceeding as expected. For instance, in our initial studies, we used a set of exclusion criteria that removed a reasonable number of participants who we felt were unlikely to benefit from the

PREREGISTRATION AND REGISTERED REPORTS

intervention. However, in our new context, these exclusion criteria eliminated more than half the study sample. When we revised the exclusion criteria to be more inclusive, we found that some of our interventions had a statistically significant result.

In years gone by, we might have chosen to simply revise our exclusion criteria post hoc and then attempt to publish our results without calling attention to the timing of our change in exclusion criteria. Under the current norms of publishing methods, no one would have been the wiser. Instead, we were fortunate to have had an experimental set up (a large series of online courses running automatically with no additional effort from us) that allowed for ongoing data collection. We also, fortunately, had a commitment to open science practices. So we decided that rather than publish our findings using our post hoc data analytic plans, we would write a new preregistration (<https://osf.io/ya594>) and then see if our results replicated in a second wave of data collection. We waited another semester and summer, reopened the data, and reanalyzed our results with refined predictions and analytic code.

Once again, our new predictions led to null results. The significant effects that we found with our modified exclusion criteria did not replicate in the new wave of the study. We continued to explore our data and to find relevant new covariates, and we developed new short-term outcomes. We ran two more waves of the study, and by the end of our experiment, we found two modest effects that we could confidently predict using preregistered study plans and subsequently find in later waves of the experiment. But throughout the four waves of the study, we regularly found statistically significant associations with certain outcomes, interventions, covariate structures, and exclusion criteria that appeared during our exploratory research in a single wave but could not be predicted and replicated in subsequent waves. In the first wave (N=96,548), for example, we found evidence that our value relevance affirmation benefited

PREREGISTRATION AND REGISTERED REPORTS

racial minorities in the United States, but we never confirmed this finding in the subsequent three waves (with N of 102,969, 41,933, and 27,719 respectively (<https://osf.io/ya594>). It was a humbling experience. And it was a powerful illustration of de Groot's (1956/2014) mentalist trap and Gelman and Loken's (2013) garden of forked paths.

For our research team, preregistration was an essential check on our inclination to tell ourselves stories about why theory might support our findings and on our incentives to write publishable articles and advance our careers as junior scholars. Preregistration also played a vital role in helping us publish confirmatory findings that we felt we could stand behind. Perhaps most importantly, it generated a record of our thinking and decisions that readers can use for themselves to ascertain how and when our data analytic strategies were shaped by new data.

Suggested Guidelines for Writing Preregistrations

In my niche of educational psychology—testing behavioral interventions in online course platforms—it was possible to compose thorough pre-analysis plans, right down to the analytic code, and these served our team well. This level of thoroughness might be inappropriate for less mature lines of research, where many factors in an experimental setting remain unknown, or in certain field settings where flexibility and adaptivity to local circumstances are important for negotiating access. Educational psychology and its various subfields will need to develop new norms and expectations around preregistration, weighing costs against the important benefits of transparency around prediction and preserving the confirmatory interpretability. The primary cost is researcher time: more writing must be done earlier in the research process (Reich et al., 2020). In experiments that are completed and published, researchers may experience this as a time shift rather than an additional time cost, since the writing done for preregistration largely overlaps with requirements for publication. For researchers who conduct many experiments,

PREREGISTRATION AND REGISTERED REPORTS

discarding the results of most and only publishing a select few, the burdens of these costs will be felt most acutely. The field may also find, however, that discouraging unpublished experiments and making more room for the publication of null results and other early field trials brings benefits of transparency to the whole field.

The practices of open education science are emerging (van der Zee & Reich, 2018), and it is difficult to make evidence-based claims about the best methods or principles for preregistering studies. In 1969, de Groot called for thorough documentation:

Foremost . . . is the recommendation to work out in advance the investigative procedure (or experimental design) on paper to the fullest possible extent. This “blueprint” should comprise: a brief exposition of the theory; a formulation of the hypothesis to be tested; a precise statement of the deductions leading up to the predictions to be verified; a description of the instruments—in the broadest sense—to be used, complete with instructions for their manipulation; detailed operational definitions of the variables to be used; a statement about the measurement scales (nominal, ordinal, interval, ratio) in which the respective variables are to be read . . . ; a clearly defined statement of the respective universes to which the hypothesis and the concrete prediction(s) apply; an exact description of the manner in which the samples are to be drawn or composed; a statement of the confirmation criteria, including formulation of null hypotheses, if any, choice of statistical test(s), significance level and resulting confirmation intervals . . . ; for each of the details mentioned, a brief note on their rationale i.e., a justification of the investigator’s particular choices. (de Groot, 1969, p. 136)

