

# Behavioral Experiment Simulation Tool

---

## Proprietary Software by Dr. Eugen Dimant

Generate publication-ready synthetic data for behavioral science experiments

---

## Overview

The Behavioral Experiment Simulation Tool is a sophisticated platform that generates high-fidelity synthetic datasets for behavioral science research. Whether you're piloting a new experiment, preparing analysis code, or teaching research methods, this tool produces data that mirrors what you would collect from actual human participants.

Unlike simple random number generators, this system applies decades of survey methodology research to create responses that exhibit realistic psychological properties—including response styles, attention patterns, and treatment effects calibrated to your specifications.

---

## Who Benefits from This Tool?

### Academic Researchers

- **Pre-registration preparation:** Generate synthetic data to develop and test your analysis scripts before collecting real data
- **Power analysis:** Validate sample size calculations with data that matches your expected effect sizes
- **Grant applications:** Include preliminary analyses in proposals using realistic simulated data
- **Pilot testing:** Test your experimental design logic and identify issues before investing in data collection

### Graduate Students and Postdocs

- **Methods training:** Learn data analysis techniques with realistic datasets that have known properties
- **Thesis preparation:** Develop analysis pipelines while awaiting IRB approval
- **Replication planning:** Generate data matching published effect sizes to plan replication studies

### Course Instructors

- **Teaching materials:** Create datasets with specific statistical properties for classroom exercises
- **Exam development:** Generate realistic data for assessment questions
- **Student projects:** Provide synthetic data for students who cannot access real participant data

## Industry Researchers

- **A/B test planning:** Simulate experiment outcomes before deploying to production
- **UX research:** Generate survey response data for prototyping analysis dashboards
- **Stakeholder presentations:** Demonstrate analysis approaches with representative data

---

## How It Works

### Step 1: Upload Your Survey

Simply upload your Qualtrics survey file (.qsf format). The system automatically extracts:

- Experimental conditions from your BlockRandomizer or branch logic
- Dependent variables (Likert scales, sliders, matrices)
- Open-ended questions and their visibility rules
- Survey flow logic determining which questions each participant sees

### Step 2: Confirm Your Design

Review the detected experimental structure:

- Verify that conditions are correctly identified
- Confirm which scales represent your dependent variables
- Adjust factorial design specifications if needed

### Step 3: Configure Parameters

Specify your simulation parameters:

- **Sample size:** How many participants to generate
- **Effect size:** Target Cohen's d for treatment effects (0.2 = small, 0.5 = medium, 0.8 = large)
- **Data quality:** Proportion of attentive vs. careless responders

### Step 4: Generate and Download

The system produces a publication-ready CSV file containing:

- Participant IDs and condition assignments
- Likert scale responses with realistic distributions

- Unique open-ended text responses
  - Demographics and metadata
  - Quality metrics and validation flags
- 

## The Science Behind the Simulation

### Persona-Based Response Generation

Rather than generating random responses, the system assigns each simulated participant a "persona" based on survey methodology research. These personas reflect actual patterns observed in human respondents:

**Engaged Responders:** High attention, thoughtful responses, full scale use. Based on Krosnick's (1991) "optimizers" who invest cognitive effort in providing accurate answers.

**Satisficers:** Lower effort responses, tendency toward agreement, restricted scale range. Krosnick's research documented this common response strategy where participants provide acceptable rather than optimal answers.

**Extreme Responders:** Consistent use of scale endpoints. Greenleaf's (1992) work identified this stable response style that varies across individuals.

**Acquiescent Responders:** Strong agreement bias regardless of item content. Billiet & McClendon's (2000) studies documented this tendency to agree with statements.

**Careless Responders:** Low attention, random patterns. Meade & Craig's (2012) research characterized these inattentive participants common in online samples.

### Effect Size Calibration

Treatment effects are calibrated using Cohen's d, the standard measure in behavioral science:

$$d = (\text{Treatment Mean} - \text{Control Mean}) / \text{Pooled Standard Deviation}$$

When you specify  $d = 0.5$ , the system adjusts response distributions so that the mean difference between conditions matches your target. This is achieved through:

1. **Semantic parsing** of condition names to determine effect direction
2. **Graduated adjustments** applied at the individual response level
3. **Validation checks** confirming achieved effects match targets

### Scale Reliability Modeling

Multi-item scales exhibit realistic internal consistency (Cronbach's alpha) through a factor model approach:

```
Response = lambda * Common_Factor + sqrt(1 - lambda^2) * Unique_Error
```

Where  $\lambda$  (factor loading) is derived from the target reliability. Items measuring the same construct share common variance while retaining item-specific variation, producing alpha values typically ranging 0.75-0.90.

## Response Style Modeling

The system models several well-documented response styles:

**Acquiescence Bias:** Tendency to agree with statements, producing higher means on positively-worded items. Modeled as a per-participant offset.

