

RESEARCH-BASED PRACTICE

Questionable Research Practices and Evidence-Based Practices in School Psychology

By Joseph D'Intino, Ekaterina Lysenko & Steven R. Shaw

pp. 8-11

Volume 47 Issue 4

By Joseph D'Intino, Ekaterina Lysenko & Steven R. Shaw

Evidence-based practices (EBP) in psychology and education are necessary for the development of best practices and required by law. However, EBP is a term that is often freely applied in these settings without considering the implications of the meaning of “evidence-based.” Evidence-based practices are defined as the process of carefully, conscientiously, and judiciously evaluating the best available research evidence in the application to clinical practice and service delivery. Similarly, the Every Student Succeeds Act (ESSA, 2015) presents four levels of EBP to guide practitioners in selecting the most appropriate interventions, tools, and techniques. Those levels are:

- Tier 1—Strong Evidence: supported by one or more well-designed and well-implemented randomized control experimental studies.
- Tier 2—Moderate Evidence: supported by one or more well-designed and well-implemented quasi-experimental studies.
- Tier 3—Promising Evidence: supported by one or more well-designed and well-implemented correlational studies (with statistical controls for selection bias).
- Tier 4—Demonstrates a Rationale: practices that have a well-defined logic model or theory of action, are supported by research, and have some effort underway by an educational system or outside research organization to determine their effectiveness.

Although the ESSA provides a framework for decision-making, the quality of the research is the foundation of whether an intervention, program, or technique meets minimum standards of quality to be implemented (Rycroft-Malone et al., 2004). Taking for granted that all research is conducted and reported in the most accurate and effective manner is easy and problematic. The common assumption is that because research findings have been found to be statistically significant and vetted by and published in a scholarly journal, the results therefore provide strong evidence for the effectiveness of a given intervention. Because of several questionable research practices, the evaluation of research for high quality support or evidence is a challenging task (Ioannidis, 2016a).

Research is often minimized in psychology and education (Biesta, 2007). The number one reference for teachers to learn about and acquire interventions is Pinterest, followed closely by Google searches and other social media outlets (Carpenter, Cassaday, & Monti, 2018). Moreover, it usually requires more than 10 years between the publication of research results and implementation of such research in the training and clinical practice of educators and psychologists (e.g., Buchs, Filippou, Pulfrey, & Volpé, 2017). Usually, this research filters through secondary sources such as workshops, textbooks, policy and regulation development, and lectures from preservice and inservice presenters. Expecting teachers and psychologists to implement interventions that are supported by quality scientific research requires that research be credible, transparent, effectively communicated, time-saving, resource efficient, generalizable, useful in diverse settings, easy to implement, and sound (Lyon et al., 2013). These factors are not related to the four tiers of evidence-based practice described in ESSA.

The improvement of interventions from school psychologists relies on research as the communicator of effective strategies, techniques, interventions, and practices (Ioannidis, 2017). Research serves to validate what works, to explore the plausibility of new ideas, and to determine which practices are effective. However, several questionable practices in education and psychology research contribute to problematic consequences for these fields (Ioannidis, 2005). Among these consequences are ineffective interventions being implemented to vulnerable populations; wasted financial, time, and labor resources; and the dissemination of inaccurate information that eventually influences future research, policy, and clinical practice in school psychology.

Although research is never without limitations, clinicians, practitioners, and professionals must maintain a critical perspective when evaluating research findings to assert quality and relevance (Ioannidis, 2016b). However, the use of questionable research practices skews interpretations, reporting, or the scientific process and, therefore, hinders the progress of education and psychology (John, Loewenstein, & Prelec, 2012). Most school psychologists are trained in a scientist–practitioner model. Part of being a scientist–practitioner is instruction in research methods. However, many of the basic threats to internal and external validity such as random assignment of participants to condition, implementers and evaluators who are blind to condition, representative samples, reliable and valid measurement of outcomes, and other basics of research design and analysis are necessary, but not sufficient, for research to be applied as evidence-based practice. To strengthen clinical practices, improve the implementation of EBPs, and to create an evidence-based profession, researchers and clinicians must have a strong understanding of all questionable research practices (QRPs) that affect the quality of research.

Questionable Research Practices

There are several well-described and common QRPs that affect most of the published papers in psychology (John et al., 2012; Open Science Collaboration, 2015). QRPs limit the inferences that can be made from results and negatively influence the practical utility of research (Van der Zee & Reich, 2018).

