

## Process Tracing for Dummies: Solutions for design, analysis and presentation

**Martijn C. Willemsen**

Eindhoven University of Technology

**Eric J. Johnson**

Columbia University

**Ulf Böckenholt**

McGill University

Research Supported by:  
NWO Veni Grant 451-02-097  
NSF Grant SES-0352062



### Process tracing

- What is process tracing?

*Process tracing is an experimental method used to capture process data during decision tasks*

- Definition of process data:

*Any data collected between the presentation of the stimuli and the final response given*

### Old School process tracing

- Verbal reports, manual retrieval (information boards)
- early computerized information boards: Mouselab for DOS
- Analysis done on ad-hoc indices
  - Aggregation over attributes/alternatives
  - Difficult to comprehend the multivariate nature of the process data
- Addresses but does not model individual differences



### New School Process tracing

- New technologies for capturing process
  - Clickstream data on the web
  - MouselabWEB
  - Eye movements (trackers)
  - fMRI
- New graphical representations (Icon Graphs)
- More sophisticated models (multilevel):
  - Allow for finer tests of theories
  - Allow for modeling of heterogeneity

### Goal of this talk

- Important recent developments make process data more useful
- Process data should be used more to enhance our theories and predictions
- Solutions for:
  - Design of process tracing experiments
  - Representations of process data
  - Statistical methods for analysis of process data

## Design: MouselabWEB

- Goal: perform Mouselab-like process tracing experiments on the web (and in the lab)
- Approach: simple HTML/javascript available in recent browsers (works in 96%+ of browser usage)
  - Operating System Independent
  - No network delays: (Client-side, 1/60th second precision)
  - Fast and Easy: No plug-ins, small pages
  - No hassle server-side scripting (php/mysql)
  - Easily extended: Open source (GNU license)

## Features of MouselabWEB

- <http://www.mouselabweb.org/>
- **Designer** program to design pages with mouselabWEB and other questions
- **Datalyser** program to retrieve and replay a movie of the process data
- Web-based means:
  - Large numbers of respondents
  - A lot of heterogeneity in participants (not quite the average 20 year old student lab participant)
  - Specificity of respondents: targeting specific groups

## Asian Disease

**Asian Disease**

Imagine that the US is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

<b>Program A</b>	200 people will be saved	probability = 1 (100%)
<b>Program B</b>	600 people will be saved	probability = 1/3 (33.3%)
	0 people will be saved	probability = 2/3 (66.7%)

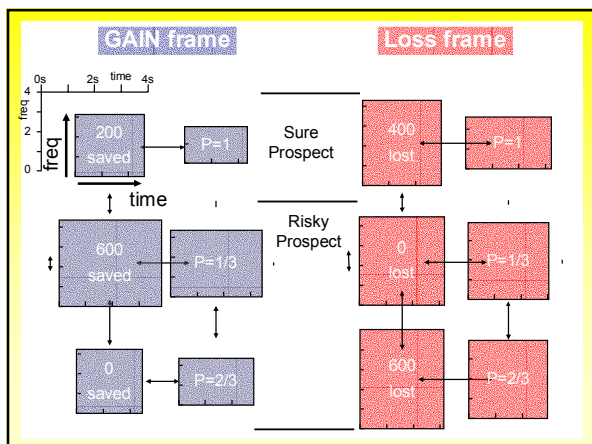
1. Which of the two programs would you favor?

Program A  
 Program B

Show data   Play data   Play Loss data

## Representation: Icon Graphs

- How to represent process data?
- Multivariate nature of process measures
  - **Attention:** acquisition frequency and looking time
  - **Search patterns:** direction and variability of search
  - **Time dynamics:** attention and search over time
- Icon Graphs can be used to display all this in one graph, using a 2D display.



## Methods: Data structure

- Many repeated observations within a participant
- Data structures are rich:
  - Attributes/alternatives
  - Time per acquisition
  - Time dependencies between acquisitions
  - Transitions between acquisitions
- Data are clustered:
  - First level: individual observations
  - Second level: design and participant variables (including individual characteristics and individual-difference measures)

## Multilevel models

- Multilevel models allow investigating relations between lower-level predictors (e.g., attention to an option) and higher-level predictors (e.g., a measure of loss aversion on the participant level).
- Individual-level parameters (e.g., intercept, slopes, transition probabilities) become dependent variables:
  - $\beta_{it} = \text{Linkfunction}(\eta_0 + \eta_1 z_i)$

## Advantages of Multilevel models

- Flexible decomposition of process measures
- Tests of random and fixed effects of individual-difference variables on decision process
- Incorporation of 'nuisance' effects (e.g., spatial dependencies in information display, reading order).
- Works with unbalanced data

## Another example

- Reference dependence (Tversky and Kahneman, 1991)

## Reference Dependence

## Example of Multilevel analysis

- Model for Reference dependence
  - Dependent variable: amount of attention to a box (frequency or time)
  - 48 observations (6 boxes x 4 quarters x 2 trials) per participant
  - Model some of the error variance by using random effects
  - Dynamics over time are modeled using a linear and quadratic terms
  - Choice is included in the model (to examine differences in process between A and B choosers)

## Results

- Attention changes dynamically over time from the reference option towards to chosen option
- Losses do not loom larger than gains

## Results

- Random factors (capture individual variance on these factors) are significant:
  - Intercept
  - trial
  - attention to loss
- Process data also predicts choice
- Multilevel model can be used to distinguish between different theoretical predictions about the process

## Summary

- Synergy of new techniques, new representations, and new analysis methods
  - Sophisticated data collection on the internet increases diversity in samples
  - New graphical representations (e.g. Icon Graphs) help identifying what factors and differences are relevant
  - New multilevel analysis methods allow for actually modeling these effects, including individual differences

## Summary

- Process data can tell why effects occur or do not occur and how individual differences might mediate this.
- Process data increases insight into cognitive processes underlying decision making
- Use process data yourself!

<http://www.mouselabweb.org/>

*Process Tracing*  
FOR  
DUMMIES<sup>®</sup>  
~~FOR~~  
*Geniuses*