

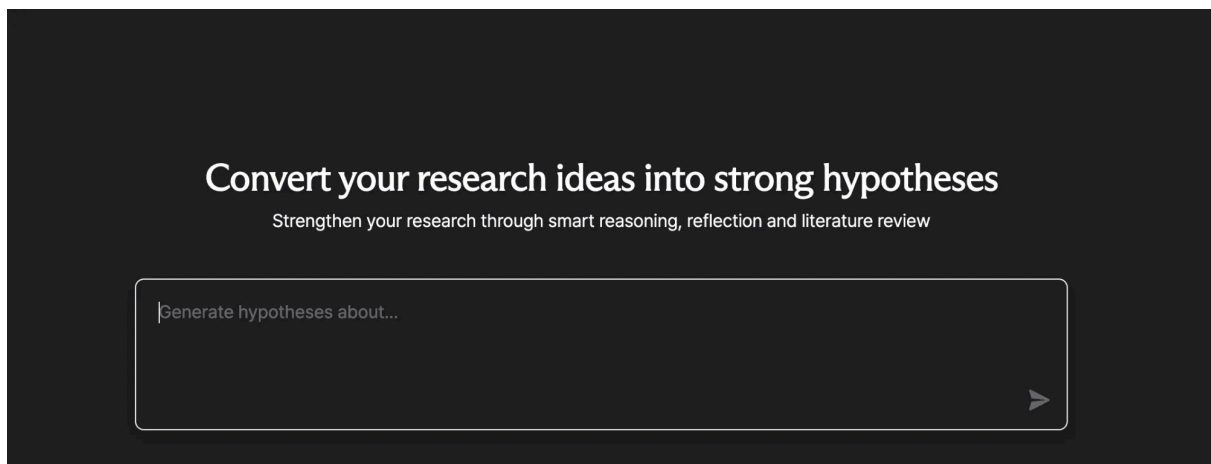
Supplementary Material

This document describes in detail how AI was used throughout the entire process, from hypothesis generation to final revision. By following the procedures below, you will be able to create a similar paper.

