Building Open Science into Graduate Training in Clinical Psychology

Jeffrey M. Girard

Department of Psychology, University of Kansas

Author Note

Jeffrey M. Girard https://orcid.org/0000-0002-7359-3746

Correspondence concerning this article should be addressed to Jeffrey M. Girard, Department of Psychology, University of Kansas, 1415 Jayhawk Blvd (Room 426), Lawrence, KS 66045, USA, Email: jmgirard@ku.edu

Abstract

Open science promises greater transparency and rigor, yet clinical psychology has lagged in integrating it into graduate training. This commentary builds on recent evidence that open-science instruction in U.S. programs remains uneven and largely conceptual. I argue that progress requires more than awareness: it depends on pedagogical scaffolding, computational infrastructure, and cultural alignment. Embedding reproducible, code-based workflows, literate programming, and ethical transparency throughout training can make openness both feasible and normative. By treating transparency as a professional competency, supported by open educational resources and institutional incentives, clinical psychology can model how to cultivate a sustainable culture of open science.

Keywords: commentary, open science, graduate training, research methods

Building Open Science into Graduate Training in Clinical Psychology

Open science offers powerful tools for transparency and rigor, yet clinical psychology has lagged behind other subfields in embedding them in graduate training. Much of this lag reflects distinct headwinds: handling sensitive patient data, coordinating with clinics and IRBs, and studying rare conditions (Tackett et al., 2017). Still, these factors do not fully explain the gap. Core practices such as preregistration, code sharing, and reproducible workflows can be taught and adopted despite them.

Drawing on syllabus audits and graduate-student surveys, Van Til et al. (in press) document that open-science training is present in some U.S. programs but uneven overall. Readings and occasional assignments appear, yet hands-on opportunities to practice core behaviors are scarce. Students are generally interested, but many remain unfamiliar with practices such as registered reports.

Writing as a clinical psychologist with a strong computational focus, I take an opinionated stance: passive exposure to open science principles is not enough to meaningfully reform clinical psychology. Graduate training must provide sustained, scaffolded, and experiential instruction that turns abstract ideals into everyday habits. Preparing students to engage in open science also means teaching computational tools and workflows that, though not historically central to psychology training, are now vital for transparency, reproducibility, and effective scientific communication in a digital age. Building on Van Til et al.'s findings, I argue that meaningful progress depends on coordinated action across three domains: pedagogy, infrastructure, and culture.

Building Open Science into Graduate Training

If psychology is to prepare the next generation of scientifically rigorous researchers, open-science principles must be woven into graduate training from the outset. The first year of training plays a disproportionate role in shaping a student's professional identity (Bruss & Kopala, 1993; Cruess et al., 2015), and early methods and statistics courses are key sites for transmitting not only design and analysis skills but also the norms of transparent, reproducible science. These courses can introduce the rationale for openness alongside hands-on activities such as structured

preregistrations, reproducible code, and transparent reporting. As students progress, they can apply these practices to increasingly independent projects: designing studies, preregistering hypotheses, analyzing original data, and depositing outputs in open repositories. By the dissertation stage, students should be fluent in open workflows and optionally experiment with advanced forms such as registered reports or collaborative version control.

Because methods and statistics are typically taught across subfields (Aiken et al., 2008), instructors can collaborate with content-area faculty to ensure open-science tools are applied in domain-relevant ways. Once students have core competencies, exercises can include parallel prompts or datasets from areas like clinical, social, or developmental psychology (graded using shared criteria). Low-effort connectors such as guest visits or interdisciplinary panels can help students appreciate both the generalizability of transparent workflows and the unique challenges of applying them within their own field. This coordinated approach keeps instruction inclusive and relevant without requiring major curricular overhauls.

For instructors, sustainable adoption depends on both pedagogical infrastructure and institutional support. Open educational resources (OER) such as Creative-Commons–licensed textbooks and syllabi, assignment templates, example repositories with de-identified or synthetic data, and instructor guides lower preparation costs while standardizing quality (Mondal, 2025; Wiley et al., 2014). Scientific societies and departments can amplify their impact by curating peer-reviewed teaching portals, issuing micro-grants for OER development, or creating starter kits aligned with accreditation standards.

