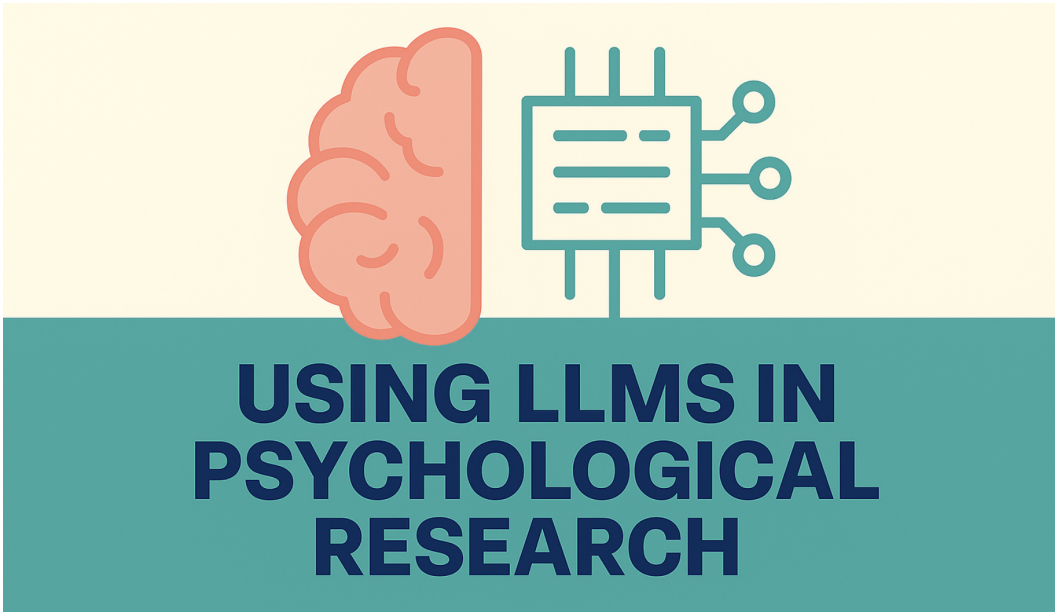
Sam Pratt
sampratt@g.ucla.edu
Updated May 2025



A Practical Guide for Researchers

Purpose

Generative artificial intelligence—especially Large Language Models (LLMs) like ChatGPT—are quickly becoming valuable tools in psychological research. This guide is designed to help researchers use these tools effectively, responsibly, and creatively. It includes useful resources for using LLMs at each stage of the research process. For other miscellaneous resources, see *Nature*'s guide to choosing the right AI tools for research.

Why This Matters

Psychological research is often time-intensive and cognitively demanding. LLMs offer a way to offload repetitive, low-stakes, or manually intensive tasks so that researchers can devote more energy to what matters most.

Understanding how to use these tools *well* is essential. Used thoughtfully, LLMs can improve productivity, clarity, and even creativity. Used carelessly, they can introduce errors, biases, or ethical oversights. This guide is designed to help you use these tools optimally and avoid common pitfalls.

Necessary Caveats

LLMs are useful but imperfect tools. Throughout this guide, you'll see section-specific "**Important Considerations**," but a few general caveats apply across the board:

1. **LLMs can hallucinate.** They may confidently generate false information or cite non-existent sources. Always double-check factual claims and references.

2. **Protect sensitive data.** Never paste identifiable participant data or confidential materials into an AI model unless you're using a secure, local instance and have IRB approval.

3. **Use outputs as drafts.** AI-generated content should be treated as a first pass—not necessarily a final product.

4. **Disclose use of AI.** Be transparent about how you used LLMs in your work. For example, this guide was drafted with the help of ChatGPT. Some journals now require disclosure of AI use (see APA's policy on generative AI and this review on the ethics of AI disclosure in scholarly work).

5. **Respect evolving norms.** Standards for using LLMs in science are still taking shape. This document will be updated as norms evolve, and I welcome your feedback as part of that process (you can reach me at sampratt@g.ucla.edu or on twitter https://x.com/sampratt99).