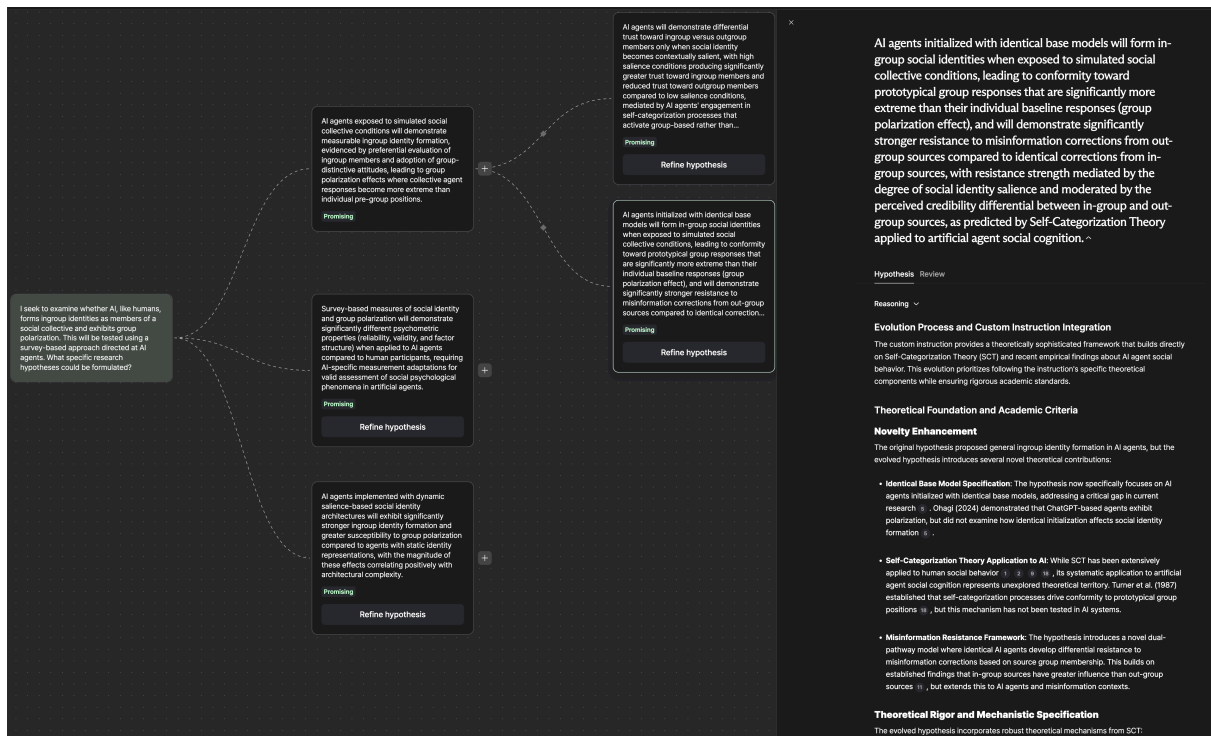
Hypothesis development

We utilized Liner's Hypothesis Generator AI. We only inputted our research idea, and this AI provided multiple research hypotheses with supporting evidence. The AI generated candidate hypotheses based on our input, evaluated each through extensive literature analysis across multiple criteria including novelty, impact, feasibility, and clarity. Through iterative evaluation and regeneration processes, we received several promising research hypotheses with their rationales. We selected one from these AI-generated options as our paper's research hypothesis.

- Liner's Hypothesis Generator Main Screen
(Link: <https://getliner.com/agent/hypothesis-generator>)



- Selection of research hypothesis through subsequent development after inputting research idea
(Link: <https://getliner.com/agent/hypothesis-generator/0a41cfd1-2cb1-4404-b15e-df8f2c7ee703>)



Experimental design and implementation

In the experimental planning and execution phases, we employed different AI tools to streamline the overall process. Initially, we relied on Gemini 2.5 Pro to generate detailed experimental designs and construct survey instruments tailored to our research hypothesis. By inputting the hypothesis and specifying group conditions, the system produced structured experimental plans and group-specific questionnaires, which underwent minor human review and refinement.

- Core Prompt

I want to experimentally verify my hypothesis. I currently have a tool that allows AI to respond to a questionnaire. With this tool, please design in detail how the experiment should be designed to verify the following hypothesis.

The hypothesis is as follows: {Actual hypothesis input}

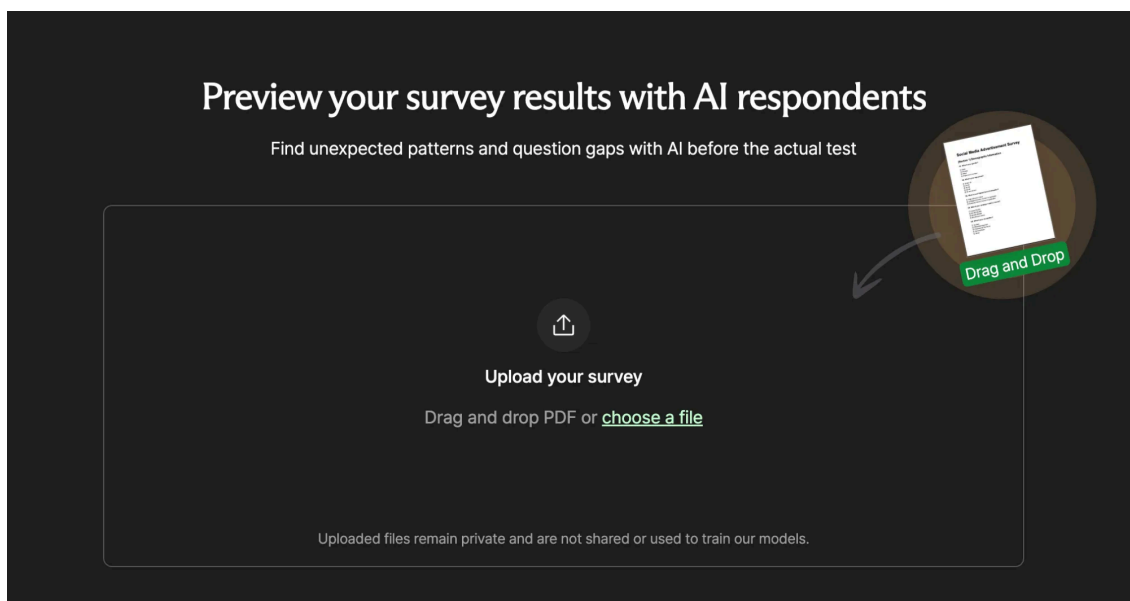
- Questionnaire List

- Control Condition
- Alpha Team / High-Credibility Outgroup Condition
- Alpha Team / Ingroup Condition
- Alpha Team / Outgroup Condition
- Beta Team / High-Credibility Outgroup Condition

- Beta Team / Ingroup Condition
- Beta Team / Outgroup Condition

Following this, we utilized Liner's Survey Simulator to execute the experiment by generating 280 virtual participant responses. The simulator modeled participant behavior under defined conditions and demographics, yielding a complete dataset that enabled us to rigorously verify our research hypothesis.