This kind of detailed preplanning might appear especially daunting to educational psychologists or learning scientists working in field and school settings with messy dynamics

PREREGISTRATION AND REGISTERED REPORTS

around negotiating access and working in the complex environments of typical schools and classrooms. Nobel laureate Esther Duflo and colleagues (2020) have recently argued for a more concise approach to pre-analysis plans, especially in the kinds of field settings that are held in common with educational psychology. They argue that in real-world settings, the time costs of documenting all possible contingencies rapidly become prohibitive. (They reserved their strongest cautions against editorial norms that privilege studies that follow their plans with complete fidelity over those that make sensible adaptations as field studies evolve.) Duflo et al. argued for more detail where study designs call for it; for instance, if subgroup effects are expected, predicted, and an important part of the theory and analytic plans, then more detail should be provided in a preregistration about analytic plans for treating subgroups. But overall, they recommended that researchers keep pre-analysis plans short. Examples of Duflo's own preregistrations can be found online, including a field study in 200 preschools and early-education centers in India (<https://www.socialscienceregistry.org/trials/1195>) and her own research in an EdX MOOC (<https://www.socialscienceregistry.org/trials/172>).

Ultimately, open science advocates in educational psychology will have to develop their own norms around preregistration for the field. I recommend three principles as a starting point:

- 1) Start with short, concise preregistrations, primarily because if the burden is lower, more people will be willing to write them (and to read them afterward).
- 2) Develop more robust preregistrations over the course of a research trajectory. The first confirmatory study in a series might have a very concise preregistration, but as researchers become more familiar with settings, experimental designs, data outputs, and analysis plans, they should be able to be more concrete and comprehensive in their plans.

PREREGISTRATION AND REGISTERED REPORTS

- 3) Preregister more detail on correspondingly important parts of the study. A study that hinges on a certain subgroup analysis, for instance, should have clear hypotheses and methods for analyzing that subgroup.

Individual researchers can use these principles to guide their own preregistrations, and journal editors and conference organizers can publish guidance on the kinds of preregistrations that would be most appropriate to their subfield. Johnson and Cook, for example, published recommended guidelines for preregistration of single-case guidelines in the journal *Exceptional Children* (2019). Researchers in special-education readers using single-case methods will find the examples and principles there more helpful, specific, and relevant than the general guidelines that I have offered here. And just as single-case study methods require specific guidance for preregistration, so will qualitative research (Haven & van Grootel, 2019), design research, and other methods. As more educational psychologists adopt these practices, future metascience studies can analyze how preregistrations are constructed, used by researchers, incorporated by reviewers and editors, and read by other scientists.

Preregistrations do not and should not constrain the kinds of analysis that researchers ultimately conduct. From de Groot to Duflo and colleagues, there is a consensus among open science advocates that once a study is preregistered, researchers should be free to make sensible adaptations, to try new constructions of variables, to test different statistical analytic methods, and so forth. The crucial point to hold firmly is that the publication of results should allow readers to identify data analytic strategies that were developed before and after the start of data analysis. One publication strategy is to have two distinct findings sections, one for prespecified analyses and one for post hoc, exploratory work (Gehlbach and Robinson, 2018). I recommend

PREREGISTRATION AND REGISTERED REPORTS

that the American Psychology Association incorporate this guideline in the next update to the *Publication Manual*.

One way to reinforce all of these research design norms is to incorporate the process into the peer review process, which leads me to registered reports.

Registered Reports

As more individual researchers preregister their studies, readers of educational psychology research will be able to more clearly distinguish between confirmatory and exploratory findings. Some problems in scholarly publishing, however, cannot be solved by the actions of individual researchers; industry-wide editor and reviewer bias against null findings is a good example (Ioannidis, 2005). The positive effects of an intervention or surprisingly large effect sizes are considered prestigious and publishable, so even if the literature were not overrepresented with illusory results, the “file-drawer” problem would still lead to an overrepresentation of positive or negative findings and an underrepresentation of null findings. One promising solution to addressing these disincentives is shifting peer review in a way that would privilege certain kinds of analytic rigor over certain kinds of results through a new publication format: registered reports.

