Behavioral Experiment Simulation Tool – Methods Summary

Version 2.0 | Created by Dr. [Eugen Dimant](#)

Overview

The Behavioral Experiment Simulation Tool generates realistic synthetic pilot data for behavioral science experiments. It automates the creation of datasets that mirror real survey response patterns, enabling researchers and students to:

- Test analysis pipelines before data collection
- Practice data cleaning and exclusion procedures
- Verify survey logic and variable coding
- Develop analysis scripts on realistic data structures

What This Tool Does

1. Automatic Survey Parsing

Upload your Qualtrics QSF file and the tool automatically:

- Detects experimental conditions from survey flow
- Identifies Likert scales and their properties
- Extracts open-ended question specifications
- Infers factorial design structure

2. Behaviorally Realistic Response Generation

The simulator generates responses that reflect actual human survey behavior:

- **Response styles**: Engaged, satisficing, extreme responding, acquiescence
- **Attention patterns**: Realistic attention check pass/fail rates
- **Individual differences**: Trait-based variation across participants
- **Sentiment alignment**: Open-ended responses consistent with numeric ratings

3. Complete Output Package

Each simulation produces:

- `Simulated.csv` — Generated dataset
- `Metadata.json` — Simulation parameters and settings
- `Schema_Validation.json` — Data quality checks
- `R_Prepare_Data.R` — Analysis-ready R script
- `Column_Explainer.txt` — Variable descriptions
- `Instructor_Report.md` — Documentation for verification

Persona-Based Simulation Methodology

Theory-Grounded Response Styles

The simulator assigns behavioral personas based on established survey methodology literature:

Persona	Weight	Basis
Engaged Responder	30%	Careful, thoughtful responding
Satisficer	20%	Krosnick (1991) satisficing theory
Extreme Responder	8%	Greenleaf (1992) extreme response styles
Acquiescent Responder	7%	Yes-saying bias patterns
Careless Responder	5%	Inattentive response patterns

Domain-Specific Personas

The tool also includes specialized personas for research domains:

- **Consumer Behavior**: Brand loyalist, deal seeker, impulse buyer, conscious consumer
- **AI/Technology**: Tech enthusiast, tech skeptic, AI pragmatist, privacy-concerned
- **Behavioral Economics**: Loss-averse, present-biased, rational deliberator
- **Organizational Behavior**: High performer, disengaged employee
- **Social Psychology**: Prosocial, individualist, conformist

Personas are automatically selected based on keywords detected in your study description.

Trait-Based Individual Variation

Each simulated participant receives unique trait values drawn from persona-specific distributions:

- Attention level
- Response consistency
- Scale use breadth (endpoint vs. midpoint tendency)
- Acquiescence tendency
- Social desirability concern

This creates realistic individual differences without requiring manual configuration.

Open-Ended Response Generation

Open-ended text responses are generated to:

- Match the sentiment of numeric responses (positive, negative, neutral)
- Reflect persona-specific verbosity and engagement
- Include natural language variation (hedging, fillers, varying sentence structure)
- Align with the study topic and experimental conditions

The system uses extensive template libraries with natural language variation to produce responses that resemble authentic participant feedback.

Recommended Uploads

Required: Qualtrics QSF File

Export from Qualtrics: **Survey → Tools → Import/Export → Export Survey**

The QSF file provides:

- Question structure and types
- Survey flow and randomization
- Response option formats
- Condition/block assignments

Optional: Survey PDF Export

Export from Qualtrics: **Survey → Tools → Import/Export → Print Survey → Save as PDF**

The PDF provides:

- Improved question wording detection
- Better domain inference for persona selection
- Visual formatting context

Research Foundations

Core LLM Simulation Research

Foundational Papers:

1. **Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., & Wingate, D. (2023)**. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3), 337–351.

   - **DOI:** [10.1017/pan.2023.2](10.1017/pan.2023.2)
   - Demonstrates LLMs can replicate human response distributions across demographic subgroups with "algorithmic fidelity."