**Extreme Response Style:** Tendency to use scale endpoints. Higher extremity = more responses at 1 or 7 (on 7-point scales).

**Social Desirability:** Inflation of socially favorable responses. Applied proportionally based on item content.

**Midpoint Avoidance:** Cultural variation in willingness to use neutral midpoint. East Asian samples typically show lower midpoint avoidance than Western samples.

## Survey Flow Logic

The system respects your experimental design by tracking which questions each participant would actually see:

- **Block-level conditions:** Participants only receive responses for their assigned condition's blocks
- **Display logic:** Questions with condition-specific visibility rules are handled appropriately
- **Factorial designs:** Crossed conditions (e.g., Al x Hedonic, Al x Utilitarian) are properly parsed

---

## Open-Ended Response Generation

Text responses are generated using domain-specific templates and word banks, ensuring:

1. **Uniqueness:** No two participants give identical responses
2. **Context-awareness:** Responses reference the experimental manipulation when appropriate
3. **Condition-specificity:** Only participants who would see a question receive a response
4. **Length variation:** Verbosity varies by persona and question type

## Example Responses

For a question "What did you think about the AI recommendations?":

#### **Treatment participant** (saw AI):

"I found the AI-generated recommendations to be helpful. The system seemed to capture my preferences well. I appreciated how it addressed my needs effectively."

#### **Control participant** (no AI):

(Empty – this participant didn't see this question)

---

## Data Quality Features

### Attention Check Simulation

Configurable proportion of participants fail attention checks, matching real-world rates. Failed attention checks are flagged for potential exclusion.

### Careless Response Detection

The system can identify (and optionally flag or exclude) simulated careless responses:

- **Straight-lining:** Same response repeated across items
- **Alternating patterns:** Systematic alternation (1-7-1-7)
- **Midpoint overuse:** Excessive neutral responses
- **Response time anomalies:** Unrealistically fast completion

### Validation Metrics

Generated datasets include quality metrics:

- Achieved effect sizes with confidence intervals
  - Condition balance verification
  - Missing data rates
  - Response distribution statistics
- 

## Technical Specifications

### Input Requirements

- Qualtrics Survey Format (.qsf) file
- Internet connection for web interface

## Output Format

- CSV file compatible with R, SPSS, Stata, Python
- Metadata JSON with simulation parameters
- Optional quality report

## Supported Question Types

Type	Example
Likert Scales	7-point agreement scales
Sliders	Visual analog scales (0-100)
Matrix Tables	Multi-item scales with shared options
Multiple Choice	Single selection questions
Text Entry	Open-ended responses
Numeric Input	Willingness to pay, quantities
Rank Order	Preference rankings
Heatmaps	Click coordinate data

## Supported Experimental Designs

- Between-subjects (2+ conditions)
- Factorial (2x2, 2x3, 3x3, etc.)
- Mixed designs with between and within factors
- Complex branching and skip logic

---

## Frequently Asked Questions

### Is this generating "fake data"?

No—this is **synthetic data generation**, a legitimate research methodology. The tool produces data with known statistical properties for specific purposes:

- Testing analysis code before real data collection
- Teaching data analysis with realistic datasets
- Power analysis and sample size planning
- Developing preprocessing pipelines

Synthetic data should never be misrepresented as real participant data in publications.

## How realistic are the responses?

The responses exhibit statistical properties matching published research on human survey behavior:

- Mean responses around 5.0–5.5 on 7-point scales (documented positive response bias)
- Standard deviations of 1.2–1.8 (typical for Likert data)
- Cronbach's alphas of 0.75–0.90 for multi-item scales
- Effect sizes within +/-0.15 of specified targets

## Can I use this for any survey?

The tool is optimized for behavioral science experiments with:

- Clear experimental conditions (treatment/control)
- Likert or similar scale DVs
- Defined effect size expectations

It may not be suitable for purely exploratory surveys or complex longitudinal designs.

## How are effect directions determined?

The system parses condition names to determine which should produce higher/lower responses:

- **Positive indicators:** "high", "treatment", "reward", "positive"
- **Negative indicators:** "low", "control", "loss", "negative"

For complex designs, effect direction can be specified manually.

---

## Getting Started

1. **Export your Qualtrics survey** as a .qsf file (Survey > Tools > Import/Export > Export Survey)
  2. **Access the tool** at the provided URL
  3. **Upload your .qsf file** and review the detected structure
  4. **Configure parameters** for your specific needs
  5. **Generate and download** your synthetic dataset
- 

## Citation

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

## Support and Contact

For questions, feature requests, or collaboration inquiries, please contact through official university channels.

---

*Version 1.0.0 | Proprietary Software | All Rights Reserved*

*Developed by Dr. Eugen Dimant*