Table of Contents

Idea Generation & Literature Review 4

- Elicit: The AI Research Assistant 4
- ReadCube Papers: The AI Reference Manager 5

Stimuli, Item, & Survey Development 6

- PsyItemGenerator 6
- Psychology Vignette Creator 7

Data Analysis and Statistical Planning 8

- Data Analyst (by ChatGPT) 8
- Using GPT for Psychological Text Analysis 9

Writing and Manuscript Preparation 10

- Psychology Peer Reviewer 10

Idea Generation & Literature Review

Elicit: The AI Research Assistant

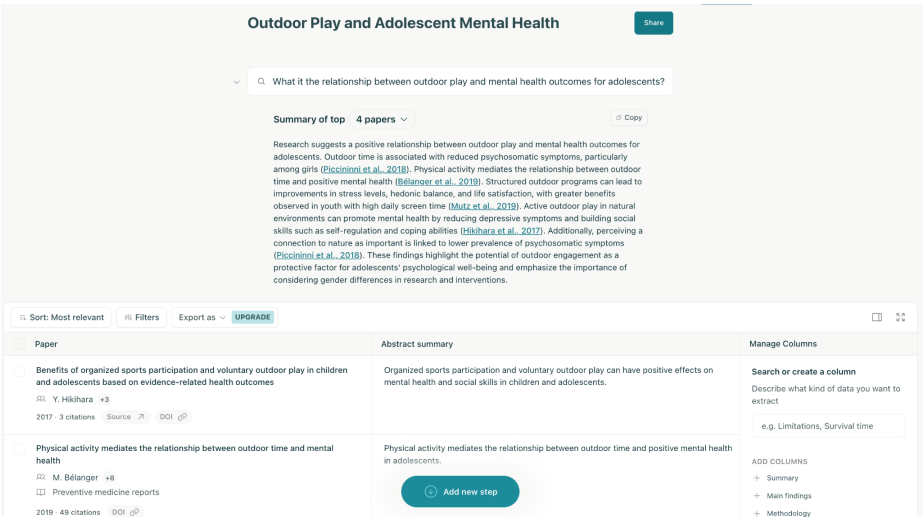
Purpose: Finds academic papers and generates structured literature reviews when prompted with a research question. Elicit uses a corpus of over 125 million open and closed access papers from Semantic Scholar. It was co-founded by a Stanford cognitive scientist and rolls out regular updates with new functions.

Link: https://elicit.com

Example Use Cases (see this guide for more):

- Find papers that answer a research question
- Summarize findings across multiple papers
- Generate a structured literature review as a downloadable PDF
- Filter papers by specific criteria (e.g., experiments, sample size, outcome)

Example Prompt:



Important Consideration:

- Avoid prompting GPT to “summarize this paper” or “find a paper that shows…” when conducting a literature review. A recent study finds that LLMs (like GPT) tend to overgeneralize or misrepresent findings when prompted directly. Tools like Elicit are better suited for this task because they extract structured information directly from real papers.

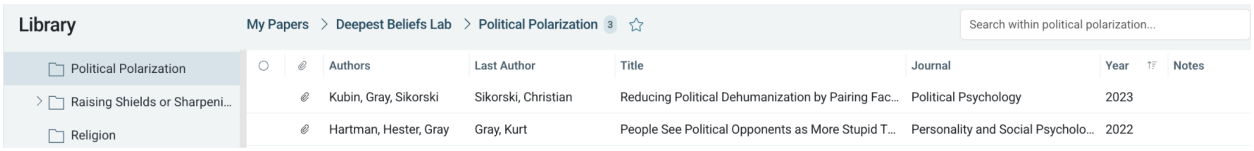
ReadCube Papers: The AI Reference Manager

Purpose: Online reference manager that includes an AI assistant for interacting with papers.

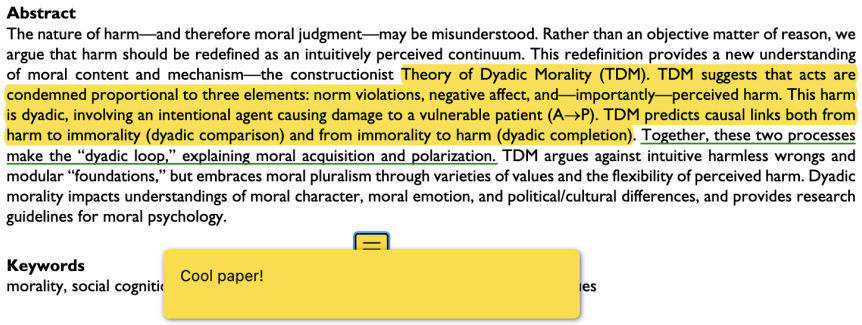
Link: https://www.papersapp.com/ ($39 per year with academic discount)

Example Use Cases

- Save PDFs and organize into project folders



- Annotate PDFs



- Chat with PDFs using the AI assistant - it cites any information it provides you directly from the paper.