2. **Horton, J. J. (2023)**. Large language models as simulated economic agents: What can we learn from Homo Silicus? *NBER Working Paper* No. 31122.

   - **DOI:** [10.3386/w31122](10.3386/w31122)
   - Shows LLMs exhibit human-like economic behaviors including downward-sloping demand and status quo bias.

3. **Aher, G. V., Arriaga, R. I., & Kalai, A. T. (2023)**. Using large language models to simulate multiple humans and replicate human subject studies. *Proceedings of the 40th International Conference on Machine Learning (ICML)*, PMLR 202:337–371.

   - **Paper:** proceedings.mlr.press/v202/aher23a.html
   - Introduces "Turing Experiments" for validating LLM simulation of human behavior across classic studies.

High-Impact Validation Studies

4. **Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023)**. Generative Agents: Interactive Simulacra of Human Behavior. *ACM Symposium on User Interface Software and Technology (UIST)*.

   - **DOI:** [10.1145/3586183.3606763](10.1145/3586183.3606763)
   - Stanford/Google research demonstrating believable autonomous agent behavior in simulated environments.

5. **Binz, M. & Schulz, E. (2023)**. Using cognitive psychology to understand GPT-3. *Proceedings of the National Academy of Sciences (PNAS)*, 120(6), e2218523120.

    - **DOI:** 10.1073/pnas.2218523120
    - Systematic comparison showing LLMs exhibit human-like cognitive patterns across decision-making and reasoning tasks.

6. **Dillion, D., Tandon, N., Gu, Y., & Gray, K. (2023)**. Can AI language models replace human participants? *Trends in Cognitive Sciences*, 27, 597-600.

    - **DOI:** 10.1016/j.tics.2023.04.008
    - Found 0.95 correlation between GPT-3.5 and human moral judgments; reviews when LLMs can substitute for human participants.

Additional Resources

7. **Brand, J., Israeli, A., & Ngwe, D. (2023)**. Using GPT for Market Research. *Harvard Business School Working Paper* 23-062.
    - **Paper:** hbs.edu/ris/download.aspx?name=23-062.pdf
    - Demonstrates LLMs can generate realistic willingness-to-pay estimates and consumer preference data for under $100.

Standardization Features

Why Standardization Matters

When multiple teams simulate pilot data, standardized defaults ensure:

- Comparable datasets across teams
- Consistent exclusion criteria
- Reproducible simulation parameters
- Fair assessment of analysis approaches

Default Settings (Simple Mode)

Parameter	Default Value
Gender distribution	50% male
Mean age	35 years
Age SD	12 years

Parameter	Default Value
Attention check pass rate	95%
Random responder rate	5%
Min completion time	60 seconds
Max completion time	1800 seconds
Straight-line threshold	10 consecutive identical responses

Advanced mode allows customization of all parameters.

Exclusion Criteria Simulation

The tool simulates realistic data quality issues and flags participants for potential exclusion:

Flag	Description
`Flag_Speed`	Completion time outside acceptable range
`Flag_Attention`	Failed attention check(s)
`Flag_StraightLine`	Excessive straight-line responding
`Exclude_Recommended`	Recommended for exclusion (any flag)

These flags allow students to practice applying pre-registered exclusion criteria.

Reproducibility

- Each simulation run receives a unique `RUN_ID` and `SIMULATION_SEED`
- Seeds are derived from study metadata plus timestamp for uniqueness
- Providing the same seed reproduces identical output
- Metadata JSON captures all parameters for replication

Technical Requirements

- **Python 3.9+**
- **Dependencies**: streamlit, pandas, numpy, reportlab, sendgrid, pypdf

Citation

If you use this tool in your research or teaching, please cite:

> Dimant, E. (2025). Behavioral Experiment Simulation Tool (Version 2.0).

Contact

For questions, issues, or feature requests:

- Website: https://eugendimant.github.io/
- Email: edimant@sas.upenn.edu