Publication bias. Journal editors prefer publishing innovative, ground breaking, and exciting research that shows positive effects for interventions (Ferguson & Brannick, 2012). This creates a bias in which statistical significance at the .05 level is preferred for journal acceptance over quality of research design. Publication bias tends to encourage poor research practice as there is a well-known phenomenon that the stronger the research design, the smaller the effect size of any intervention (Ferguson & Heene, 2012). This creates a perverse incentive for researchers to prioritize the likelihood of a statistically significant intervention over high quality research design (Anderson, Kelley, & Maxwell, 2017). In evidence-based practice, comprehensive research reviews and meta-analyses of multiple research studies are considered the highest level of evidence. Yet, because of publication bias, comprehensive literature reviews and meta-analyses are simply a compilation of studies with weak research designs and statistically significant findings (Franco, Malhotra, & Simonovits, 2014). It is unknown exactly how many nonsignificant research studies for a given intervention were never published due to publication bias.

Underpowered studies. Low-powered studies, often due to small sample sizes, are especially common in educational and psychological research. Anyone conducting research in special education is aware that gaining a sufficient number of participants for a reasonably powered study is daunting. However, low-powered studies dramatically limit inferences that can be made about the findings (Greenland et al., 2016). One of the limitations concerns inferences about how the information can be applied and implemented with other samples and other populations.

P-hacking. This QRP is also known as data dredging, fishing, and snooping (Mark, Lee, & Harrell, 2016). Researchers collect or select data or conduct statistical analyses until nonsignificant results

become significant. These practices lead to true effect sizes in published studies to be inflated or misreported. Common practices include conducting analyses midway through experiments to decide whether to continue to collect data to achieve statistical significance, collecting data on a large number of variables with a large sample and searching for relationships at the .05 level and then explaining them in a post hoc fashion, recording many response variables and deciding which to report after the analysis, and deciding whether to include or drop outliers in the analysis (Chambers, 2013). Studies using P hacking are simply capitalizing on error and have little meaning or implication for practice.

Harking. Hypothesizing After the Results are Known, also referred to as harking, is another QRP that creates opportunity for the reader to falsely judge the quality and relevance of a given study (Kerr, 1998). The major concern regarding harking is that the practice completely defeats the scientific process whereby hypotheses and research questions serve to guide and identify the gaps in the research by asking relevant questions that are answered through well designed research methods and data collection. Writing the hypotheses after knowing the results misleads the reader and serves only to bolster the impression that the authors were initially correct in their assertions and hypotheses. Moreover, the valuable information regarding the initial hypotheses and how those hypotheses were informed by existing research and theories is lost through the process of harking. There are common cases where journal editors or reviewers request that authors change the hypothesis to be more consistent with the data. This is clearly inappropriate. Adhering to the scientific method and presenting hypotheses as they were initially conceptualized provides useful information to the reader and the field.

Outlier management. Outlier refers to extreme values that abnormally lie outside the overall pattern of a distribution of variables. Outliers result from various factors, including participant response errors and data entry errors. Managing outliers involves removing data points that fall within the extreme ranges. In most research, addressing outliers is an important part of data management and analysis strategies. However, in school psychology, most of us refer to the outliers as our clientele. Ignoring outliers in research may have the effect of excluding people with extreme behaviors, academic performance, or other functioning (Hagiliassis & Di Marco, 2017). Some forms of outlier management may deprive clinicians of useful inferences on extreme behaviors.

Conflating confirmatory with exploratory research. For research to support an intervention to be applied in educational psychological settings, there must be some evidence supporting the probability of causation. If the intervention does not cause the desired effect, then when implemented, the probability of the desired effect is lower. Exploratory (often correlational) research is valuable and essential in school psychology. However, interpreting exploratory data as if it were confirmatory or causal data is a common mistake that has negative implications for implementation.

Conflicts of interest. Another QRP that can mislead readers occurs when the researcher has a conflict of interest in which there is an evident claim or stake in the result of the results. Just as in cases where pharmaceutical companies fund research regarding the medication that they manufacture, researchers may find themselves in situations in which they have a direct financial, or sometimes reputational, stake in the success of an intervention or instrument (Coyne, 2016). In most cases these conflicts of interest are disclosed; nevertheless, the conflict may prompt some researchers to report more favorable results or interpretations as they stand to benefit if the intervention or instrument results in significant positive outcomes. Potential benefits are not strictly financial; secondary nonfinancial benefits can include career advancement, publication, or providing support for a colleague or friend, among others (Coyne, 2016). Moreover, many of these biases may be unconscious and, despite the best intentions, researchers might still be engaging in a conflict of interest. As such, increased regulation regarding potential conflicts of interest are important for disseminating quality research.

