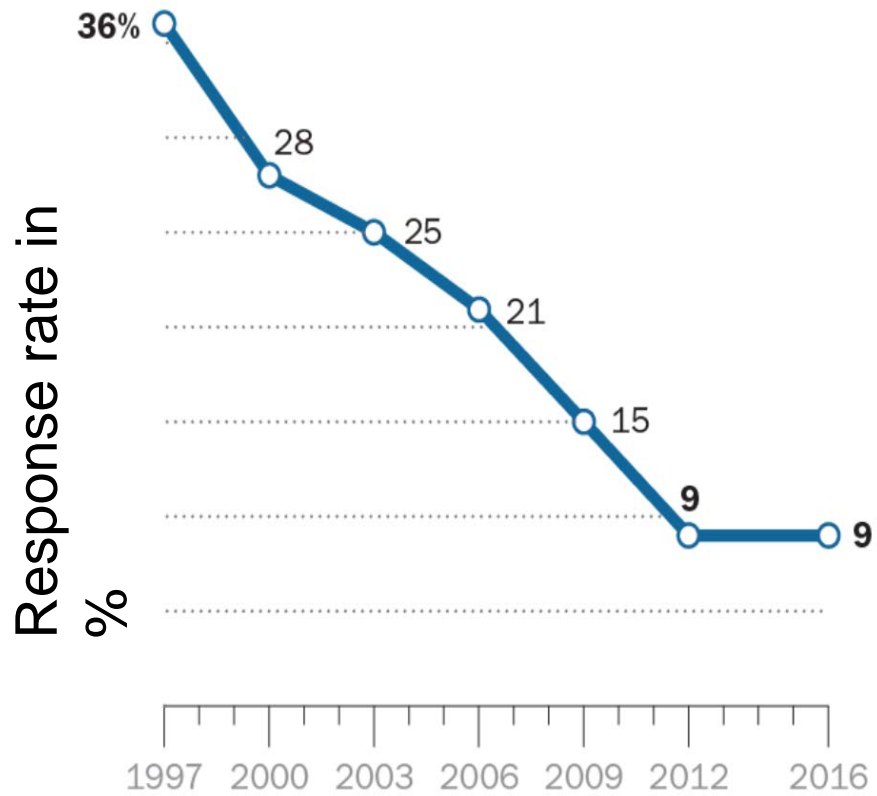


Using Paradata to Evaluate Online Performance

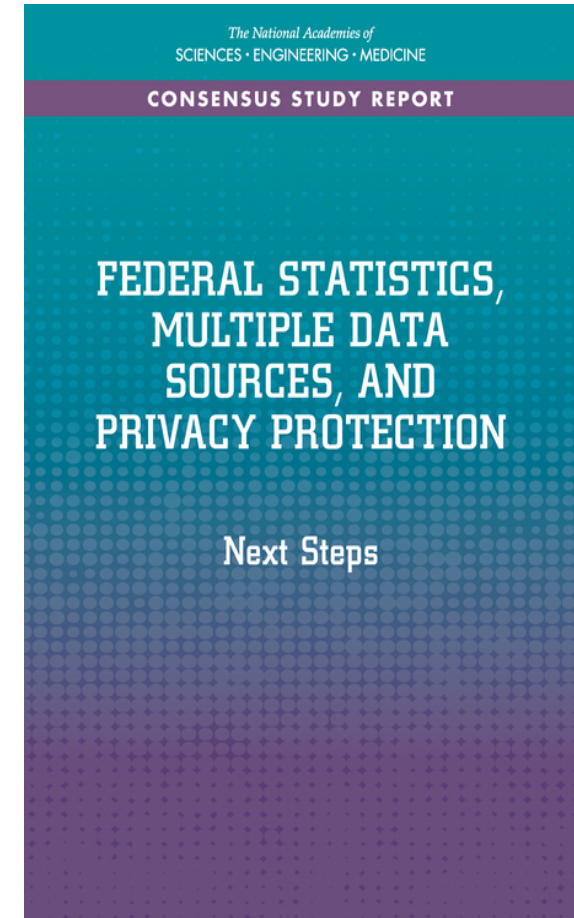
Frauke Kreuter – Beyond Results, 6/17/2020

Acknowledgement: DFG (GR 3793/2-1 and KR 2211/5-1)
Sonja Greven, Amanda Fernández-Fontelo, Felix Henninger, Pascal Kieslich,

