

Open science and preregistration

Amanda Huensch, University of Pittsburgh, amanda.huensch@pitt.edu

Abstract

This chapter presents the practice of preregistration and its role in the open science movement. It begins with a definition and explanation of the different types of and approaches to preregistration along with the logic and justification of preregistering a study. Arguments both for and against preregistration are presented along with some common critiques and misconceptions of its purpose and implementation. Next, the chapter turns to the more practical considerations of how a researcher might participate in the preregistration process. This section highlights how preregistration impacts the research cycle and provides recommendations for those who wish to preregister their studies. Finally, the chapter ends with a discussion of future considerations, including a call for professional training and considerations of how to evaluate the successfulness of preregistration efforts.

Introduction

In line with a “methodological turn” (Byrnes, 2013, p. 82) resulting in an increased focus on study quality and rigor, the field of applied linguistics has joined psychology and many other scientific disciplines in an effort to engage in more transparent and open scientific practices (Gass et al., 2021). Open science initiatives have been gaining increased momentum as they offer solutions to combat the so-called replication or reproducibility crisis. Many complex and intertwined factors have contributed to this ‘crisis’ in the sciences. In psychology, Pashler and Wagenmakers (2012) pointed to (a) a publication (Bem, 2011) in a well-respected journal

providing empirical evidence of extrasensory perception; (b) reports of data fraud by prominent researchers (Stroebe et al., 2012); and (c) publications demonstrating how researcher degrees of freedom (i.e., the flexibility inherent in making methodological and analysis decisions during the research process) can produce problematic significant results, for instance that listening to the Beatle's song *When I'm Sixty-Four* will cause one to decrease in age by an average of 18 months (Simmons et al., 2011). In the field of applied linguistics, another impetus has been an increase in methodological syntheses (e.g., Norris & Ortega, 2000; Plonsky, 2013) calling for improvements based on their examinations of research and reporting practices (Marsden & Plonsky, 2018). The reproducibility crisis has also been more optimistically labeled the "credibility revolution" (Vazire, 2018, p. 411) because of its impact on fostering improved scientific practices. Take, for example from our own field, the establishment of the IRIS database (Marsden et al., 2016), a repository for materials and data sharing for second language research, the adoption of open science badges¹ by journals, and an increased focus on replication (Porte, 2012; Porte & McManus, 2019) including an amendment to the American Association for Applied Linguistics (AAAL) Tenure and Promotion Guidelines explicitly stating the value of replication studies.

The focus of this chapter is on preregistration or the process of placing a time-stamped research plan outlining hypotheses, methods, and analyses into a repository before any data are collected. This practice is designed to increase transparency as it makes it possible to differentiate preplanned analyses and methodological decisions from those that were not. Before turning to a discussion of the conceptual background, purpose, and benefits of preregistration, it

¹ Open science badges were created by the Center for Open Science as a way of acknowledging and rewarding researchers participating in the open science practices of materials sharing, data sharing, and preregistration. At the time of publication, 2703 journals were listed as offering one or more of the badges. <https://www.cos.io/initiatives/badges>

is worthwhile to acknowledge the relative ‘newness’ of this initiative in our field,² even in comparison to other open science initiatives (e.g., materials sharing). For instance, at the time this chapter was written, 15 journals in the Linguistics and Language discipline on topfactor.org were listed as offering open science badges, and while all of those journals offered the open materials and open data badges the same was not true for the preregistered badge.³ This means that we are still very much in the process of deciding upon norms and establishing best practices for preregistration. The goal of this chapter is twofold: to introduce readers to the conceptual underpinnings of preregistration highlighting its benefits and criticisms, and to provide some practical guidance and resources for study preregistration.