Overgeneralizing findings to practice. Even a high-quality study that supports a specific practice should only be referred to as a proof of concept. Whether the effect sizes reported in a refereed article are specific to the context, system, resources, language, implementers, students of interest, measurement, developmental level, or other factors is unclear from a single study (Makel et al., 2016). There are few practices or interventions that are universally positive. Most practices or

interventions that yield positive results are effective for specific age groups at specific times in specific contexts with specific resources. Effectively matching high quality published research that shows positive effects to the actual needs and contexts is a critical aspect of implementing EBPs.

Lack of data sharing. Not having data available forces the reader to assume that the researchers have engaged in the most effective data management, data analysis, and reporting (Longo & Drazen, 2016). There are now repositories available so that data can be publicly posted and interested researchers can reanalyze data. Not all data are appropriate for sharing. However, the improved transparency of research will increase the credibility and improve the likelihood of the implementation of these evidence-based practices (Nosek et al., 2015).

Accessibility. Universal EBPs are impossible to realize when most of the practices appear in scholarly journals, which may only be accessed if there is a subscription or individual purchase of articles. The individual purchases can be as much as \$35 per article. Unless articles are available and easily accessible for teachers, therapists, and psychologists, implementation of the most up-to-date practices supported by research cannot be expected.

Lack of replication. In nearly all fields of science, when a phenomenon is discovered, the next step is to reproduce the result in the same lab with the same measures and procedures (Nosek & Errington, 2017). The step after that is to replicate the result in a different lab with slightly different procedures and samples. In education, when a phenomenon is discovered, it is immediately published, sold by consultants, turned into a business, and used to influence regulation and training, among other attempts to implement the new phenomena.

Over the last 5 years, many classic experiments in psychology have been found not to replicate (Baker, 2016). The foundation of psychological science is in crisis because of a lack of replication. A recent review of all database papers from three school psychology journals since 2007 finds that 1.2% of evaluative studies are exact replications, conceptual replications, or were subsequently replicated (Shaw, Lysenko, Wang, & Bitencourt, 2018). The foundation of evidence-based practice is not based on robust scientific processes (Munafò et al., 2017).

Implications and Remedies

Evidence-based practice is an outstanding idea that emphasizes the importance of research to practice. There is also the potential to provide powerful tools to teachers, therapists, and psychologist in clinical practice. The idea that research and practice are two independent and unrelated components is clearly incorrect—they are reciprocally dependent.

Noting QRPs is part of identifying current challenges that prevent the concept of EBPs from reaching its potential as useful in practice. There are approaches to remedy many of these problems. School psychology would be wise to recognize and consider some of these remedies.

Registered reports. This involves authors submitting the introduction, methodology, and data analysis strategies to a journal for peer review (Chambers, 2013). The peer review focuses on the quality of rationale, methods, and analysis plan used prior to data collection. Hypotheses and research methods are registered, archived, and publicly available. If the submission is acceptable for the journal, then data are collected. The article will be accepted no matter the direction of results. This approach eliminates several questionable research practices, including low statistical power, selective reporting of results, p-hacking, harking, low power, and publication bias. This eliminates the need for researchers to have statistically significant results by any means. There are currently 118 scholarly journals that accept registered reports. There is one in school psychology (i.e., *Canadian Journal of School Psychology*; Shaw, 2018).

Data sharing. This is the practice of making anonymous data used for scholarly research available to other investigators (Longo & Drazen, 2016). Data sharing allows for the reproducibility of study results and the reuse of old data for new research questions, thereby promoting future projects. An

important goal of data sharing is transparency of data analysis approaches. Many universities have resources for archiving data and there are other publicly available archives for data sharing.