Background: Concern in Survey Research

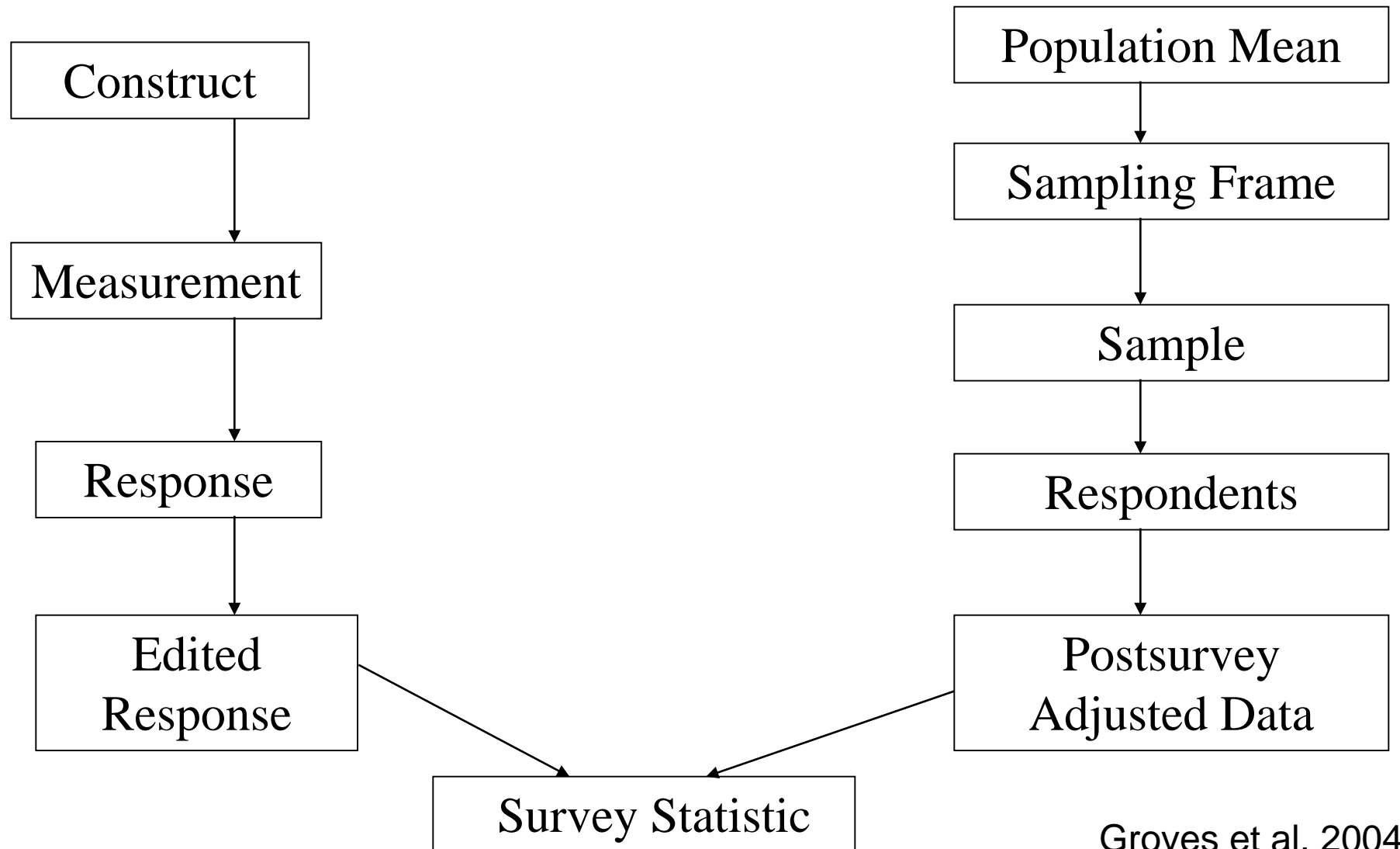


Source: Pew Research Center surveys conducted 1997 to 2016. "What Low Response Rates Mean for Telephone Surveys"