Conceptual Background, Purpose, and Benefits

In its most basic sense, preregistration is the practice of registering a specified research plan before data collection has occurred. The amount of detail included in the plan can vary from being quite skeletal to including a full write-up of the introduction, literature review, and methods section. While this type of document might seem unfamiliar at first, the basic concept has been likened to research proposals such as those for dissertations or theses created and defended at the graduate level (Roberson et al., 2020) or those for grant proposals. This is not to say that all preregistrations are as comprehensive as those documents might be, but importantly the preregistration should include enough detail to document critical decisions in advance of carrying out the study. Simmons et al. (2021b) provided a useful discussion of “good vs. bad” preregistration information (p. 13). To model this in an applied linguistics discipline, consider

² Preregistration, while an arguably newer initiative in applied linguistics, has been a mandatory component of clinical trials research in the United States since 2007.

³ Several journals in the field (e.g., *Bilingualism: Language, and Cognition*, *Journal of Child Language*, and *Language and Speech*) offer publication using a registered report format, a type of preregistration discussed in the Conceptual Background, Purpose, and Benefits section, despite not offering open science badges.

the examples in Table 1 that might be used in a study on L2 pronunciation. In both cases, the ‘bad’ responses are underspecified: There are multiple ways the variable comprehensibility could be defined and measured and many possible picture story tasks that could be used to elicit speech.

Table 1. Example information included in a preregistration

Preregistration item	Bad	Good
Dependent variable	comprehensibility	Listeners will rate speech samples on a scale of 1 = extremely easy to understand to 9 = extremely difficult to understand
Speaking task	speakers will retell a picture story	Speakers will have 30 seconds to plan their narrations of the Suitcase story before being given 90 seconds to retell the story in their own words

To appreciate the benefits of preregistration it is necessary to recognize the difference between empirical endeavors that generate hypotheses from those which tests hypotheses, or in other words between exploratory and confirmatory research. The danger in not doing so leads to placing too much confidence in the probability that results can be replicated (Nosek & Lindsay, 2018) because in treating exploratory work as confirmatory, “*p*-values lose their meaning due to an unknown inflation of the alpha-level” (Nosek & Lakens, 2014, p. 138). With a timestamped preregistration document and explicit distinctions made by authors identifying which methodological decisions and analyses were preplanned and which were not, preregistrations provide a public record for others to independently evaluate and distinguish confirmatory versus exploratory tests. It is important to note that preregistering a study does not cause it to be intrinsically of higher quality nor does it combat outright fraud (see Common Critiques and

Misconceptions section), but rather its main benefit is a potential increase in transparency (Wagenmakers & Vazire, 2020).

The need for increased transparency has come about in part from a combination of researcher degrees of freedom in conjunction with publication bias. Publication bias refers to the likelihood that studies with novel, clean, and statistically significant (dare we say ‘exciting’?) results are more likely to be published in comparison to those with null results or those that are ‘messy’ (e.g., perhaps some tests support predictions while others do not). Evidence for this bias has been provided in a variety of fields, including applied linguistics (see e.g., Chiu et al., 2012; Lee et al., 2015; Plonsky, 2013) and often comes in the form of demonstrating larger effects in published vs. unpublished literature (e.g., as part of a meta-analysis). A related factor is the pressures put on researchers to be prolific publishers in order gain tenure and promotion or secure research grants. This combination of a strong desire and need to publish frequently coupled with the higher likelihood of statistically significant results being published can lead to the widespread occurrence of questionable research practices (QRPs) whose impact means that it is possible, if not likely, that the cumulative results in our published literature do not accurately reflect reality. Preregistration might offer a solution to this state of affairs by increasing transparency of which methodological and analytical decisions were preplanned. For instance, in their analysis of effect sizes reported in published studies in the field of psychology, Schäfer and Schwarz (2019) demonstrated that the size of effects differed between published research with and without a preregistration component (median = 0.36 vs. median = 0.16, respectively) such that the former was inflated.