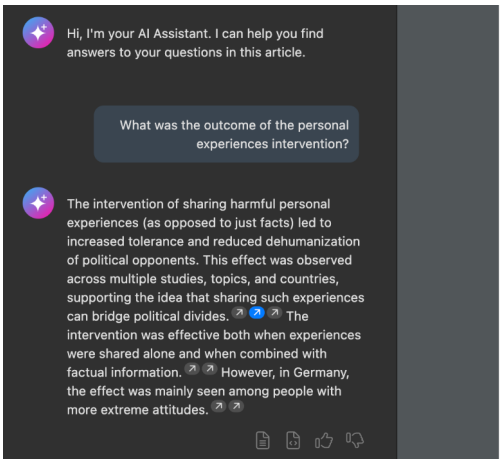




Figure 1. Study 1 mean differences by condition. ***$p < .001$, **$p < .01$.

Dehumanization Towards Opponent

Participants responded to an 8-item measure (Bastian et al., 2013), using a 7-point scale from 1(*not at all*) to 7 (*extremely*). This measure is frequently used to assess levels of dehumanization (e.g., Kteily et al., 2016; Kubin et al., 2022). Example item: "I felt like the person in the story lacked self-restraint, like an animal." Higher scores indicated greater dehumanization ($\alpha = .69$).[3]

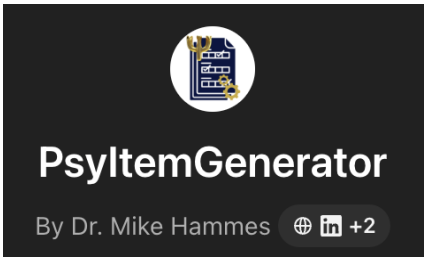
Results

Participants were more tolerant, and dehumanized less, when opponents based their attitudes on harmful personal experiences as compared to facts. See Figure 1.

Using PROCESS mediational analysis[4,5] (Hayes & Rockwood, 2020), we tested Hypotheses 1 and 2, which posited harmful personal experiences (vs. facts) would increases tolerance and reduce dehumanization. We did not pre-register these analyses but did pre-register them in Studies 2–4. Analyses supported both hypotheses and replicated the *outgroup-experience*

Stimuli, Item, & Survey Development

PsyItemGenerator



Purpose: A CustomGPT that generates items for psychometric instruments and discusses psychological constructs

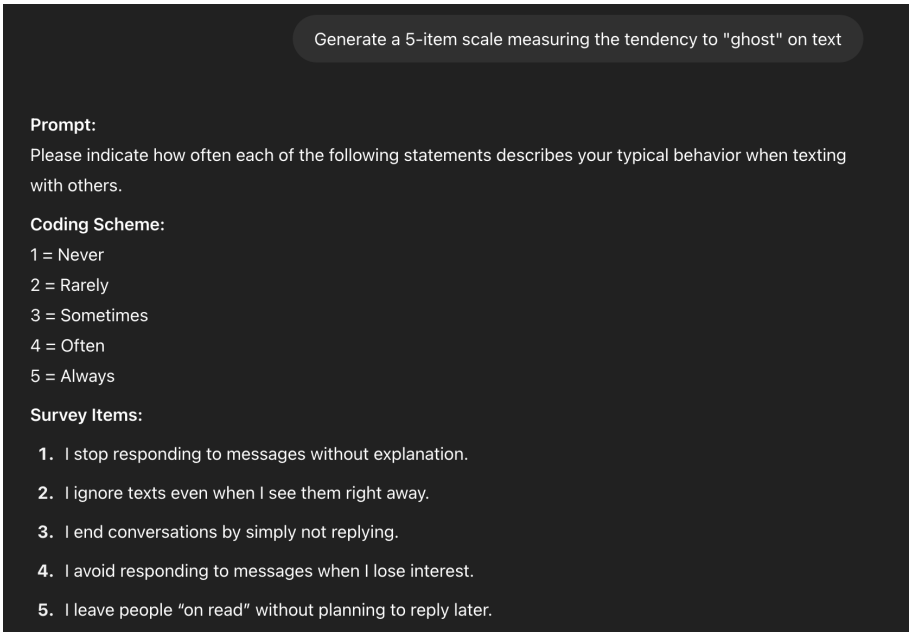
Link: https://chatgpt.com/g/g-jgnEb7vHp-psyitemgenerator