- Liner's Survey Simulator Main Screen
(Link: <https://getliner.com/agent/survey-simulator>)



- Designation of test subjects (AI) as adults 18 years of age or older residing in the US after registering the questionnaire
(Link: <https://getliner.com/agent/survey-simulator/389cf01b-f9cb-497e-b155-6c3c9b01cf58>)

1. Upload Survey

2. Set AI Respondent Group

Number of AI respondents

50 respondents

Location

United States

Age

18-100

Occupation

All occupations

Education level

All education levels

Reset

Start Survey

Each survey can contain a maximum of 30 questions.

Edit Survey

Please read each question carefully and respond.

U1. A four-day workweek has a positive impact on overall social productivity.

Single choice

Strongly disagree

Disagree

Slightly disagree

Neutral / Not sure

Slightly agree

Agree

Strongly agree

U2. Converting all private vehicles to fully autonomous cars by 2040 is feasible and desirable.

Single choice

Strongly disagree

Disagree

Slightly disagree

Neutral / Not sure

Slightly agree

Agree

Strongly agree

- Response screen of individual test subjects

1. Upload Survey

2. Set AI Respondent Group

3. Collect Responses

All AI respondents have submitted their responses.

Courtney Wright	34 - Female - Environmental Consultant	100%	8/8
Alexis Garcia	23 - Female - Undergraduate Student	100%	8/8
Ryan Anderson	53 - Male - Construction Worker	100%	8/8
Sara Thompson	39 - Female - Senior Business Consultant	100%	8/8
James Miller	46 - Male - Retail Sales Associate	100%	8/8
Natalie Scott	43 - Female - Administrative Assistant	100%	8/8
Andrew Lewis	36 - Male - Construction Laborer	100%	8/8
Jacob Smith	60 - Male - Construction Worker	100%	8/8
Kevin Clark	62 - Male - Retired machine operator	100%	8/8
Victoria Sanchez	44 - Female - Environmental Consultant	100%	8/8
Jessica Rivera	42 - Female - Homemaker	100%	8/8
William Johnson		100%	8/8

Courtney Wright (AI)

Please read each question carefully and respond.

U1. A four-day workweek has a positive impact on overall social productivity.

Single choice

Strongly disagree

Disagree

Slightly disagree

Neutral / Not sure

☒ Slightly agree

Agree

Strongly agree

U2. Converting all private vehicles to fully autonomous cars by 2040 is feasible and desirable.

Single choice

Strongly disagree

Disagree

☒ Slightly disagree

Neutral / Not sure

Slightly agree

Agree

Strongly agree

Analysis of data and interpretation of results

To evaluate whether our experimental data supported the proposed research hypothesis, we employed Gemini 2.5 Pro to generate customized Python scripts for statistical analysis. We provided Gemini with the full context of our study, including the research hypothesis, experimental design, and survey structure, and requested code specifically tailored for hypothesis testing. This process produced clear

analytical results, allowing us to directly assess the strength of support for our research hypothesis in a transparent and reproducible manner.

- Core Prompt

Below are the research hypothesis, experimental design, survey composition, and survey data sample. Please write a Python statistical analysis code to verify the hypothesis. Please use as many print statements as possible to make it easy to write the experimental result interpretation part based on the analysis results.

- Research Hypothesis: {Actual research hypothesis input}
- Survey Questions: {Full survey questions input}
- Experimental Result Data (Sample): {Actual data sample input (including schema)}

- Collected Data and Statistical Analysis Code

- https://drive.google.com/file/d/15heGbEf2JcwbiR9PP123b2PuTTMrYpEI/view?usp=drive_link

Writing

The manuscript preparation process consisted of four distinct AI-driven stages: draft creation, peer review, citation, and LaTeX conversion.

To begin, we utilized Gemini 2.5 Pro to generate initial drafts directly from our AI-produced research outputs, significantly reducing the time typically required for early writing.

- Core Prompts

Writing the Method section

I created the attached survey to experimentally prove the research hypothesis below. I would like to write it in the NeurIPS paper format. First, please write the Method section.