QRPs include *p*-hacking (Simonsohn et al., 2014) or attempting “multiple analyses to obtain statistical significance” (p. 534), HARKing (Hypothesizing After Results are Known;

Kerr, 1998) or “presenting a post hoc analysis (i.e., one based on or informed by one’s results) in one’s research report as if it were, in fact, an a priori” hypothesis (p. 196), and selective reporting of results (e.g., reporting only those analyses that obtained statistically significant results and omitting those that failed to reach statistical significance). Isbell et al. (2022) reported that almost half of the respondents to their survey on methodological practices in applied linguistics said they had withheld methodological details to make the study seem cleaner, and a similar number chose to omit results that were not statistically significant. This is not to imply that our field is full of immoral and fraudulent scientists, rather, that by our very human nature, we inadvertently participate in practices that threaten the trustworthiness and credibility of reported outcomes. Throughout the process of collecting and analyzing data, we make many decisions whose varied options are often justifiable. For instance, imagine that we plan to conduct a statistical test with a requirement that our data are normally distributed. How will we test for normality? By visually inspecting histograms or Q-Q plots? By running a Shapiro-Wilk test? Doing all three? What if one or all of these approaches provides evidence of a non-normal distribution? Should we transform the data? Run a different statistical test? Report the results ‘as is’ with a cautionary note? The reality is that making any of these decisions is justifiable and therefore which decision is ‘best’ is ambiguous. This becomes problematic when the outcomes of these decisions differentially lead to results that either are or are not statistically significant. Simmons et al. (2011) referred to this as researcher degrees of freedom and argued that “when we as researchers face ambiguous analytic decisions, we will tend to conclude, with convincing self-justification, that the appropriate decisions are those that result in statistical significance” (p. 1360). Of critical importance, it is necessary not to think about QRPs as something that only ‘evil’ or ill-intending researchers do because if we do, then we will conceptualize them as something that only ‘others’

do and will not accept seeing ourselves as potential contributors to the problem (which we likely are). By preregistering a study, we provide a public record of which of these decisions we made ahead of time, thus increasing transparency and potentially decreasing QRPs.

There are two main types of preregistration that differ with regard to whether a peer review component is incorporated prior to data collection. In its most basic form, the term preregistration simply implies that a research plan has been registered in some way (see Practical Considerations and Resources for Study Preregistration section) and does not necessarily include any type of peer review. Another type of preregistration, known as registered reports, does include a peer review stage and represents a modified approach to the publication process (see e.g., Marsden et al., 2018). Registered reports implement a peer review stage *prior* to data collection during which time the study justification (i.e., literature review) and methods are evaluated by peer reviewers (often referred to as Stage 1). If the outcome of the peer review process is positive, a journal offers ‘in-principle acceptance’ to the authors indicating that regardless of the outcomes of the study the manuscript will be published so long as the preregistered plan is adhered to and the reporting of results and subsequent discussion are reasonable. This is not to say that no modifications are allowed (see Common Critiques and Misconceptions section), but any deviations must be documented, and unplanned analyses should be explicitly identified as such. The manuscript then goes through a second round of peer review after data is collected and results and discussion are reported (referred to as Stage 2). Critically, the study cannot be rejected simply because null results are found. Some journals (e.g., *Bilingualism: Language, and Cognition*, *Journal of Child Language*) have incorporated a Stage 0 phase in which authors submit a short (typically 1-2 page) letter of intent justifying their pursuit of publication via registered report which the editors then approve or reject.

The registered report approach to publication is argued to offer a solution to the publication bias problem in that it shifts incentives for both authors and reviewers (Nosek, 2020). If a study is being evaluated before any data have been collected or results are known, then reviewers can focus on judging whether the research questions are interesting and relevant and whether the methods are high quality as opposed to being influenced by whether the results were statistically significant or not. Similarly, authors are no longer incentivized to report clean, statistically significant results because in the registered report format, they would receive manuscript acceptance before outcomes are known. If publication bias is an issue, then registered reports should yield an increase in the publication of null (or at least non-significant) results. Indeed, Nosek (2020) reported that a comparison of publications in psychology journals offering the registered report format indicated that only 10% of publications following the traditional format indicated null results while that number jumped to 60% for registered reports in the same journals. While this modified approach to the publication process might appear to require greater time and resources (see Practical Considerations and Resources for Preregistration section), some initiatives have been proposed to combine the peer review process of registered reports with grant applications (Nosek, 2020).

