Getting Research Done with ChatGPT Plus and Modern AI

A one hour training for quantitative social sciences

Juan Pablo Posada Aparicio

Institute for Replication, University of Ottawa

October 2025

Session goals

- Align on the Replication Games storyline, schedule, and roles before event day.
- Clarify randomization, treatment arms, and deliverables drawn from the preregistered design.
- Preview the AI training components that support literature, data, coding, writing, and review work.
- Reinforce responsible and reproducible practices across human-only and AI-assisted teams.

Roadmap, about 60 minutes

- 5 minutes. Pre-games context and storyline.
- 10 minutes. Study design, randomization, and assignments.
- 15 minutes. Event-day operations, deliverables, and support.
- 20 minutes. ChatGPT Plus feature tour across the research workflow.
- 5 minutes. Guardrails, reproducibility, and Q and A.
- 5 minutes. Codex overview at the end.

Pre-games storyline

- Kick off the Replication Games cohort and align on the narrative before event day.
- Friendly human versus AI-augmented challenge tests speed, accuracy, and issue-spotting.
- We study vertical gaps across expertise tiers and horizontal gaps across disciplines.

Team and support network

- Institute for Replication with Abel Brodeur coordinates the University of Ottawa hub.
- Support crew: Ghina Abdul Baki, Juan Pablo Aparicio, Bruno Barbarioli, Lenka Fiala, Derek Mikola, David Valenta.
- University of Ottawa hosts in person; virtual teams rely on Zoom and Slack. Email: instituteforreplication@gmail.com.

Study design essentials

- Around 300 participants stratified by expertise tier and discipline tag.
- Randomized 1:1 within strata to AI-assisted access versus human-only control.
- Task pool spans Economics, Political Science, and Psychology with assignments balancing in- and out-of-discipline exposure.

Treatment arms and tiers

- **Human control**. No external AI during the work window; document everything manually.
- **Cyborg arm**. ChatGPT Plus with Advanced Data Analysis, Deep Research, Agent Mode, and Codex CLI support.
- Tiers from Undergraduate to Professor; we log discipline tags, coding experience, and AI familiarity for heterogeneity analyses.

Participant prep checklist

- Complete this orientation and skim the reporting workbook (GitHub template).
- Accept the ChatGPT Team invite promptly; access stays scoped to event participants.
- Confirm hardware, software licenses (R/Stata/Python), VPN access, and data permissions before event day.
- Review the assignment email so you know your tier, discipline tag, arm, and team roster.

Event-day timeline and workflow

- 8:45 local check-in or remote login; 9:00 Dropbox folder and reporting sheet unlock.
- Read instructions, identify the focal result, and catalog received files.
- Reproduce the assigned result, logging timestamps; audit code for major and minor errors.
- Run robustness checks and keep the reporting sheet updated throughout the seven-hour window.

Deliverables, compliance, and support

- Submit the reproduced result, error log, and reporting workbook by 16:00, plus qualitative notes if helpful.
- Control arm pledges no AI; AI arm logs prompts, files, and outputs for internal usage analyses.
- Primary outcomes cover success, timing, error counts, and robustness; secondary outcomes review narratives and recommendations.
- Technical issues: contact Derek Mikola, Ghina Abdul Baki, Bruno Barbarioli, or Juan Pablo Aparicio. Design decisions: Abel Brodeur or Juan Pablo Aparicio.

Post-event follow-up

- Focus groups by treatment capture qualitative experience across arms.
- De-identified outputs enter the replication archive once the preregistration lock lifts.
- Participants receive summary results before journal submission and can provide feedback.

LLM basics

- Large language models predict the next token in context to follow instructions for language and code tasks.
- Strengths: structured generation, pattern completion, code drafting, translation. Weaknesses: hallucinations, stale knowledge, need for clear constraints.
- Your edge: provide roles, context, examples, and verification loops to steer outputs.

ChatGPT Plus toolkit

- **Advanced Data Analysis**. Run Python, upload files, and produce figures or tables in chat.
- **Browsing & Deep Research**. Reach current sources with citations and credibility checks.
- **Custom GPTs & Agent Mode**. Tailor assistants and supervise multistep execution inside your workflow.

Research workflow highlights

- Literature scaffolds surface theories, debates, data sources, and structured reference tables.
- Data documentation drafts dictionaries, README structures, and polished tables from messy notes.
- Analysis support covers exploratory summaries, cleaning scripts, reusable helpers, and visualization guidance.
- Writing and review assistance interprets results, drafts abstracts, flags causal language risks, and generates referee-style feedback.

Advanced features in practice

- Deep Research orchestrates medium-depth web reviews with explicit credibility and coverage criteria.
- Agent Mode sequences browsing, analysis, and drafting steps while keeping humans in the approval loop.
- Works best on scoped, auditable tasks where intermediate outputs remain inspectable.

Guardrails and checkpoints

- Track prompts, data inputs, timestamps, and versions so provenance stays transparent.
- Verify claims against trusted sources and store scripts or logs with version history.
- Keep workflows reproducible and aligned with data governance, disclosure, and consent policies.

Codex overview

- Codex is OpenAI's coding agent that works inside git-managed projects alongside your local tools.
- Natural-language instructions drive code exploration, edits, and command execution with human approvals.
- Terminal-first workflow surfaces plans, diffs, and checkpoints so changes stay aligned with existing practices.

Q and A

Thank you.