Still, even well-intentioned instructors often lack the time, training, or institutional backing to redesign courses. Adoption becomes far more feasible when departments provide shared teaching assets: template-based preregistrations, staged code assignments, workshop materials, and quick-start guides for relevant tools. Regular peer-teaching circles, curated assignment repositories, and a designated "open-science lead" (or small working group) can help maintain internal documentation, host demos, and troubleshoot curricular updates.

Tangible incentives further reinforce adoption: micro-grants, summer stipends, targeted

TA support, and one-time course releases for substantial redesigns. Recognizing open-science teaching artifacts and OER creation as scholarly contributions in annual reviews and promotion criteria closes the loop. Embedding transparency in graduate education thus requires not only sound pedagogy but also the technical and institutional infrastructure that makes open practices practical, visible, and valued (Mabile et al., 2025).

Tools and Workflows that Empower Transparency

Open science is grounded in values and policies, but it is sustained by tools and technologies that make those values actionable. The tools students learn early on often become the scaffolding for their scientific careers; ensuring they support reproducibility, transparency, and collaboration is therefore essential.

Code-Based Workflows

Perhaps the most transformative shift in recent research practice has been the move from costly, point-and-click statistical software to open-source, code-based workflows. Whether students use R, Python, or another programming language, writing and documenting analytic scripts fosters a level of transparency and reproducibility that manual workflows cannot match. Analyses encoded in readable, versioned scripts can be reviewed, rerun, and built upon by others. The analysis itself becomes an object of scientific discourse, not a black box (Sandve et al., 2013).

Some point-and-click tools can export "syntax" or logs of the steps taken. This is better than nothing and can aid transparency, but exported syntax is often incomplete, hard to reuse, and brittle across software versions. In contrast, fully code-based workflows produce clear, shareable scripts that plug into later practices we describe (e.g., keeping a clean history of changes and packaging the computational environment) making verification, collaboration, and long-term reuse more reliable.

Analytic code should be treated as part of the scientific method, not an optional add-on. While data may justifiably be protected for legal or ethical reasons, there is no scientific basis for withholding code. To do so is to obscure essential elements of the analysis pipeline. Yet many students are never taught to see code as a core scientific output. This pedagogical omission is both

consequential and fixable. Journals can accelerate progress by requiring open code (or at least syntax) alongside publications, reinforcing the norm that transparency extends to all parts of the research process.

Literate Programming

Literate programming treats code and narrative text as equally important parts of scientific communication. Open-source publishing systems like Quarto and Jupyter Notebook implement this approach by allowing researchers to seamlessly combine text, code, and output (figures, tables, diagnostics) in a single document (Perkel, 2018; Scheidegger et al., 2025). Among these, Quarto stands out for its support of multiple languages (including R and Python) and scholarly features such as citations, equations, and cross-references, and its ability to produce a variety of reproducible formats ranging from reports and manuscripts to slides, websites, or entire books.[1]

In graduate courses, literate programming encourages students to document their reasoning, explain decisions, and produce reports that can be rerun at any time. This approach is ideal for methods assignments, replication projects, and theses. It also reflects how high-quality research is increasingly disseminated—not just through manuscripts but through richly documented supplemental materials that promote transparency and scientific communication. These same practices prepare students to collaborate on interdisciplinary research teams and open doors to roles beyond academia, including in data science and applied research, where reproducible workflows are essential.

Beyond the Basics

While more advanced, two computational frameworks further enhance transparency and reproducibility. **Version control** helps you keep a running history of your work so you can see what changed in a document, undo mistakes, and work with others without overwriting each other. Today, Git is the standard tool for version control, and services like GitHub and GitLab make it easy to store projects online (either publicly or privately) and to share them with collaborators

---

[1] This article was created using Quarto. See github.com/jmgirard/cps2025 for the code.