Beyond the potential for increased transparency, there are a variety of additional benefits of preregistering a study. For all forms of preregistration, one benefit is the careful attention to making critical study design decisions in advance and having a record of them. While this seems like something we should be doing anyway, it has been suggested that the accountability inherent in knowing the plan will be publicly available may increase the attention given to this important task (van't Veer & Giner-Sorolla, 2016). Having a written record of our own decision-making also decreases our need to remember decisions (Lakens, 2019) which can be important if time

has passed between our planning and data collection. Preregistering a study also demonstrates a certain confidence in one's work and in some cases allows one to receive a COS badge upon publication which shows readers a commitment to a priori decision making and transparency. Finally, if pursuing the registered report form, a major benefit is receiving feedback prior to data collection as well as the potential of in-principle acceptance.

Common Critiques and Misconceptions

While the previous section described a number of benefits and potential advantages of preregistration this is not to say that there are no concerns about or arguments against the practice (see e.g., Goldin-Meadow, 2016; Oberauer & Lewandowsky, 2019; Pham & Oh, 2021a, 2021b; Szollosi et al., 2020). Nevertheless, many of the common critiques found in current discussions appear to be better characterized as misconceptions about preregistration rather than actual reasons not to preregister. A recent exchange in the *Journal of Consumer Psychology* (Pham & Oh, 2021a, 2021b; Simmons et al., 2021a, 2021b) detailed arguments for and against preregistration. This section draws from this exchange and highlights four common critiques and misconceptions: Preregistration stifles creativity, Preregistration is too onerous (for both authors and reviewers), Preregistration is not suitable for all research types, and Preregistration prevents fraud and increases study quality but will result in reduced productivity.

Preregistration stifles creativity

One of the main critiques of preregistration is that the practice stifles creativity and exploration by 'locking' researchers into a plan that they cannot deviate from and discouraging unplanned analyses because they do not carry the same evidentiary status (Pham & Oh, 2021b). This is problematic because exploration is important and necessary for scientific advancement. One reason this critique might be better characterized as a misconception is because

preregistering a study does not force a researcher to never deviate from their plan nor does it disallow exploration. What preregistration does do is provides, in a transparent way to the rest of the research community, information about which decisions were preplanned and adhered to and which were not. This is a very critical distinction and deserves repeating. Preregistration neither disallows nor devalues exploratory work – it just has to be labelled as such. In fact, in some cases it would be problematic not to deviate from the preregistered plan, such as if authors realize there is a problem with or reason not to conduct a preplanned analysis (Simmons et al., 2021b).

Similarly, when analyzing the data, a researcher might notice an unexpected trend or finding in the data and choose to explore it with additional analyses. This is not only ‘allowable’ but likely desirable, again, with the caveat that such an exploration is clearly labeled as such.

Preregistration is too onerous (for both authors and reviewers)

Another common critique of preregistration is that it is onerous to implement for both authors and reviewers in that it is likely to be overly time-consuming and to require additional resources. For instance, Pham and Oh (2021b) pointed out “the overall administrative cost that a preregistration system entails in terms of (a) researchers’ time for preparing and submitting the preregistration documents [and] (b) gatekeepers’ (editors and reviewers) time for evaluating the alignment between submitted manuscripts and corresponding preregistrations” (p. 165).

Regarding the former, there are a variety of resources (see Practical Considerations and Resources for Study Preregistration section) in the form of technologies, infrastructure, templates, and guidelines that have surfaced to make the process of preregistration not only feasible but relatively simple and straightforward for authors (Lakens, 2019). Simmons et al. (2021b) argued that preregistration could not be that onerous for authors given that over 20,000

people had used the site AsPredicted.org (a freely available platform for study preregistration) without having an extrinsic reward for doing so.