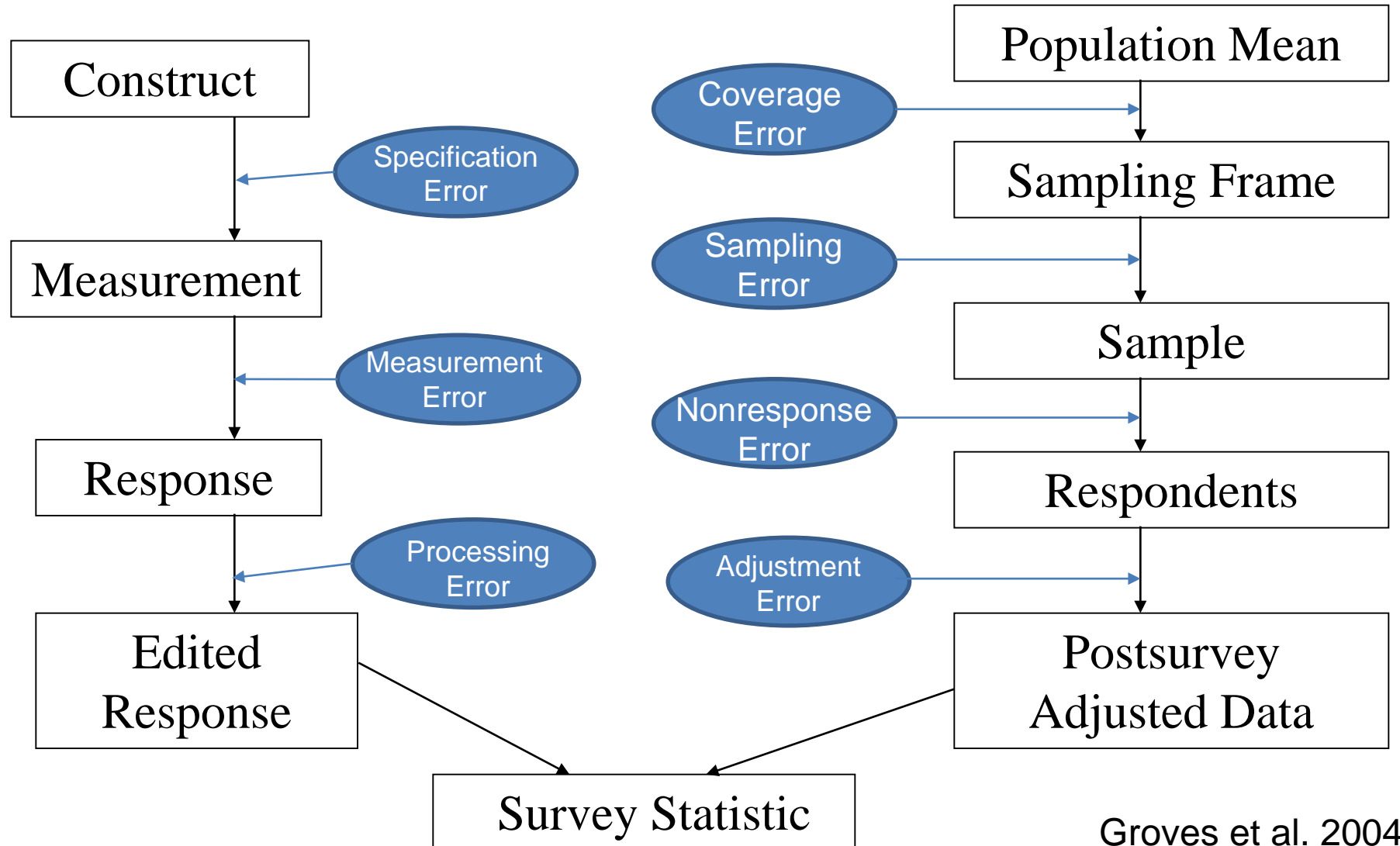
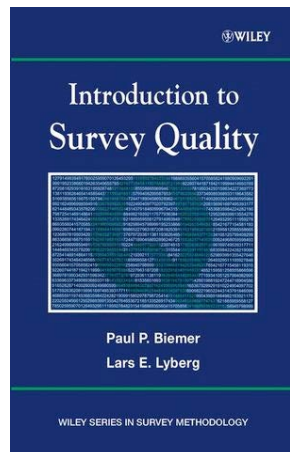
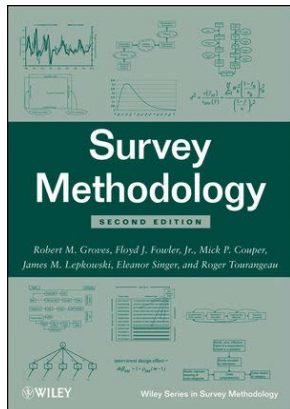