Valuing replication. Encouraging replication of studies increases confidence that the results of the research are due to the treatment or intervention rather than error or bias (Makel & Plucker, 2014). Journals, reports, and student theses and dissertation often value uniqueness. Replication is considered by some to be a lesser form of science. Given how much the major findings in psychology have been challenged by recent attempts to replicate, replication is an essential aspect of social and physical sciences (Baker, 2016). Scholarly journals would make a much larger contribution to the development of EBPs if they valued and accepted replications rather than searching for the next big and exciting thing (LeBel, Campbell, & Loving, 2017). Some approaches could be a special section of journals that included exact or conceptual replications, encouraging students conducting masters theses to engage in replications, and otherwise valuing replications as a critical and essential part of scientific enterprise.

Increasing journal access. Open access journals allow for researchers and practitioners to access the article without having to pay fees, which increases the visibility and accessibility of the research (Hagiliassis & Di Marco, 2017). However, most open access journals involve the author paying sometimes significant publication fees to get their papers published. In addition, many predatory or junk journals are often open access, and it is easy to conflate legitimate open access journals with predatory journals. Many traditional journals have an open access option that involves the authors paying a fee so that everyone can access the specific article free of charge. Another option is to submit articles to a pre-print service so that scholars and clinicians can review the article before it is reviewed and published. Nearly all major journals allow an article to be submitted to a preprint service and then later submitted to the journal. This is a good opportunity to receive feedback before the article submitted to a journal and it is also a method by which access to the preprint journal article is always available even if the article later appears in a closed or subscription limited journal.

Conclusions

Clinicians and practitioners rely on high quality research when selecting evidence-based practices and interventions. However, QRPs can lead to questionable clinical practices and flawed decision-making. The effectiveness of the implementation of EBPs to improve the lives of children and their families will largely be dependent how researchers address QRPs.

References

- Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: A method adjusting sample effect sizes for publication bias and uncertainty. *Psychological Science*, 28(11), 1547–1562. doi:[10.1177/0956797617723724](https://doi.org/10.1177/0956797617723724)
- Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604), 452–454. doi:[10.1038/533452a](https://doi.org/10.1038/533452a)
- Biesta, G. (2007). Why “What Works” won’t work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory*, 57(1), 1–22. doi:[10.1111/j.1741-5446.2006.00241.x](https://doi.org/10.1111/j.1741-5446.2006.00241.x)
- Buchs, C., Filippou, D., Pulfrey, C., & Volpé, Y. (2017). Challenges for cooperative learning implementation: Reports from elementary school teachers. *Journal of Education for Teaching*, 43(3), 296–306. doi:[10.1080/02607476.2017.1321673](https://doi.org/10.1080/02607476.2017.1321673)
- Carpenter, J., Cassaday, A., & Monti, S. (2018). Exploring how and why educators use Pinterest. In E. Langran & J. Borup (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference* (pp. 2222–2229). Washington, DC: Association for the Advancement of Computing in Education (AACE). Retrieved from <https://www.learntechlib.org/primary/p/182833/>

Chambers, C. D. (2013). Registered reports: A new publishing initiative at Cortex. *Cortex*, 49, 609–610. doi:[10.1016/j.cortex.2012.12.016](https://doi.org/10.1016/j.cortex.2012.12.016)

Coyne, J. C. (2016). Replication initiatives will not salvage the trustworthiness of psychology. *BMC Psychology*, 4, 28. doi:[10.1186/s40359-016-0134-3](https://doi.org/10.1186/s40359-016-0134-3)

Every Student Succeeds Act of 2015, Pub. L. No. 114-95 § 114 Stat. 1177 (2015-2016).

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods*, 17(1), 120–128. doi:[10.1037/a0024445](https://doi.org/10.1037/a0024445)

Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, 7(6), 555–561. doi:[10.1177/1745691612459059](https://doi.org/10.1177/1745691612459059)

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345(6203), 1502–1505. doi:[10.1126/science.1255484](https://doi.org/10.1126/science.1255484)

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology*, 31(4), 337–350. doi:[10.1007/s10654-016-0149-3](https://doi.org/10.1007/s10654-016-0149-3)

Hagiliassis, N., & Di Marco, M. (2017). Improving access to focused psychological interventions for people with intellectual disability. *InPsych: The Bulletin of the Australian Psychological Society Ltd*, 39(2), 11. Retrieved

from <https://search.informit.com.au/documentSummary;dn=922460951839643;res=IELAPA>

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Med*, 2(8), e124. doi:[10.1371/journal.pmed.0020124](https://doi.org/10.1371/journal.pmed.0020124)