Regarding the time necessary to evaluate alignment between manuscripts and preregistrations, this does represent a cost worth acknowledging and raises important, practical questions of implementation. Who, if anyone, should be responsible for managing the evaluation? Should this be the job of journal editors? Or, perhaps it is more likely that it will fall to the shoulders of (volunteer) reviewers. We do not have answers to these questions yet, and this cost – a concern that the process will be burdensome for editors, and especially reviewers – is probably a contributing factor to the fact that only a few journals in our field offer the registered report format. But this cost alone is not a sufficient reason to reject the practice of preregistration. In fact, if we focus specifically on the registered report format of preregistration, there have been arguments made that it might even reduce reviewer workload (Simmons et al., 2021b). For example, registered reports might reduce the number of overall submissions and rejections a manuscript goes through in its publication history given that feedback can be provided before data are collected and analyzed.

Preregistration is not suitable for all research types

Another critique of preregistration is that it is not suitable for all research types (e.g., qualitative research, exploratory research) and therefore is problematic because it will result in the devaluation of such work. Similarly, it seems inherent in some critiques of preregistration that an assumption is that those in favor of preregistration envision that eventually all studies will have to be preregistered. To address the final critique first which can be boiled down to ‘we should not support preregistration because not everything can be preregistered’ does not appear

to be a valid argument. Even if not all studies could be preregistered that does not provide a logical or reasonable argument against not preregistering those studies that would benefit.

At the foundation of the other two critiques is the idea that some research types cannot be preregistered and that preregistering some will devalue the work that is not preregistered. Regarding the former, Haven and van Grootel (2019) and Haven et al., (2020) argued that preregistration can also be useful for qualitative researchers. OSF has provided a template for qualitative researchers along with additional resources <https://osf.io/j7ghv/>. As for exploratory research, this too can likely benefit from preregistration. In addition to the fact that, “if a finding relies on *p*-values, confidence intervals, or Bayes factors, it purports to be confirmatory” (Simmons et al., 2021a, p. 178), other aspects of exploratory research such as sample size, sampling criteria, methods used to detect outliers, etc. can be preregistered. Finally, regarding the devaluation of other work, clearly that is not the intent of preregistration, and again, it is not a reason not to endorse preregistration.

Preregistration prevents fraud and increases study quality but will result in reduced productivity

The purpose of preregistration is neither to prevent outright fraud nor to increase study quality. These misconceptions about preregistration have been used to argue against their usefulness. For instance, critics state that PARKing (Preregistering After Results are Known; Yamada, 2018) could be a concern or authors could register multiple versions of a study. Yes, certainly these are possibilities, but (a) it seems unlikely that many would really take the time to do so, but (b) more importantly, preregistration is not attempting to combat outright fraud. Outright fraud is (unfortunately) always possible.

It is also a misconception that the purpose of preregistration is to increase study quality. The purpose of preregistration is to increase transparency which potentially allows others a better

opportunity to judge quality (Wagenmakers & Vazire, 2020). Nevertheless, it is not claimed that preregistration will inherently impact study quality. Finally, a concern about preregistration is that it might result in reduced productivity. While I am not aware of studies investigating this, (Wagenmakers & Vazire, 2020) stated this might happen, but she actually thinks it would be a potential benefit to the field to help establish a culture that encourages a slower and more thoughtful research process instead of one that rewards more and more prolific publishing.

Practical Considerations and Resources for Study Preregistration

In this section, I first discuss some of the practical considerations of preregistering a study, focusing on my experience with a registered report and continue by providing information about some of the resources available for preregistering studies. The registered report discussed in this section is Huensch and Nagle (2021) published with *Language Learning*. All materials, protocols, data, analysis code, and manuscript versions are publicly available via the following link on OSF: <https://osf.io/4j5cr/>.

One of the main considerations to keep in mind when preregistering a study, and writing a registered report in particular, is that it might represent a potential shift in the research process for some. One possible approach to conducting a study is to first plan the study, obtain any necessary human subjects or research ethics approval, collect and analyze the data, and then write up the results for publication. When completing a registered report, it is likely necessary to front load more of the writing and planning time as the typical timeline would begin with planning the study and writing up the first sections of the manuscript (introduction, literature review, methods). While some might be concerned about the additional time, in my own experience, the timeline lasted approximately the same duration. Table 2 provides a timeline comparison of the registered report to that of another publication/project I completed. When

considering the timelines, they are actually remarkably similar: Both manuscripts had two years between initial conception and manuscript acceptance and required approximately one year for IRB approval, data collection, and analysis. The benefits of the registered report approach were receiving feedback and input on the design via the peer review process and the comfort of knowing the study had already received in-principle acceptance.