Registered reports were developed by the journal *Cortex* for use in the publishing of neuroscience articles (Chambers et al., 2015). Like many studies in educational psychology, neuroscience involves a great deal of quantitative research primarily conducted using statistical analyses that build on null-hypothesis significance testing. In registered reports, peer review happens at least before data analysis, and ideally before data collection begins. Authors submit a preregistered Phase I manuscript with an introduction, background, and method section—including an analysis plan, and editors send these out for peer review. Reviewers, authors, and

PREREGISTRATION AND REGISTERED REPORTS

editors collaborate to refine the quality of research questions and analytic plans for a study without knowing the results of the analysis. Positive reviews lead to in-principle acceptance, where editors agree to publish articles, regardless of the direction or magnitude of findings, if the original plans are adhered to, with reasonable modifications documented.

Registered Reports: Case Studies in Education and Educational Psychology Research

From 2018 to 2020, I edited a special topic at *AERA Open* about registered reports. I entered the experience with the sense that registered reports were a promising approach to publishing, and I concluded the endeavor believing that it would lead to a more enjoyable, more transparent, more productive future for educational research. When we asked our authors to reflect on the process of publishing registered reports, one author team argued that there was no going back: “It was like upgrading from a typewriter to a computer” (Reich et al., 2020).

The most distinctive shift in the review process was the feeling that the work was collaborative rather than adversarial and was focused on building the best possible research rather than serving as a gatekeeping function for particular kinds of theories or findings. When reviewers do not know the results of the study, the review process includes a greater sense of humility. Reviewers offered constructive suggestions alongside critiques. Reviewers could also propose substantial modifications to analytic plans without fear of asking researchers to rewrite an entire finished paper. In a registered report, the paper is not yet written, the analyses are not yet conducted, and much more room exists for making substantial changes and improvements in research design. One team argued that registered reports were more efficient than regular publications: “We saved so much time on analyses by not having to re-do ten different analyses to address reviewer criticisms” (Reich et al., 2020).

PREREGISTRATION AND REGISTERED REPORTS

A second important benefit of the registered report format is that authors claimed that the format let them be more candid with their audience about their findings. Knowing that the paper had already been evaluated on the merits of the design, authors felt more freedom to candidly describe limitations or uncertainties in their research. In most editorial processes, authors feel pressure to have a certain tidiness in their findings and theorizing. In registered reports, the questions and methods must be tidy, but if the findings are confusing, contradictory, or problematic, authors can candidly report those issues rather than sanding down their rough edges. After receiving an in-principle acceptance, researchers can report their findings and be straightforward about null effects, possible confounds, and additional possible interpretations.

Finally, making research design public and available for commentary earlier in the research process means that researchers have an additional incentive to do a better job in those important phases. Of course, many labs have internal processes—from advisor meetings to lab meetings to peer collaborations—that create opportunities for accountability, but public accountability means that researchers have an additional incentive to sharpen their designs.

Suggested Guidelines for Registered Reports

As with preregistration, there are many details to be ironed out and norms to be developed for registered reports to be a valued component of publishing research in educational psychology. Probably the most important consideration involves field studies. While lab studies can often be carefully designed in advance and rigorously controlled, field studies often require negotiation and modification of study plans with local stakeholders. Doctoral student Jenna Kramer (2020) published a registered report in our special topic based on an intervention that she had conducted to send text messages to matriculating college freshman to reduce “summermelt,” the phenomenon of accepted high school seniors’ not registering for and attending college in the

PREREGISTRATION AND REGISTERED REPORTS

subsequent fall. Her research timeline was dictated by her field implementation partner and her doctoral studies calendar, and she needed to launch her field study before getting editorial approval for a Phase I manuscript. Still, she wanted to publish her study as a registered report, so after completing data collection, she waited to conduct data analysis until receiving feedback on her Phase I submission. One risk of this approach is that reviewers might have demanded substantial changes to her experimental methods (as they might in a regular journal submission), and Kramer would have been stuck waiting for reviews. As it turned out, two experts in the field provided external reviews on her research design and made constructive suggestions that she could implement, her Phase I manuscript was granted in-principle acceptance, and the entire review and publication calendar lined up reasonably well with her doctoral studies. Still, one could imagine other scenarios in which writing and implementing a rigid preregistration are incompatible with the kinds of negotiations required for many field studies.

