

Zachary J. Schroeder, PhD

Minneapolis, MN · zachschrue@gmail.com · linkedin.com/in/zachary-schroeder-1816a2241 · github.com/zjschroeder

SUMMARY

Quantitative researcher (PhD, Psychology; Educational Data Science specialization) who turns structured and unstructured data into findings people act on. Six years of independent research: text analytics over 7.8M social-media posts, a three-study preregistered experiment program (N = 1,125), and validation of automated transcription, text-classification, and sentiment measures against human coding. 500+ hours teaching statistics to non-statisticians. NSF Graduate Research Fellow. R, Python, SQL, Qualtrics.

EXPERIENCE

Research Scientist and Statistics Instructor (Graduate Research Fellow), University of Oregon, Eugene, OR *Sep 2019 to Jun 2026*

- Consolidated six legacy studies (20 years of video-recorded interviews, 74 hours of conversation) into one analytic dataset: transcribed with Whisper, corrected by a research-assistant team I trained, and segmented with a model I fine-tuned on 5,000 hand-coded utterances (96.8% accuracy on a holdout sample).
- Built an ensemble of four transformer classifiers (weighted-majority voting) to tag dialogue acts at scale; reached 93.4% accuracy on a 1,000-utterance human-coded holdout, above published single-model benchmarks of 86 to 92%.
- Modeled which features of what people say predict how accurately others read them, using sentence embeddings (Python) and Bayesian multilevel Gaussian process models pooled across studies (R, Stan); checked robustness with prior and kernel sensitivity analyses and leave-one-out cross-validation.
- Independently built a Python and R text-analytics pipeline over 7.8M unstructured social-media posts (topic modeling, sentiment and affect scoring, engagement modeling by account tier); found anger and impoliteness drove engagement most sharply among mid-tier accounts.
- Designed and ran a three-study experiment program in Qualtrics (N = 1,125) on how sense of belonging changes the way people interpret ambiguous feedback: exploratory study, preregistered direct replication with power analysis and exclusion criteria, and preregistered extension; published in *Social Psychology of Education* (2025).
- Modeled academic and disciplinary outcomes for 6,351 students across two national longitudinal datasets with preregistered multilevel logistic regressions; found suspension disparities tracked adult judgment more closely than students' self-reported behavior.
- Delivered 500+ hours of statistics and research-methods instruction across 13 terms to classes of 20 to 150, through graduate statistics taught in R; trained and supervised ~50 research assistants on data collection, coding, and quality checks.

Executive Administrator, Medtronic, Mounds View, MN

May 2019 to Oct 2019

- Coordinated the merger of two departments (100+ employees) and produced an all-staff meeting for 1,500 employees, owning content, filming, live stream, and materials.

SELECTED PROJECTS

[A step-by-step SQL tutorial on a live SQLite database](#), [an interactive t-test and ANOVA teaching app with automated tests](#), and [a constrained-optimization schedule builder for 251 LA28 Olympic sessions](#) (all run in the browser). More at github.com/zjschroeder.

SKILLS

Languages and tools: R (expert: tidyverse, brms/Stan, lme4, ggplot2, Shiny, Quarto/R Markdown), Python (advanced: pandas, scikit-learn, Hugging Face transformers, sentence-transformers, NLTK), SQL (SQLite), Qualtrics, SPSS, Git/GitHub, SLURM/HPC

Analytics: Bayesian multilevel and Gaussian process modeling, logistic and multivariate regression, meta-analysis, psychometric scale validation, supervised and unsupervised machine learning, ensemble classification and holdout validation, data visualization, automated and reproducible reporting

Text analytics and NLP: speech-to-text transcription, sentence embeddings and semantic similarity, dialogue-act and topic classification (LDA, STM), sentiment and affect scoring (LIWC-22), TF-IDF, qualitative coding at scale

Research methods: exploratory-to-confirmatory designs, preregistration, replication and power analysis, survey instrument design, controlled experiments (online and in-person), human-in-the-loop annotation, mixed methods, interview coding

EDUCATION AND PUBLICATIONS

PhD, Psychology, University of Oregon, April 2026. Specialization in Educational Data Science.

BA, Psychology (honors, cum laude) and Theatre, Northwestern University, June 2018.

Two peer-reviewed articles (*Social Psychology of Education*, 2025; *Frontiers in Psychology*, 2025); eight conference presentations.

[Google Scholar](#)