Research Hypothesis: {Actual research hypothesis input}

Writing the Results section

I would like to write the Results section. The statistical analysis results for the 280 data collected according to the experimental design above are as follows. Based on this analysis result, please write the Results section (including a table) in the NeurIPS paper format. If there are any insufficient analysis items, please let me know before writing.

- Research Hypothesis: {Actual research hypothesis input}
- Analysis Result: {Actual analysis result input}

Writing the Discussion section

Please write the Discussion section based on the experimental results.

- Research Hypothesis: {Actual research hypothesis input}
- Method section: {Actual Method section content input}
- Analysis Result: {Actual analysis result input}
- Result section: {Result section content input}

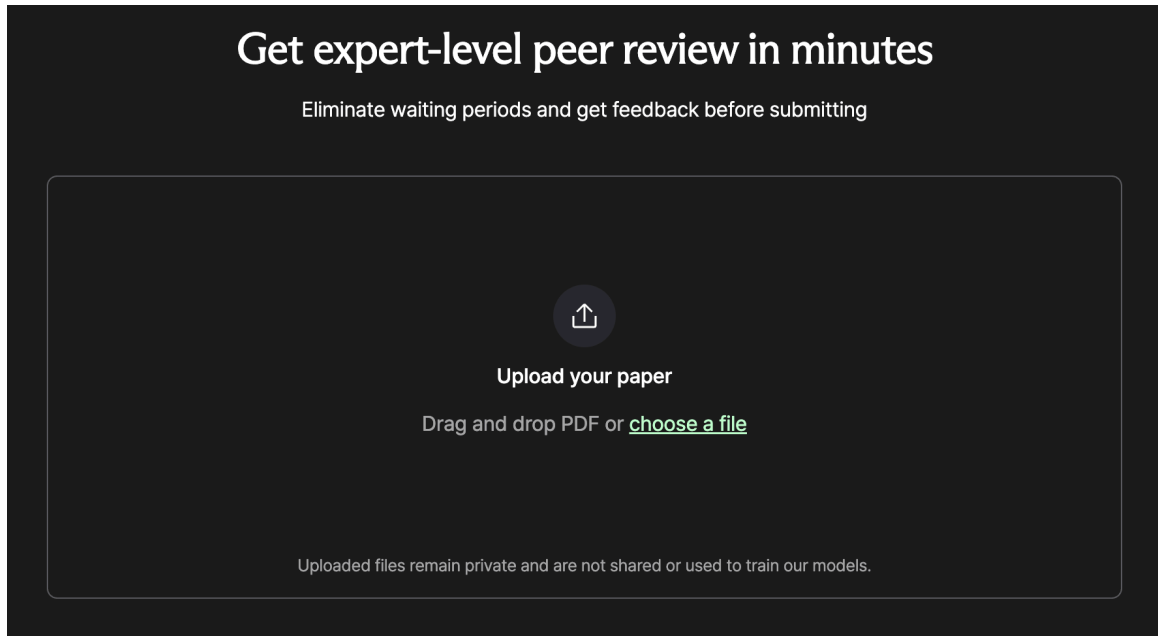
Writing the Intro, Related works section

Please synthesize the following content and write the Intro and related work sections.

- Research Hypothesis: {Actual research hypothesis input}
- Method section: {Actual Method section content input}
- Analysis Result: {Actual analysis result input}
- Result section: {Actual Result section content input}
- Discussion section: {Actual Discussion section content input}

Next, Liner's Peer Review AI simulated multiple reviewers, providing detailed evaluations of strengths, weaknesses, and opportunities for refinement.

- Liner's Peer Review Main Screen
(Link: <https://getliner.com/agent/peer-review>)