Ioannidis, J. P. A. (2016a). The reproducibility wars: Successful, unsuccessful, uninterpretable, exact, conceptual, triangulated, contested replication. *Clinical Chemistry*. doi:[10.1373/clinchem.2017.271965](https://doi.org/10.1373/clinchem.2017.271965)

Ioannidis, J. P. A. (2016b). Why most clinical research is not useful. *PLOS Med*, 13(6), e1002049. doi:[110020490.1371/journal.pmed](https://doi.org/110020490.1371/journal.pmed)

Ioannidis, J. P. A. (2017). Acknowledging and overcoming nonreproducibility in basic and preclinical research. *JAMA*. doi:[10.1001/jama.2017.0549](https://doi.org/10.1001/jama.2017.0549)

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

Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. *Personality and Social Psychology Review*, 2, 196–217. doi:[10.1207/s15327957pspr0203_4](https://doi.org/10.1207/s15327957pspr0203_4)

LeBel, E. P., Campbell, L., & Loving, T. J. (2017). Benefits of open and high-powered research outweigh costs. *Journal of Personality and Social Psychology*, 113(2), 230–243. doi:[10.1037/pspi0000049](https://doi.org/10.1037/pspi0000049)

Longo, D. L., & Drazen, J. M. (2016). Data sharing. *New England Journal of Medicine*, 374(3), 276–277. doi:[10.1056/NEJMe1516564](https://doi.org/10.1056/NEJMe1516564)

Lyon, A. R., Ludwig, K., Romano, E., Leonard, S., Stoep, A. V., & McCauley, E. (2013). “If it's worth my time, I will make the time”: School-based providers' decision-making about participating in an evidence-based psychotherapy consultation program. *Administration and Policy in Mental Health and Mental Health Services Research*, 40(6), 467–481. doi:[10.1007/s10488-013-0494-4](https://doi.org/10.1007/s10488-013-0494-4)

Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty replication in the education sciences. *Educational Researcher*. doi:[10.3102/0013189X14545513](https://doi.org/10.3102/0013189X14545513)

Makel, M. C., Plucker, J. A., Freeman, J., Lombardi, A., Simonsen, B., & Coyne, M. (2016). Replication of special education research necessary but far too rare. *Remedial and Special Education*. doi:[10.1177/0741932516646083](https://doi.org/10.1177/0741932516646083)

Mark, D. B., Lee, K. L., & Harrell, F. E. (2016). Understanding the role of P values and hypothesis tests in clinical research. *JAMA Cardiology*, 1(9), 1048–1054. doi:[10.1001/jamacardio.2016.3312](https://doi.org/10.1001/jamacardio.2016.3312)

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Sert, N. P. du, ... Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1, 0021. doi:[10.1038/s41562-016-0021](https://doi.org/10.1038/s41562-016-0021)

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–1425. doi:[10.1126/science.aab2374](https://doi.org/10.1126/science.aab2374)

Nosek, B. A., & Errington, T. M. (2017). Making sense of replications. *ELife*, 6, e23383. doi:[10.7554/eLife.23383](https://doi.org/10.7554/eLife.23383)

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

Rycroft-Malone, J., Seers, K., Titchen, A., Harvey, G., Kitson, A., & McCormack, B. (2004). What counts as evidence in evidence-based practice? *Journal of Advanced Nursing*, 47(1), 81–90. doi:[10.1111/j.1365-2648.2004.03068.x](https://doi.org/10.1111/j.1365-2648.2004.03068.x)

Shaw, S. R. (2018). CJSP: On reaffirming a Canadian school psychology. *Canadian Journal of Psychology*, 33, 3–7. doi:[10.1177/0829573517748212](https://doi.org/10.1177/0829573517748212)

Shaw, S. R., Lysenko, E., Wang, L., & Bitencourt, A. N. (2018). *Improving the scientific foundations of evidence-based practices in school psychology: The role of registered reports and replications*. Poster presented to the International Congress of Applied Psychology. Montreal, QC, Canada.

Van der Zee, T., & Reich, J. (2018). Open education science. *AERA Open*, 4, 1–15. doi:[10.1177/2332858418787466](https://doi.org/10.1177/2332858418787466)

JOSEPH D'INTINO is a doctoral student, EKATERINA LYSENKO is a masters student, and STEVEN R. SHAW is an associate professor in the School and Applied Child Psychology program at McGill University.