([Bryan, 2016](#); [Chacon & Straub, 2014](#)). Using version control builds accountability and clarity: anyone can trace what was done and when. As a practical student project, publishing a basic researcher website built with Quarto and hosted for free on GitHub introduces versioned, public-facing scholarship in a low-stakes, hands-on way.[2]

Environment management aims to make the same code run the same way anywhere by reproducing the software environment (i.e., operating system, system libraries, programming-language version, and packages). Tools like R's *renv* and Python's *virtualenv* record and restore the exact package versions used and, for even more reliability, other tools can create a self-contained copy of the entire environment. For R users, *rang* ([Chan & Schoch, 2023](#)) makes it straightforward to build such copies with Docker ([Merkel, 2014](#)), locking the operating system, R version, and packages to known snapshots. Packaging the environment with the project (e.g., alongside a manuscript or preregistration) makes complex work easier to share, rerun, and verify years later, strengthening replication and extension.

Though more technical, these methods future-proof scientific work and are increasingly vital for interdisciplinary and collaborative projects.

Tools as Values in Action

These tools do more than ensure technical accuracy: they teach a way of thinking. Code-based workflows, literature programming, version control, and environment management (Table [1](#)) foster habits of documentation, foresight, and humility. They signal that science is not only about reaching results but about making the process intelligible and reviewable. By teaching these tools explicitly and modeling their use, programs can help students build workflows that are not only reproducible but also resilient and collaborative. In a world where scientific claims face growing scrutiny, these capacities are not optional; they are part of what it means to do science well.

Beyond technical robustness, open tools reduce inequities in access to knowledge.

---

[2] See my academic lab's website at https://affcom.ku.edu and its Quarto source code at https://github.com/jmgirard/affcomlab.

Open-source software, shared teaching resources, and transparent workflows allow smaller or underfunded programs to match the standards of well-resourced institutions. By emphasizing openness as both a scientific and an equity value, graduate training can broaden who participates in and benefits from psychological science.

These same values extend beyond the classroom and into publication. Journals and editors occupy a powerful position in normalizing open-science workflows. When submission systems reward transparency (e.g., through preregistration badges, open-data links, or replication features), they reinforce the habits cultivated in training. Aligning incentives across education and publication ensures that the same principles guiding data collection and analysis also guide dissemination and recognition.

Yet even the best tools cannot overcome the ethical, institutional, and cultural barriers that shape how openness is practiced. These realities are particularly salient in clinical psychology.

Ethical and Institutional Realities in Clinical Training

Clinical psychology occupies a distinctive position at the intersection of scientific training and professional regulation. Because most programs operate under formal accreditation standards, accrediting bodies such as the American Psychological Association (APA) and the Psychological Clinical Science Accreditation System (PCSAS) exert substantial influence over curricular priorities. Depending on how they define competencies in research and ethics, these organizations can either facilitate or hinder the adoption of open-science practices. Clinical psychology thus functions as a model-setting domain for embedding transparency within professional education, where accreditation requirements, competency benchmarks, and supervisory structures can translate principles into enforceable expectations.

At the same time, clinical training entails challenges few other areas face (Tackett et al., 2017). Ethical and legal imperatives of patient confidentiality complicate sharing of raw data, preregistration of sensitive protocols, and replication of clinical trials (e.g., Ness, 2007). Research partnerships with hospitals, clinics, and community agencies are governed by data-use agreements that often limit dissemination. Students must also balance research with coursework and intensive

clinical practica, leaving limited bandwidth to master new methods. These constraints can make open-science ideals seem aspirational rather than achievable.

Yet the very features that complicate openness also create opportunities for innovation. Clinical psychology can pioneer frameworks for *ethical transparency* that balance rigor with privacy. Examples include using de-identified or synthetic datasets for teaching, secure data enclaves for authorized replication, and preregistration formats that allow embargoes until sensitive work concludes (e.g., Howison et al., 2024; Liu et al., 2025). Because clinical training already integrates ethics and reflective supervision, open-science competencies can be framed as extensions of professional responsibility rather than external mandates. Linking transparency to ethical conduct reinforces its relevance to both science and care.