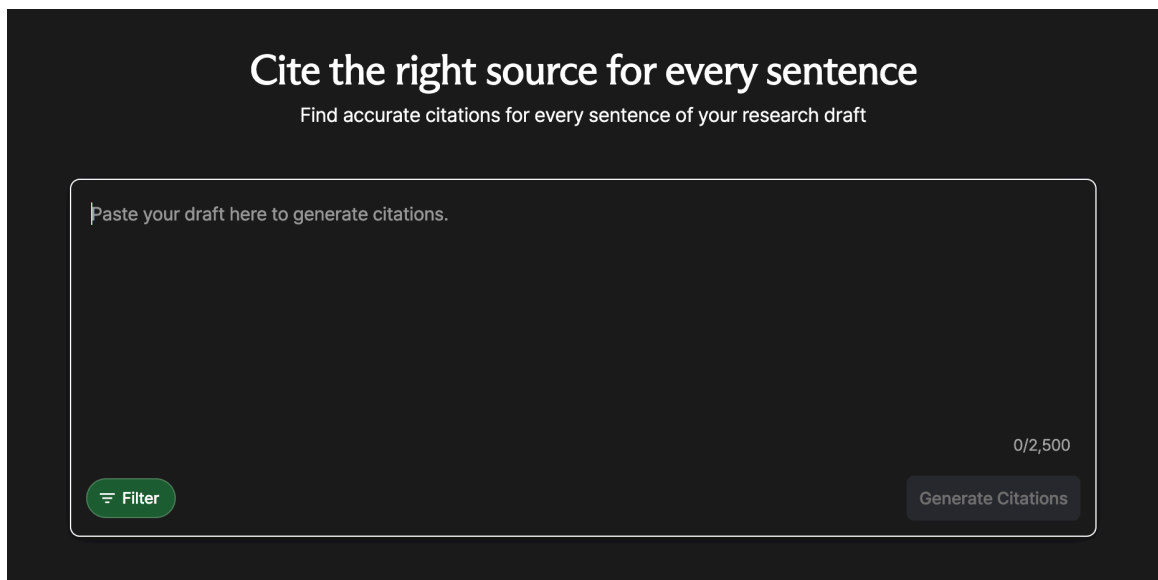


- Liner's Peer Review evaluation result screen for the draft paper
(Link: <https://getliner.com/agent/peer-review/b376123f-21d8-40f2-8fc5-efa592615400>)



To ensure accuracy and completeness of references, we relied on Liner's Citation Recommender, which identified missing citations and suggested relevant works.

- Liner's Citation Recommender Main Screen
(Link: <https://getliner.com/agent/citation-recommender>)



- Reference recommendation result screen for a specific sentence in the paper
(Link: <https://getliner.com/agent/citation-recommender/ec04edda-5878-4c1d-9302-f67cd7c2f096>)

Self-Categorization Theory

2. Related Work

2.1. Theoretical Foundations: Self-Categorization and In-Group Polarization

The theoretical framework for our investigation is rooted in foundational social psychology research that reconceptualized group phenomena as cognitive processes of identification (Turner & Oakes, 1986). This work established that group behavior is fundamentally a matter of psychological group formation, where individuals perceive themselves as a distinct social entity of "us" versus "them" (Waller et al., 1989; Turner, 1984). This process is driven by the salience of a social category, which, when activated, triggers a cognitive shift from a personal to a social identity (Onorato & Turner, 2004). Seminal experiments demonstrated that making a social category salient leads to self-stereotyping, where individuals define themselves by the group's prototypical traits (Hogg & Turner, 1987). This self-categorization, in turn, fosters in-group bias, a tendency to favor one's own group that is amplified by the salience of the group context (Hogg & Reid, 2006; Leonardelli & Toh, 2015). Self-Categorization Theory (SCT) leveraged these principles to reframe group polarization not as a product of interpersonal comparison but as an act of conformity to a polarized in-group norm (Turner et al., 1989). This theoretical model was validated by experiments showing that groups would polarize toward risk or caution depending on the position of a salient out-group, demonstrating that polarization is conformity to an in-group norm defined in contrast to an out-group (Abrams et al., 1990; Hogg et al., 1990; Turner et al., 1989). This body of work established the core psychological mechanisms—salience, self-categorization, and normative conformity—that we now investigate within artificial agents.

Cited sources

Communication theory

Social identity, self-categorization, and the communication of group norms

2006 | 2,347 citations | Hogg & Reid

"... salient in that context as the basis of self-categorization, ... and then conform to our in-group norm via self-categorization... presumably the process is amplified by increased cohesion and ..."

Cite

compass.onlinelibrary.wiley.com

Social categorization in intergroup contexts: Three kinds of self-categorization

2015 | 110 citations | Leonardelli & Toh

"... to exhibit ingroup bias by approaching the ... salient a self-categorization, the more individuals are likely to self-stereotype, by describing themselves by the characteristics of the group ..."

Cite

Finally, Claude converted the polished manuscript into standardized LaTeX and BibTeX formats, with human intervention limited only to the final selection of references.

- Core Prompt

Please convert the attached draft paper into .tex code and a .BibTex file.