Example Use Cases:

- Brainstorming scale items for a new measure
- Learning about psychometrics

Example prompts:

- What does an item analysis do?
- What are common quality criteria for psychometric instruments?
- Generate a 5-item scale measuring empathy



Important Considerations: Always check for unintended confounds or overly leading language

Psychology Vignette Creator



Link: https://chatgpt.com/g/g-6737ad24a8388190b59988b7fe66f359-psychology-vignette-creator

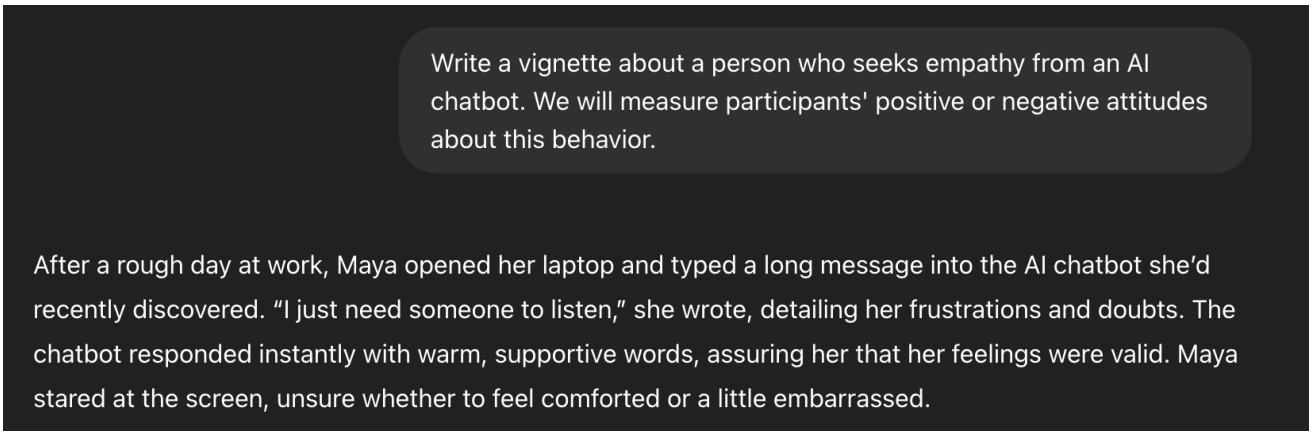
Purpose: A CustomGPT that creates short vignettes for psychology experiments.

Example Use Cases:

- Creating stimuli for experimental manipulations
- Generating condition-matched stories or scenarios

Example Prompts:

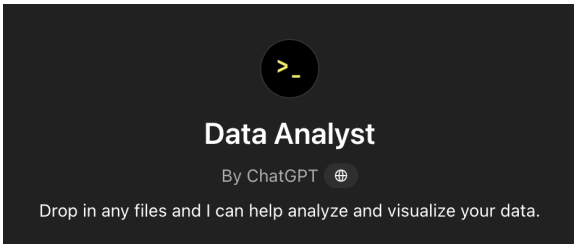
- “Create a scenario in which a person chooses whether or not to tell a lie, varying the social consequences.”
- “Generate three versions of this vignette: (1) a control condition, (2) a version in which the target experiences a positive emotion (happiness), and (3) a version in which the target experiences a negative emotion (sadness). Ensure that all other details remain consistent across conditions.”



Important Considerations: This GPT is designed to hold all aspects of a vignette constant except the key manipulation, but high-quality output may require multiple iterations and clear, specific prompts.

Data Analysis and Statistical Planning

Data Analyst (by ChatGPT)



Description: Helps with statistical planning, writing code, and interpreting code output.

Link: https://chatgpt.com/g/g-HMNcP6w7d-data-analyst

Example Use Cases:

- Troubleshooting R code for common psych analyses (e.g., t-tests, regression, ANOVA)
- Writing new code to perform a specific function
- Interpreting statistical output and ensuring accurate reporting
- Spotting logical inconsistencies or overlooked assumptions in models

Tips for Use:

- Treat GPT like a personal stats consultant throughout your workflow. Ask it to check your logic, code style, and interpretation on the fly.
- Paste your preregistration into the chat and ask GPT to flag any deviations as you go.