Even with these structures in place, real research is messy. Students still learning statistics may lack the knowledge to preregister analyses in full detail, and projects often evolve in response to data patterns or reviewer feedback (Nosek et al., 2019; Simmons et al., 2021). Rigid preregistration can backfire if it fosters box-checking over critical reflection. Instead, students should be taught to view preregistration as a flexible tool for clarifying thought, not constraining it: distinguishing exploratory from confirmatory work, documenting changes transparently, and recognizing that deviation is part of learning. Similarly, registered reports should be seen not as universal solutions but as gold standards for certain kinds of confirmatory research.

Another subtle barrier is the culture of perfectionism that pervades many graduate programs. When open science is framed in idealized terms, students may equate rigor with flawlessness and transparency with self-exposure. For trainees already navigating imposter syndrome (Cohen & McConnell, 2019), this can breed paralysis rather than curiosity. Faculty (in both teaching and research roles) can counter this by modeling vulnerability—sharing imperfect code, troubleshooting errors in front of students, and emphasizing that open science is about clarity, not perfection (Strand, 2025).

Finally, not everyone shares the same values or goals that motivate open-science advocacy (e.g., Bazzoli, 2022). Priorities can vary widely from program to program, faculty to faculty, and

lab to lab; thus, students may encounter mixed messages when course expectations emphasize openness but local mentoring norms do not (or vice versa). While many view transparency as central to scientific integrity, others prioritize innovation, idiographic understanding, or community impact. Some are skeptical that practices like preregistration and data sharing improve understanding; others worry these practices may delay student progress or disadvantage those working with limited resources. These differences reflect genuine pluralism in how scientists understand the purpose of research. Because clinical psychology operates at the nexus of scientific and professional ethics, conversations about open science must engage (rather than override) this diversity of commitments. Promoting openness, therefore, is as much a cultural project as a technical or pedagogical one. It requires empathy, dialogue, and respect for principled disagreement.

Conclusion

Van Til and colleagues have provided a valuable service by empirically documenting what many open-science advocates in clinical psychology have long suspected: while interest in transparency and rigor is growing, graduate training has yet to catch up. Their registered report offers both a clear-eyed assessment of the current state and a hopeful reminder that students are eager to engage when given the opportunity.

But awareness is only the first step. For open science to take root in clinical psychology, programs must invest in the pedagogy, infrastructure, and culture that make transparency not just possible, but normative. That means treating reproducibility as a core competency, integrating open-science principles and tools early in the curriculum, supporting faculty with resources and time, and acknowledging that transparency will look different across projects. Imperfection is not a flaw but a feature of honest science.

At the same time, accrediting bodies, journals, administrators, and funders can align incentives by treating transparency as evidence of rigor and rewarding the labor that sustains it. Graduate students themselves can be catalysts for change through peer-learning groups and shared repositories. Open science thus becomes a collective practice, not a solitary ideal—sustained by

aligned structures, supportive mentorship, and everyday habits of openness.

Open science is a practice we choose, repeatedly, in the face of complexity and constraint. It is a commitment to make our work intelligible to others and accountable to the communities we serve. For clinical psychological science, that commitment is essential—and so is the work of teaching it. A sustainable culture of open science begins in the classroom, is reinforced through mentorship, and is sustained by institutions that value good science not only for its outcomes but for its openness in process.

Acknowledgments

Author Contributions

Conceptualization: J. Girard; Writing – Original Draft Preparation: J. Girard.

Conflicts of Interest

The author declares that there were no conflicts of interest with respect to the authorship or the publication of this article.

References

Aiken, L. S., West, S. G., & Millsap, R. E. (2008). Doctoral training in statistics, measurement, and methodology in psychology: Replication and extension of Aiken, West, Sechrest, and Reno's (1990) survey of PhD programs in North America. *American Psychologist*, *63*(1), 32–50.

Bazzoli, A. (2022). Open science and epistemic pluralism: A tale of many perils and some opportunities. *Industrial and Organizational Psychology*, *15*(4), 525–528.