Another issue concerns publication norms and the reporting of both preregistered and exploratory findings. Right now, in many research publications, the findings sections may include results that were preplanned alongside exploratory work, and only weak norms exist to clearly delineate the two. In our special issue, we encouraged authors to clearly distinguish between preregistered work and exploratory work. Robinson and colleagues (2019), for instance, published results from an effort to replicate in a higher education setting an intervention for improving teacher-student relationships that had positive effects in a K–12 pilot. The experimental intervention had null results, but in conducting their surveys, the researchers found some of the first correlational evidence connecting teacher-student relationships and student outcomes in higher education. In a traditional publication, the authors might have felt pressure to minimize the broader experimental context and publish only the “significant” survey findings. In

PREREGISTRATION AND REGISTERED REPORTS

a registered report, the field benefits from the publication of a well-designed experimental study with null results, and the exploratory survey findings are presented in a fuller context of the larger initiative

The editors at journals for educational psychology and its subfields should launch multiyear experiments with publishing registered reports, following up on existing efforts from *AERA Open* and the support for preregistration in *Exceptional Children*. If the experience in this field is similar to what we found in publishing our special topic, the initiative will produce high-quality, transparent research in a process that is rigorous but also supportive and collaborative. We were very pleased that three of the seven articles that we published in our special topic had graduate students as lead authors, and I believe that junior scholars will have an important role to play in moving the field forward.

One irony of open science advocacy for preregistration and registered reports is that it will be difficult to prove conclusively that these practices work to improve transparency and reduce illusory results. We cannot randomly assign 50% of the field to adopt these practices for two years and see how effectively they change the literature and scholarship of educational psychology. But early evidence from efforts like the *AERA Open* special topic are promising; there is a cadre of early career scientists excited about these new approaches, which hold substantial promise for improving research in the field. There is much work to be done to identify the most effective norms and guidelines for these practices, but the first steps are to have the field invest in systematic efforts to test implementations of these practices. Some of the most important publishing experiments are likely to be conducted in subfields that primarily use qualitative and design methods, conduct research in complex field settings, or analyze large secondary datasets. The initial open science norms that have begun to coalesce in quantitative

PREREGISTRATION AND REGISTERED REPORTS

and experimental research may prove to be quite valuable with minimal adaptation, they may require substantial adaptation, or it may be that increased transparency is best achieved through other means.

While I hope that leaders in educational psychology will heed this call, to my fellow researchers I say that there is no need to wait for leaders or institutions to adopt these practices—you can start right now. For your next study, publish a preregistration at the Open Science Foundation or with the Society for Research on Educational Effectiveness. Don't worry about doing it "perfectly"; choose a format and granularity of preregistration that fits your research agenda, and see what you learn about the experience that could benefit others in your subfield. If preregistering an entire study seems daunting, remember that open science practices exist on a spectrum not a binary: one of several hypotheses could be preregistered, study three of a four-study paper could be preregistered, all measures and outcomes could be listed in a preregistration even if hypotheses are presented as exploratory; a research team's first Phase I registered report manuscript submission might seem less daunting after trying a few incremental steps.

Along with efforts in individual research groups, talk with others in your subfield about trying these experiments. Host a symposium or a discussion about preregistration at your next conference or faculty meeting. If you have tried preregistration, discuss your experience. If not, have participants read some of the studies in our special topic or other preregistrations or registered reports and discuss how these practices could be implemented in your subfield. Ask the editors of the journals in which you publish if they would be willing to publish a special issue on registered reports, and volunteer to help as editor. You do not need to be an expert; I myself have yet to publish a registered report. Change in the field will not be led by experts in applying open science methods to educational psychology, because there are not any, but by well-

PREREGISTRATION AND REGISTERED REPORTS

intentioned colleagues who raise their hand to try to make change and improve scientific processes for the field and for the public. Mellor (2021/this issue) offers additional ideas for how changing research norms in research culture can help move educational psychology towards more open science practices.