Example prompt: *"I'm analyzing data for an empirical psychology study using R. Below is my preregistration. Please help me write clear, accurate code, interpret the output when asked, and ensure that each step aligns with the preregistered analysis plan. Flag any deviations or potential issues as we go."*

Important Considerations:

- GPT can help write, troubleshoot, and explain code—but you are responsible for choosing valid methods and understanding the tests you're running.
- Don't upload raw participant data or any identifiable responses. GPT can assist without seeing the dataset.

Using GPT for Psychological Text Analysis

Purpose: GPT can help automate psychological text analysis tasks like coding open-ended responses, identifying emotions or moral themes in text, and standardizing qualitative data.

Key Insights from Research: A recent *PNAS* study by Rathje et al. (2024) found that GPT-4 can accurately detect psychological constructs (e.g., sentiment, discrete emotions, moral foundations) in text across 12 languages. It consistently outperformed traditional dictionary methods (e.g., LIWC, NRC) and approached the performance of fine-tuned machine learning models—all with simple prompts and no training data.

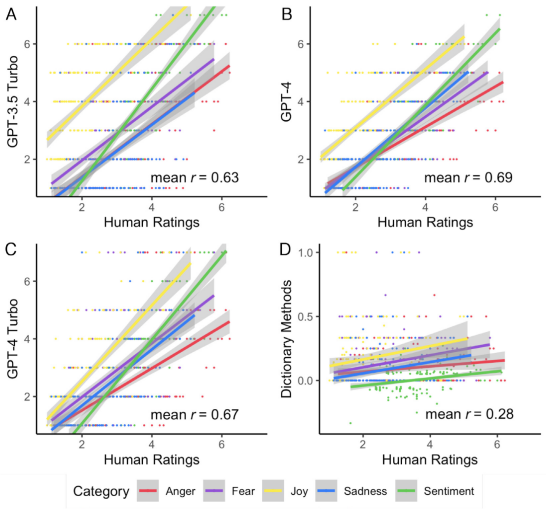


Fig. 1. Scatterplots showing correlations between human ratings and ratings predicted by different text analysis methods. (*A*) GPT-3.5 ratings; (*B*) GPT-4 ratings; (*C*) GPT-4 Turbo Ratings (the most recent model as of February 2024), and (*D*) ratings computed using dictionary methods (LIWC and NRC dictionaries with negation handling). Data are from 213 manually annotated headlines (measured on a Likert scale from 1 to 7). Each line represents a separate correlation between GPT output and manual annotators for a separate construct.

Example Use Cases:

- Classifying open-text survey responses into psychological themes
- Measuring sentiment or emotion in tweets, headlines, or open-ended responses
- Generating initial codebooks for qualitative coding

Tips for Use:

- Start by reading Rathje et al., 2024 and watching their short YouTube tutorial on using GPT for text analysis.

Important Considerations:

- LLMs are updated frequently and results can vary across model versions. Document the model and parameters you use for reproducibility.

Writing and Manuscript Preparation

Psychology Peer Reviewer



Purpose: This CustomGPT simulates an academic peer review process by generating three expert-level reviews of a psychology manuscript. Each review is detailed, critical, and focused on the paper's scientific rigor and fit for publication.

Link: https://chatgpt.com/g/g-682dee460fd881919e4f79748128ba0b-psychology-peer-reviewer

Example Use Cases:

- Getting pre-submission feedback on a manuscript draft
- Identifying weaknesses reviewers might flag
- Simulating peer reviews to guide revisions before resubmitting a paper
- Practicing how to respond to reviewer comments

Example Prompts:

- "Please review the uploaded manuscript as if submitting to *Psychological Science*. Provide three distinct reviews from experts in the topic area."
- "One reviewer should focus on theoretical framing, one on methodology, and one on interpretation of results."
- "Evaluate the manuscript's fit for *JESP*. Provide detailed comments on the analysis plan and provide specific and actionable suggestions for improvement."

Important Considerations:

- The feedback is only as good as the manuscript and prompt. Be specific about the target journal and your goals.
- Treat AI-generated reviews as helpful suggestions, not final verdicts. You still need human eyes on your work.
- Always verify that the suggestions align with your study's goals and theoretical framework.
- Peer reviews from this GPT are great for internal revisions, but should not be submitted as actual reviews or passed off as human-authored. *If you're interested in using LLMs to help write reviews, consider reading this guide posted by Nature.*