Table 2. Study timeline comparison between a registered report and a traditional publication

Registered report: Huensch and Nagle (2021)	Traditional Publication: Huensch (2019)
Idea: Nov 2018	Idea: March 2016
Stage 1 first submission: March 15, 2019	IRB approval, data collection and analysis
Stage 1 first decision: May 7, 2019	First submission: May 4, 2017
Stage 1 second submission: July 16, 2019	First decision: July 22, 2017
Stage 1 second decision: Aug 31, 2019	Second submission: Sept 22, 2017
Stage 1 third submission: Sept 30, 2019	Second decision: Dec 26, 2017
Stage 1 third decision: Oct 1, 2019	Third submission: Jan 15, 2018
In-principle acceptance: Oct 5, 2019	Final acceptance: March 6, 2018
IRB approval, data collection and analysis	
Stage 2 first submission: Oct 5, 2020	
Stage 2 first decision: Nov 17, 2020	
Final acceptance: Nov 30, 2020	

Another consideration when completing a registered report has to do with responding to reviewer critique and requests for changes or additions to the study. The difference has to do with the fact that in the registered report format, the study has not been conducted and the data have not yet been collected. In our stage 1 interactions with reviewers this resulted in some requests for changes to study design which reviewers might have been less inclined to request had the study already been carried out. For instance, our initial plan was to collect rating data from 40 listeners using Amazon Mechanical Turk. Reviewers requested both that the number of listeners be increased and that an additional listener group be included using face to face data collection. The latter request represented a major deviation from our original intent, which was

not to compare different approaches to collecting rating data. This would have added an additional research question outside of the scope of the submitted manuscript. Nevertheless, as authors wanting to gain acceptance of our registered report, it was difficult not to feel pressured to give in to the reviewer requests as data had not yet been collected. In the end, it was important to remind both ourselves and the reviewers of the goals of the study and not to continue adding additional data and research questions simply because the study was in the planning stages. Similarly, if reviewers request additional analyses at stage 2, while this might be reasonable in some cases, they should be clearly marked as not originally included in the preregistration. Regarding the request for an increased AMT listener sample, however, we did incorporate this request, and the outcome was positive in that our final sample was more robust, and we were able to allay any power concerns without having to collect additional data and re-complete analyses at a later stage.

A third consideration is related to documenting and justifying any deviations from the stage 1 plan and/or additional analyses conducted. As discussed in the Common Critiques and Misconceptions section, the intention of preregistration is not to force researchers to commit to a plan without any deviations or modifications. In our case, we found it useful to submit a document which summarized and explained any changes between the Stage 1 and Stage 2 manuscripts (this document is available on OSF and as part of the supporting materials on the *Language Learning* website). This allowed us the opportunity to maintain transparency and justify our decision-making when deviations did occur.

A final practical consideration related to preregistration is likely the need to plan for open sharing of materials, protocols, and data. For those working with human subjects, this would entail planning for such sharing as part of the research ethics approval. Sharing of data and

materials also requires a commitment to making public documents understandable and useable for the broader research community. In the remaining portion of this section, I provide information about several resources that are useful in this regard.

In terms of creating a preregistration document, there is no single agreed upon format. For registered reports, follow the guidelines provided by the journal. For non-peer reviewed preregistrations, as a start I would recommend browsing examples and templates to get a better idea of what others have done both in and outside of our field. One helpful resource is a template created by the American Psychological Association (APA) available [here](#). OSF provides a variety of information about and resources for preregistration [here](#). There are also sites that offer free housing of preregistration, for example <https://aspredicted.org/> and <https://osf.io/prereg/>. These sites also provide templates and have examples of registered studies. Examples of my study preregistrations can be found at <https://osf.io/w4gj2> and <https://osf.io/h592g>.