Conclusion

Across psychology and the social sciences, there is a growing acknowledgment of serious problems with the incentives in our field and how those incentives shape our conduct as researchers. Our understanding of these problems remains imperfect and incomplete. One anonymous reviewer of this paper suggested that open science advocates should audit recent research in educational psychology to identify the proportion of published studies that might be improved with open science methods. This would indeed be a worthy future study. The same anonymous reviewer suggested publishing a wider variety of case studies of preregistrations and registered reports, and I agree that more of these kinds of personal research narratives would help diverse researchers see themselves in new open science practices.

But we need not have a perfect understanding of the problems in research publishing to experiment with new approaches to addressing these issues. Preregistration and registered reports are two of the most promising tools for addressing the root causes of many of our scholarly ills, and they have great potential to lead to a more transparent, more accurate scholarship of educational psychology and a field that offers greater knowledge and wisdom to the teachers and learners whom we serve.

References

PREREGISTRATION AND REGISTERED REPORTS

Allen C., & Mehler D. M. A. (2019). Open science challenges, benefits and tips in early career and beyond. *PLOS Biology*, *17*(5), e3000246.

<https://doi.org/10.1371/journal.pbio.3000246>

Chambers, C. D., Dienes, Z., McIntosh, R. D., Rotshtein, P., & Willmes, K. (2015). Registered reports: Realigning incentives in scientific publishing. *Cortex*, *66*, A1–A2.

Chambers, C. D., Feredoes, E., Muthukumaraswamy, S. D., & Etchells, P. (2014). Instead of “playing the game” it is time to change the rules: Registered reports at AIMS Neuroscience and beyond. *AIMS Neuroscience*, *1*(1), 4–17.

Cook, B. G., Lloyd, J. W., Mellor, D., Nosek, B. A., & Therrien, W. J. (2018). Promoting open science to increase the trustworthiness of evidence in special education. *Exceptional Children*, *85*(1), 104-118.

de Groot, A. D. (1969). *Methodology: Foundations of inference and research in the behavioral sciences*. Mouton.

de Groot, A. D. (2014). The meaning of “significance” for different types of research (E.-J. Wagenmakers, D. Borsboom, J. Verhagen, R. Kievit, M. Bakker, A. Cramer, D. Matzke, D. Mellenbergh, and H. L. J. van der Maas, Trans.). *Acta Psychologica*, *148*, 188–94.
(Original work published 1956)

Dechartres, A., Ravaud, P., Atal, I., Riveros C., & Boutron, I. (2016). Association between trial registration and treatment effect estimates: A meta-epidemiological study. *BMC Medicine*, *14*(100). <https://doi.org/10.1186/s12916-016-0639-x>

Duflo, E., Banerjee, A., Finkelstein, A., Katz, L. F., Olken, B. A., & Sautmann, A. (2020). In praise of moderation: Suggestions for the scope and use of pre-analysis plans for RCTs in

PREREGISTRATION AND REGISTERED REPORTS

- economics (working paper 26993). National Bureau of Economic Research.
<https://www.nber.org/papers/w26993>
- Gehlbach, H., & Robinson, C. D. (2018). Mitigating illusory results through preregistration in education. *Journal of Research on Educational Effectiveness, 11*(2), 296–315.
- Gehlbach, H., & Robinson, C. D. (2021/this issue). From old school to open science: The implications of new research norms for educational psychology and beyond. *Educational Psychologist, 56*(2), ###-###).
- Gelman, A., & Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. Department of Statistics, Columbia University.
<http://www.stat.columbia.edu/~gelman/research/unpublished/forking.pdf>
- Hardwicke, T.E., Ioannidis, J. P. A. (2018). Mapping the universe of registered reports. *Nature Human Behavior 2*, 793–96. <https://doi.org/10.1038/s41562-018-0444-y>
- Haven, T., & Van Grootel, D. L. (2019). Preregistering qualitative research. *Accountability in Research, 26*(3), 229–44.
- Identifying and implementing educational practices supported by rigorous evidence: a user friendly guide.* (2003). Institute of Education Sciences.
<https://www2.ed.gov/rschstat/research/pubs/rigorousavid/rigorousavid.pdf>
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLOS Medicine, 2*(8), e124.
- Johnson, A. H., & Cook, B. G. (2019). Preregistration in single-case design research. *Exceptional Children, 86*(1), 95–112.