ERRORS IN SURVEYS

Data Generating Process

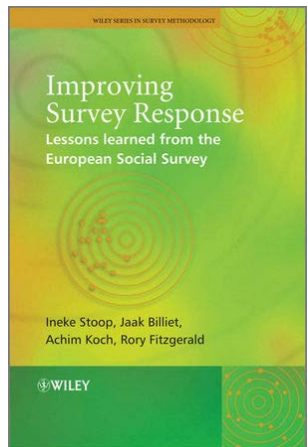
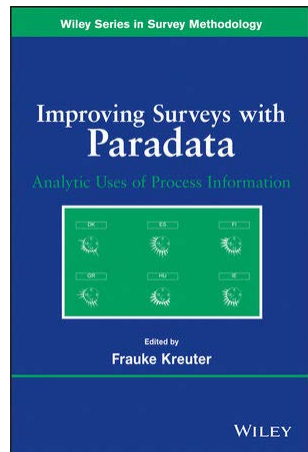
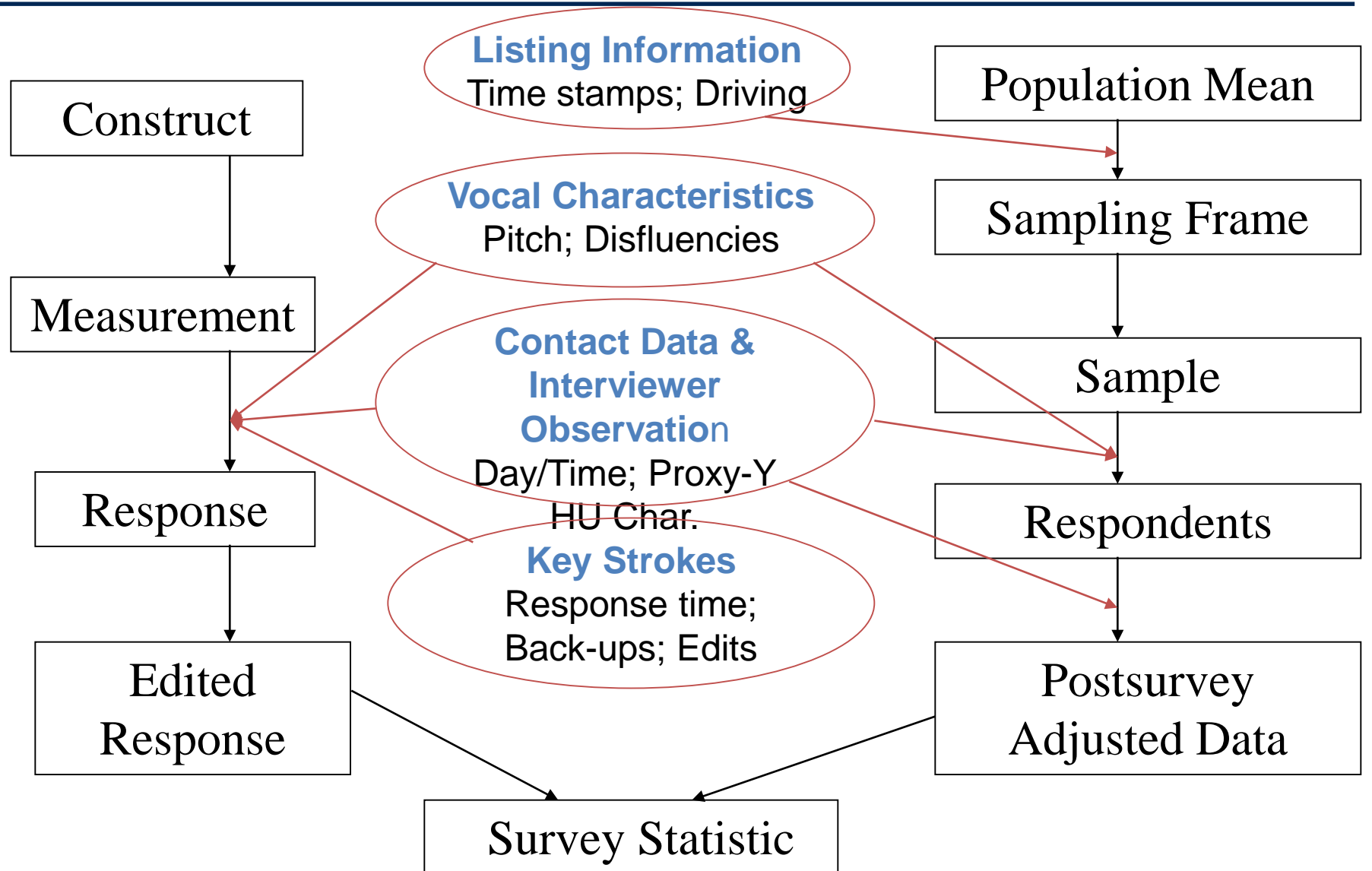