Future Considerations

As mentioned previously, preregistration is a relatively new practice in our field. This means that, as is the case for other open science initiatives, training and education both in the practical and theoretical considerations for researchers from graduate students to late career scholars is important (see also chapters in this volume by Campbell and Koessler and Hui and Huntley). Beyond disseminating information about what preregistration is, its purpose and benefits, offering workshops at conferences such as AAAL or incorporating aspects of preregistration into graduate seminars would be useful.

For students and seasoned researchers alike, we might also consider ways to continue to incentivize participation in these practices. For example, in their Early Career Research Grant competition, *Language Learning* gives priority to applications that indicate a registered report

format. Other ways to incentivize might be the continued adoption of open science badges, the organization of special issues in academic journals, and other types of scholarly recognition. In their [Statement on the Scholarly Merit and Evaluation of Open Scholarship in Linguistics](#), the Linguistic Society of America provides a section on incentivizing open scholarship which emphasizes that key documents such as annual reviews, job advertisements, and tenure and promotion reviews should acknowledge researchers' contributions to open scholarship.

As a closing thought, it is important to remember that preregistration is a relatively new initiative, particularly in the field of applied linguistics. While there is some empirical evidence for claims for and against preregistration, such as Schäfer and Schwarz (2019) who documented differences in effect sizes between preregistered studies and those that were not, we are mostly relying on logical and philosophical argumentation. The upshot of this is that we need more evidence, likely related to both the need for preregistration and the resulting outcomes. As we continue to implement and encourage open science practices in applied linguistics, we must simultaneously consider how to evaluate whether we have reached our goals and how to move forward based on the outcomes in the coming decades and beyond.

References

- Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, 100, 407–425.
- Byrnes, H. (2013). Notes from the editor. *Modern Language Journal*, 97, 825–827.
- Chiu, Y. H., Kao, C. W., & Reynolds, B. L. (2012). The relative effectiveness of digital game-based learning types in English as a foreign language setting: A meta-analysis. *British Journal of Educational Technology*, 43, E104–E107.
- Gass, S., Loewen, S., & Plonsky, L. (2021). Coming of age: The past, present, and future of quantitative SLA research. *Language Teaching*, 54, 245–258.
- Goldin-Meadow, S. (2016, September). Preregistration, replication, and nonexperimental studies. *Observer*. Retrieved from <https://www.psychologicalscience.org/observer/preregistration-replication-and-nonexperimental-studies>
- Haven, T. L., Errington, T. M., Gleditsch, K. S., van Grootel, L., Jacobs, A. M., Kern, F. G., Piñeiro, R., Rosenblatt, F., & Mekkink, L. B. (2020). Preregistering qualitative research: A Delphi study. *International Journal of Qualitative Methods*, 19, 1–13.
<https://doi.org/10.1177/1609406920976417>
- Haven, T. L., & van Grootel, L. (2019) Preregistering qualitative research. *Accountability in Research*, 26, 229–244, <https://doi.org/10.1080/08989621.2019.1580147>
- Huensch, A. (2019). Pronunciation in foreign language classrooms: Instructors' training, classroom practices, and beliefs. *Language Teaching Research*, 23, 745–764.
<https://doi.org/10.1177/1362168818767182>

- Huensch, A., & Nagle, C. (2021). The effect of speaker proficiency on intelligibility, comprehensibility, and accentedness in L2 Spanish: A conceptual replication. *Language Learning*, 71, 626–668. <https://doi.org/10.1111/lang.12451>
- Isbell, D., Brown, D., Chan, M., Derrick, D., Ghanem, R., Gutiérrez Arvizu, M. N., Schnur, E., Zhang, M., & Plonsky, L. (2022). Misconduct and questionable research practices: The ethics of quantitative data handling and reporting in applied linguistics. *Modern Language Journal*, 106, 172–195. <https://doi.org/10.1111/modl.12760>
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, 2, 196–217.
- Lakens, D. (2019, November 18). The Value of Preregistration for Psychological Science: A Conceptual Analysis. <https://doi.org/10.31234/osf.io/jbh4w>
- Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, 36, 345–366.
- Marsden, E., Mackey A., & Plonsky, L. (2016). The IRIS repository: Advancing research practice and methodology. In A. Mackey, & E. Marsden (Eds.), *Advancing methodology and practice: The IRIS repository of instruments for research into second languages* (pp. 1–21). Routledge.
- Marsden, E., Morgan-Short, K., Trofimovich, P., & Ellis, N. C. (2018). Introducing registered reports at Language Learning: Promoting transparency, replication, and a synthetic ethic in the language sciences, *Language Learning*, 68, 309–320.
- Marsden, E., & Plonsky, L. (2018). Data, open science, and methodological reform in second language acquisition research. In A. Gudmestad, & A. Edmonds (Eds.), *Critical reflections on data in second language acquisition* (pp. 219–228). John Benjamins.

- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
- Nosek, B. A. (2020, December). Culture change toward more open, rigorous, and reproducible research. Keynote address at Research Reproducibility.
<https://www.youtube.com/watch?v=FLYffIOmKl4>
- Nosek, B. A., & Lakens, D. (2014). Registered reports: A method to increase the credibility of published results. *Social Psychology*, 45, 137–141.
- Nosek, B. A., & Lindsay, D. S. (2018, February). Preregistration becoming the norm in psychological science. *Observer*. Retrieved from
<https://www.psychologicalscience.org/observer/preregistration-becoming-the-norm-in-psychological-science>.
- Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, 26, 1596–1618.
- Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, 7, 528–530. <https://doi.org/10.1177/1745691612465253>
- Pham, M. T., & Oh, T. T. (2021a). On not confusing the tree of trustworthy statistics with the greater forest of good science: A comment on Simmons et al.'s perspective on preregistration. *Journal of Consumer Psychology*, 31, 181–185
- Pham, M. T., & Oh, T. T. (2021b). Preregistration is neither sufficient nor necessary for good science. *Journal of Consumer Psychology*, 31, 163–176.

- Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analysis, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition*, 35, 655–687.
- Porte, G. (Ed.) (2012). *Replication research in applied linguistics*. Cambridge University Press.
- Porte, G., & McManus, K. (2019). *Doing replication research in applied linguistics*. Routledge.
- Roberson, A. J., Farmer, R. L., Shaw, S., Upton, S., & Zaheer, I. (2020, December 11). Opening Doors to Open Science and Scholarship for School Psychology Research, Training, and Practice. <https://doi.org/10.31234/osf.io/su53p>
- Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, 10, 813.
- 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, 1359–1366.
- Simmons, J., Nelson, L. D., & Simonsohn, U. (2021a). Pre-registration is a game changer. but, like random assignment, it is neither necessary nor sufficient for credible science. *Journal of Consumer Psychology*, 31, 177–180.
- Simmons, J., Nelson, L. D., & Simonsohn, U. (2021b). Pre-registration: Why and how. *Journal of Consumer Psychology*, 31, 151–162.
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: a key to the file-drawer. *Journal of Experimental Psychology: General*, 143, 534–547.
- Stroebe, W., Postmes, T., & Spears, R. (2012). Scientific misconduct and the myth of self-correction in science. *Perspectives on Psychological Science*, 7, 670–688.

Szollosi, A., Kellen, D., Navarro, D. J., Shiffrin, R., van Rooij, I., Van Zandt, T., & Donkin, C.

(2020). Is preregistration worthwhile. *Trends in Cognitive Sciences*, 24, 94–95.

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.

Vazire, S. (2018). Implications of the credibility revolution for productivity, creativity, and

progress. *Perspectives on Psychological Science*, 13, 411–417.

Wagenmakers, E. J., & Vazire, S. (2020). Slides: Preregistration Standards for Psychology. ZPID

(Leibniz Institute for Psychology) <https://doi.org/10.23668/PSYCHARCHIVES.4464>

Yamada, Y. (2018). How to crack pre-registration: Toward transparent and open

science. *Frontiers in Psychology*, 9, 1831. <https://doi.org/10.3389/fpsyg.2018.01831>