PREREGISTRATION AND REGISTERED REPORTS

- Kizilcec, R. F., & Cohen, G. L. (2017). Eight-minute self-regulation intervention raises educational attainment at scale in individualist but not collectivist cultures. *Proceedings of the National Academy of Sciences, 114*(17), 4348–53.
- Kizilcec, R. F., Reich, J., Yeomans, M., Dann, C., Brunskill, E., Lopez, G., . . . Tingley, D. (2020). Scaling up behavioral science interventions in online education. *Proceedings of the National Academy of Sciences, 117*(26), 14900–05.
- Kizilcec, R. F., Saltarelli, A. J., Reich, J., & Cohen, G. L. (2017). Closing global achievement gaps in MOOCs. *Science, 355*(6322), 251–52.
- Kramer, J. W. (2020). Experimental evidence on the effects (or lack thereof) of informational framing during the college transition. *AERA Open, 6*(1).
<https://doi.org/10.1177/2332858420908536>
- Makel, M. C., Hodges, J., Cook, B. G., & Plucker, J. (2019, October 31). Questionable and Open Research Practices in Education Research. <https://doi.org/10.35542/osf.io/f7srb>
- McBee, M. T., & Field, S. H. (2017). Confirmatory study design, data analysis, and results that matter. In M. C. Makel & J. A. Plucker (Eds.), *Toward a more perfect psychology: Improving trust, accuracy, and transparency in research* (pp. 59–78). American Psychological Association. <https://doi.org/10.1037/0000033-004>
- Merk, S., & Rosman, T. (2019). Smart but evil? Student-teachers' perception of educational researchers' epistemic trustworthiness. *AERA Open, 5*(3).
<https://doi.org/10.1177/2332858419868158>
- Mellor, D. T. (2021/this issue). Changing norms in research culture to value transparency over novelty. *Educational Psychologist, 56*(2), ###-###.

PREREGISTRATION AND REGISTERED REPORTS

- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–06.
- Pernet, C. (2017). Null hypothesis significance testing: A guide to commonly misunderstood concepts and recommendations for good practice. *F1000 Research*, *4*(621).
<https://doi.org/10.12688/f1000research.6963.5>
- Peters, S. J., Rambo-Hernandez, K., Makel, M. C., Matthews, M. S., Plucker, J. A. (2019). Effect of local norms on racial and ethnic representation in gifted education. *AERA Open*, *5*(2).
<https://doi.org/10.1177/2332858419848446>
- Reich, J., Gehlbach, H., & Albers, C. J. (2020). “Like upgrading from a typewriter to a computer”: Registered reports in education research. *AERA Open*, *6*(2), 2332858420917640.
- Robinson, C. D., Scott, W., Gottfried, M. A. (2019). Taking it to the next level: A field experiment to improve instructor-student relationships in college. *AERA Open*, *5*(1).
<https://doi.org/10.1177/2332858419839707>
- Soderberg, C. K., Errington, T. M., Schiavone, S. R., Bottesini, J. G., Singleton Thorn, F., Vazire, S., Esterling, K., Nosek, B. A. (2020, November 16). Initial Evidence of Research Quality of Registered Reports Compared to the Traditional Publishing Model.
<https://doi.org/10.31222/osf.io/7x9vy>
- Tukey, J. W. (1980). We need both exploratory and confirmatory. *The American Statistician*, *34*(1), 23–25.
- van der Sluis, F., van der Zee, T., & Ginn, J. (2017, April). Learning about learning at scale: Methodological challenges and recommendations. *Proceedings of the Fourth ACM Conference on Learning@ Scale*, 131–40.

PREREGISTRATION AND REGISTERED REPORTS

van der Zee, T., & Reich, J. (2018). Open education science. *AERA Open*, 4(3), 2332858418787466.

van 't Veer, A. E., & Giner-Sorolla, R. (2016). Pre-registration in social psychology—A discussion and suggested template. *Journal of Experimental Social Psychology*, 67, 2–12.

Yeomans, M., & Reich, J. (2017). Planning prompts increase and forecast course completion in massive open online courses. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 464–73.