Bruss, K. V., & Kopala, M. (1993). Graduate school training in psychology: Its impact upon the development of professional identity. *Psychotherapy: Theory, Research, Practice, Training*, *30*(4), 685–691.

Bryan, J. (2016). *Happy Git and GitHub for the useR*. https://happygitwithr.com/

Chacon, S., & Straub, B. (2014). *Pro Git* (2nd ed.). Apress.

Chan, C., & Schoch, D. (2023). Rang: Reconstructing reproducible R computational environments. *PLOS ONE*, *18*(6), e0286761.

Cohen, E. D., & McConnell, W. R. (2019). Fear of fraudulence: Graduate school program environments and the impostor phenomenon. *The Sociological Quarterly*, *60*(3), 457–478.

Cruess, R. L., Cruess, S. R., Boudreau, J. D., Snell, L., & Steinert, Y. (2015). A schematic representation of the professional identity formation and socialization of medical students and residents: A guide for medical educators. *Academic Medicine*, *90*(6), 718.

Howison, M., Angell, M., & Hastings, J. S. (2024). Protecting sensitive data with secure data enclaves. *Digital Government: Research and Practice*, *5*(2), 1–11.

Liu, Q., Shakya, R., Jovanovic, J., Khalil, M., & de la Hoz-Ruiz, J. (2025). Ensuring privacy through synthetic data generation in education. *British Journal of Educational Technology*, *56*(3), 1053–1073.

Mabile, L., Shmagun, H., Erdmann, C., Cambon-Thomsen, A., Thomsen, M., & Grattarola, F. (2025). Recommendations on open science rewards and incentives: guidance for multiple stakeholders in research. *Data Science Journal*, *24*(15), 1–9.

Merkel, D. (2014). Docker: Lightweight linux containers for consistent development and deployment. *Linux Journal*, *2014*(239), 2.

Mondal, H. (2025). Copyright and creative commons licenses in scholarly publishing: A practical guide for researchers. *Journal of Clinical Imaging Science*, *15*, 32.

Ness, R. B. (2007). Influence of the HIPAA privacy rule on health research. *JAMA*, *298*(18), 2164–2170.

Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., van 't Veer, A. E., & Vazire, S. (2019). Preregistration is hard, and worthwhile. *Trends in Cognitive Sciences*, *23*(10), 815–818.

Perkel, J. M. (2018). Why Jupyter is data scientists' computational notebook of choice. *Nature*, *563*(7729), 145–146.

Sandve, G. K., Nekrutenko, A., Taylor, J., & Hovig, E. (2013). Ten simple rules for reproducible computational research. *PLOS Computational Biology*, *9*(10), e1003285.

Scheidegger, C., Woodhull, G., Dervieux, C., Teague, C., Allaire, J. J., & Xie, Y. (2025). *Quarto*. Posit, PBC. https://quarto.org

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2021). Pre-registration is a game changer. But, like random assignment, it is neither necessary nor sufficient for credible science. *Journal of Consumer Psychology*, *31*(1), 177–180.

Strand, J. F. (2025). Error tight: Exercises for lab groups to prevent research mistakes. *Psychological Methods*, *30*(2), 416–424.

Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., Oltmanns, T. F., & Shrout, P. E. (2017). It's time to broaden the replicability conversation: Thoughts for and from clinical psychological science. *Perspectives on Psychological Science*, *12*(5), 742–756.

Van Til, K., Phillips, N., Du, T., Rose, L., Miller, J., & Lynam, D. (in press). Open science training in APA-accredited clinical psychology programs: A registered report. *Clinical Psychological Science*.

Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (2023). *R for Data Science* (2nd ed.). O'Reilly Media. https://r4ds.hadley.nz/

Wiley, D., Bliss, T. J., & McEwen, M. (2014). Open educational resources: A review of the literature. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 781–789). Springer.

Table 1

Recommended Tools and Readings

Topic	Tool	Reading
Code-based workflows	tidyverse	Wickham et al. (2023)
Literate programming	Quarto	Scheidegger et al. (2025)
Version control	Git	Bryan (2016)
Environment management	rang	Chan and Schoch (2023)