Data Generating Process

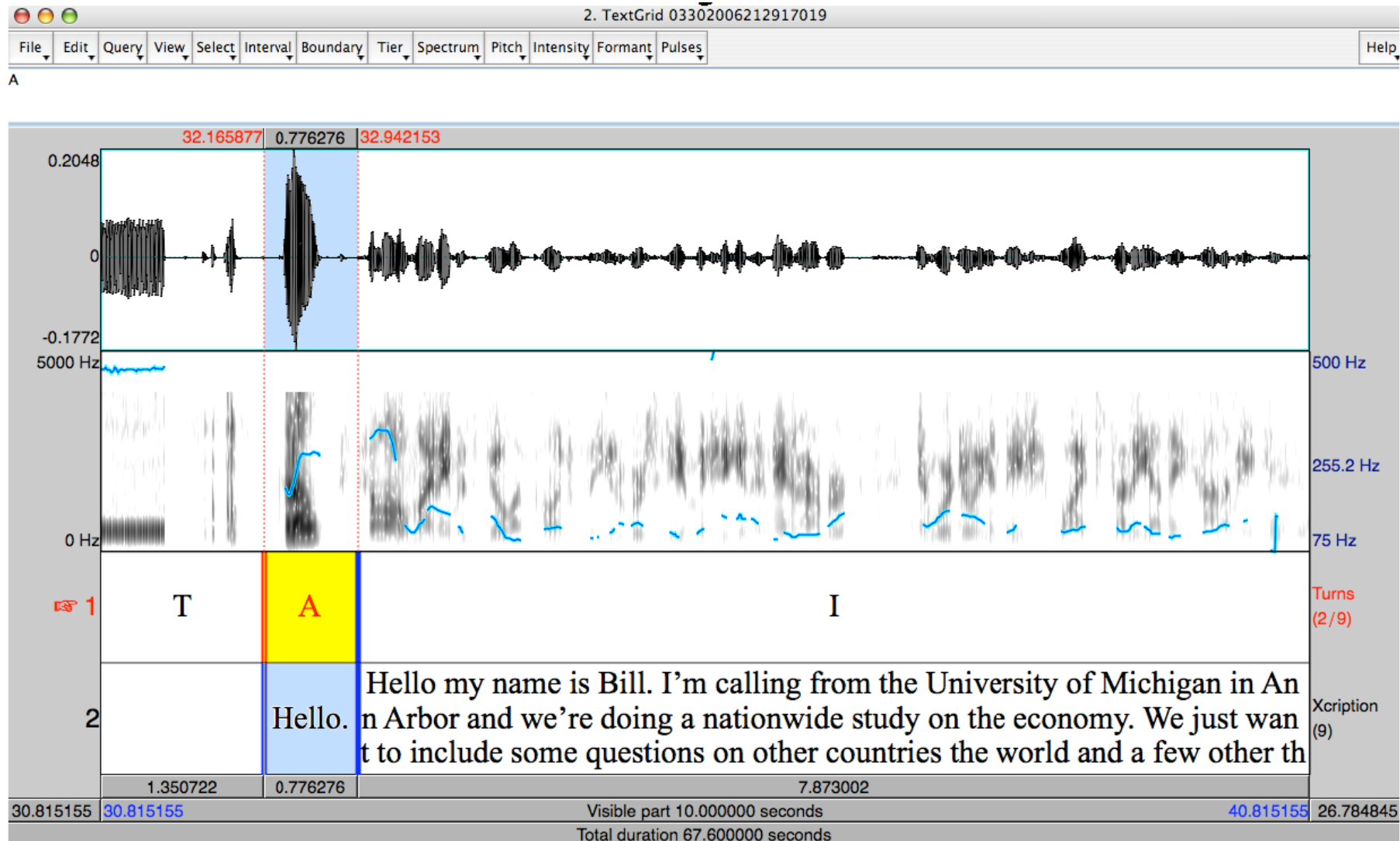


PARADATA

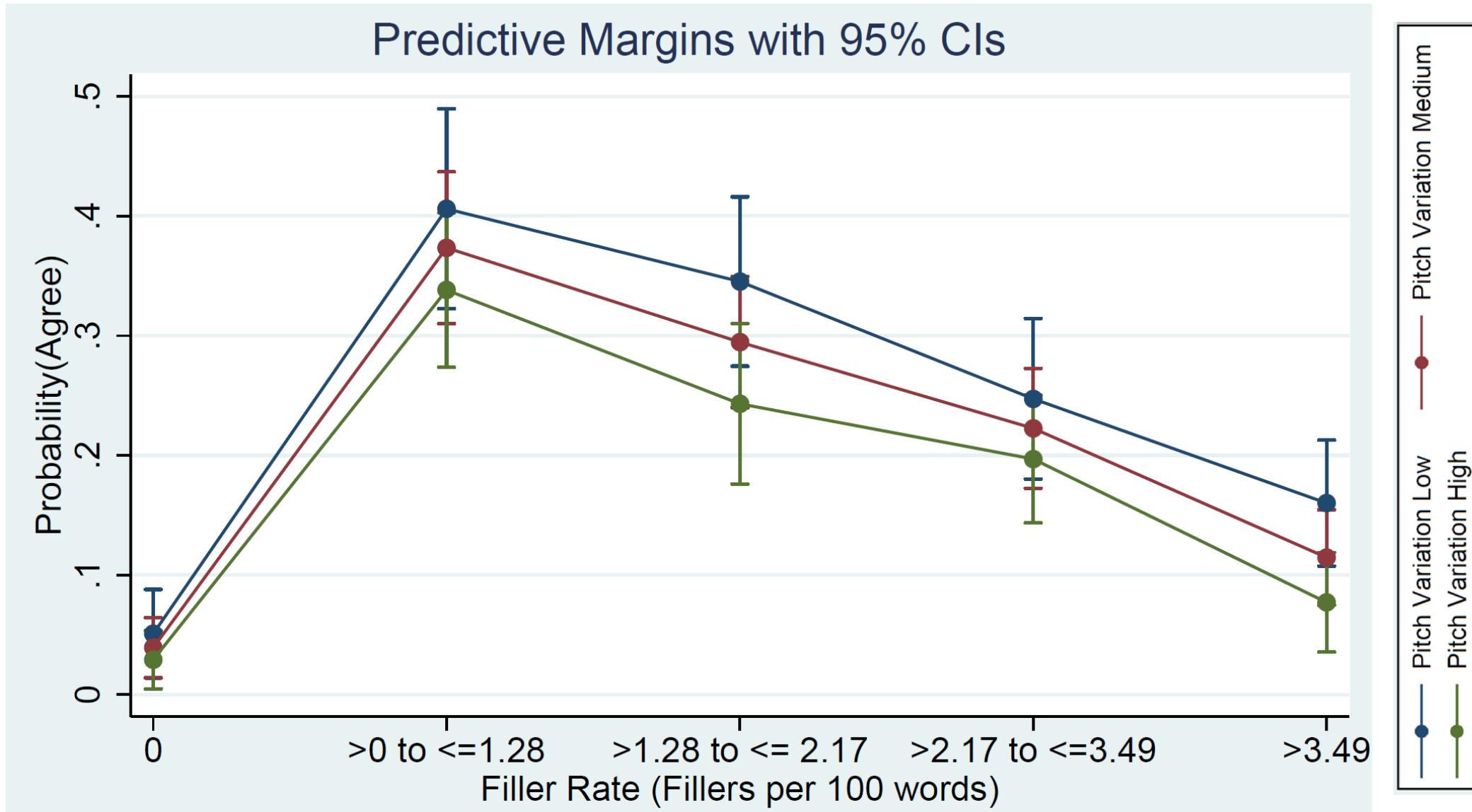
Data Generating Process



Vocal Characteristics -> Nonresponse



Vocal Characteristics -> Nonresponse

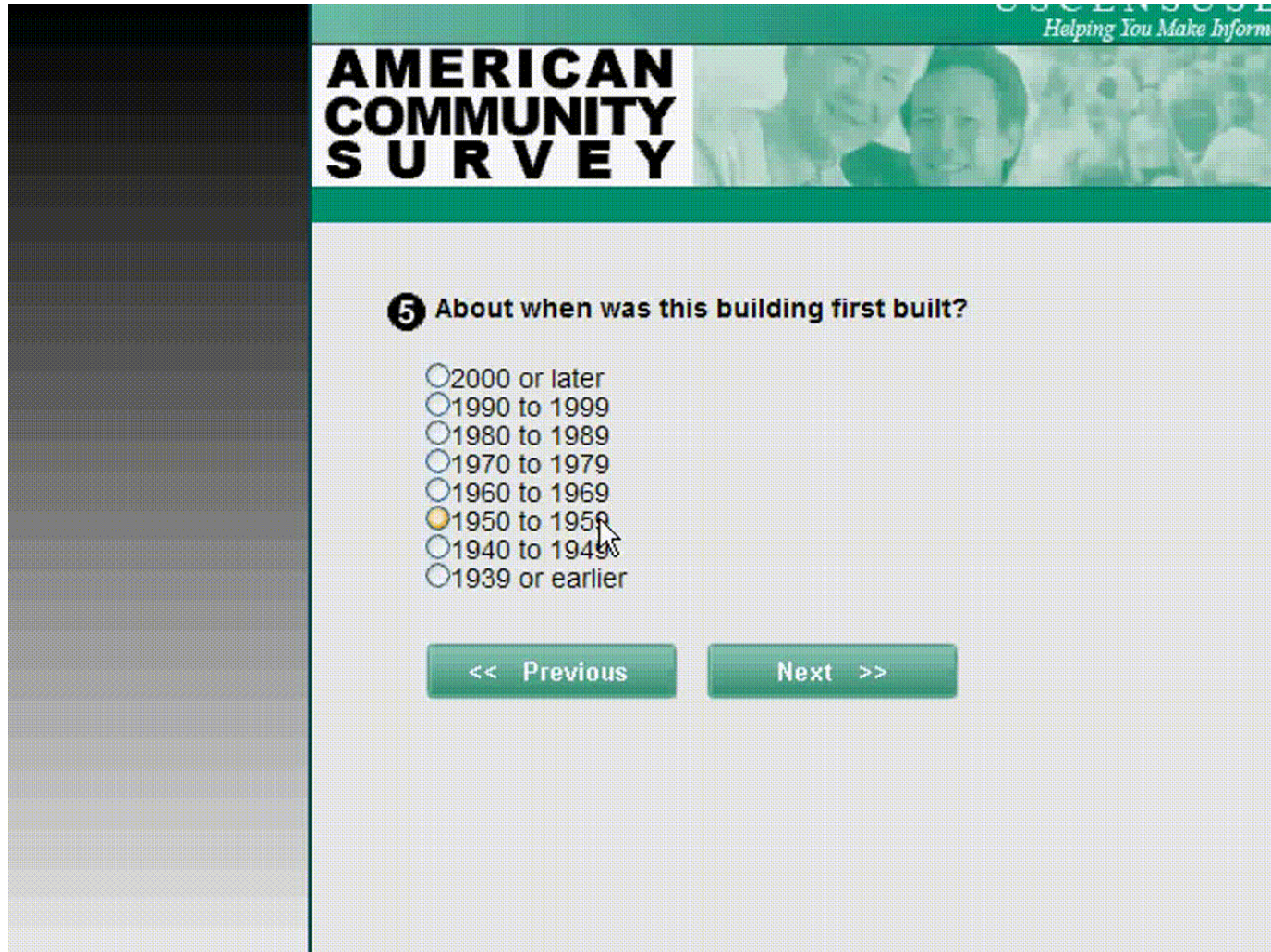


MOUSE TRACKING

Mouse Tracking

- Mouse movements are “real-time traces of the mind”
(Freeman et al., 2011)
- Across disciplines, researchers look at similar patterns of movement to understand how users interact with a webpage
 - Trajectories, hovers, regressions
- Used to measure:
 - Uncertainty (Cox & Silvia, 2006; Zushi et al., 2012)
 - Cognitive conflict (Duran 2010; Freeman 2010)
 - Interest (Mueller & Lockerd, 2001; Rodden et al., 2008)

Mouse Tracking - Example



US CENSUS BUREAU
Helping You Make Informed Decisions

AMERICAN COMMUNITY SURVEY

5 About when was this building first built?

- 2000 or later
- 1990 to 1999
- 1980 to 1989
- 1970 to 1979
- 1960 to 1969
- 1950 to 1959
- 1940 to 1949
- 1939 or earlier

<< Previous Next >>


Mouse Tracking in Web Surveys

- Mouse tracking in survey research has focused on:
 - Total distance traveled and its relationship to data quality (Steiger and Reips, 2010)
 - How specific movements are related to respondent difficulty (Horwitz et al., 2016)
- Gaps
 - Bringing patterns and applications to a survey context
 - Finding meaningful indicators
 - Predictive link between specific cognitive processes and movements
- RQ: Are specific types of mouse movement associated with specific types of difficulty when answering survey questions?

Papers summarized for this keynote

- Horwitz, R., Brockhaus, S., Henninger, F., Kieslich, P. J., Schierholz, M., Keusch, F., & Kreuter, F. (2019). Learning from mouse movements: Improving questionnaire and respondents' user experience through passive data collection. In P. C. Bearry, A. Wilmot, D. Collins, L. Kaye, J. L. Padilla, & G. Willis (Eds.), *Advances in Questionnaire Design, Development, Evaluation and Testing* (pp. 403–426). Wiley.
<https://doi.org/10.1002/9781119263685.ch16>
- Horwitz, R., Kreuter, F., & Conrad, F. (2017). Using mouse movements to predict web survey response difficulty. *Social Science Computer Review*, 35(3), 388–405.
<https://doi.org/10.1177/0894439315626360>

Study Design

- Sampling Frame: 1,627 *Rs* to a 2014-2015 survey conducted by Institute for Employment Research (IAB) who agreed to be contacted again
- Sample cases invited to web survey using 3 postal mailings and one email
 - First mailing included 5€ incentive
- September 5 – October 23, 2016
- Final response rate: 76.8% (AAPOR RR1)
- Removed from analysis:
 - Cases that did not use a mouse
 - Cases where mouse tracking never worked  n=853
 - Cases that took longer than 7 minutes to answer specific *Q*
 - Cases with item nonresponse on specific screens

Study Design

- 36 questions about opinions, employment, and demographics
- Experiments
 - Ordered vs. unordered response options
 - Check-all-that-apply vs. Yes/No grid
 - Straightforward vs. difficult response options
 - One answer vs. multiple answers possible
 - Not sensitive vs. sensitive question
- Web questionnaire programmed in SocSci Survey (Leiner, 2014)
- Mouse movements
 - Collected using JavaScript
 - Analyzed using R package mousetrap* (Kieslich et al., 2016)

*See also <https://github.com/PascalKieslich/mousetrap>

Analysis Measures

- Total distance traveled:
 - Sum of pixels covered by the mouse
- Regressions:
 - Any movement back and forth in the x- or y-direction
- Hovers:
 - Holding the mouse still for more than two seconds; frequency of however, and total time of hovers on a question

Examples

Total Distance Traveled

Sind Sie derzeit...

Falls Sie momentan in mehreren Beschäftigungsverhältnissen stehen, denken Sie bitte bei dieser und den folgenden Fragen an Ihre hauptsächliche Beschäftigung. Das ist diejenige Erwerbstätigkeit, in der Sie die meisten Stunden arbeiten.

- angestellt bei einem privaten, profitorientierten Unternehmen
- angestellt bei einem privaten, nicht profitorientierten Unternehmen
- angestellt im öffentlichen Dienst des Bundes
- angestellt im öffentlichen Dienst eines Bundeslandes
- angestellt im öffentlichen Dienst einer Kommune
- selbständig in einem freien Beruf
- selbständig in Handel, Gewerbe, Industrie, Dienstleistung, Landwirtschaft
- Freier Mitarbeiter/Freie Mitarbeiterin
- Mithelfende/r Familienangehörige/r

Weiter

Vertical Regression

Sind Sie derzeit...

Falls Sie momentan in mehreren Beschäftigungsverhältnissen stehen, denken Sie bitte bei dieser und den folgenden Fragen an Ihre hauptsächliche Beschäftigung. Das ist diejenige Erwerbstätigkeit, in der Sie die meisten Stunden arbeiten.

- angestellt bei einem privaten, profitorientierten Unternehmen oder einer Privatperson, das/die regelmäßig Löhne oder Gehälter auszahlt
- angestellt bei einem privaten, nicht profitorientierten Unternehmen, welches gemeinnützig ist oder von Steuervergünstigungen profitiert
- angestellt im öffentlichen Dienst des Bundes (Arbeitgeber ist die Bundesrepublik)
- angestellt im öffentlichen Dienst eines Bundeslandes (Arbeitgeber ist ein Bundesland)
- angestellt im öffentlichen Dienst einer Kommune (Arbeitgeber ist zum Beispiel eine Gemeinde, eine Stadt, ein Landkreis, ...)
- selbständig in einem freien Beruf (z.B. Arzt/Ärztin, Rechtsanwalt/Rechtsanwältin oder Architekt/in)
- selbständig, jedoch nicht in einem freien Beruf im Bereich Handel, Gewerbe, Industrie, Dienstleistung, Landwirtschaft
- Freier Mitarbeiter/Freie Mitarbeiterin, der aufgrund eines Dienst- oder Werkvertrags für ein Unternehmen Aufträge persönlich ausführt
- Mithelfende/r Familienangehörige/r ohne Bezahlung

Weiter

H1: Ordered vs. Unordered Response Options

Ordered Educational Attainment

UNIVERSITÄT
MANNHEIM

Institut für Arbeitsmarkt-
und Berufsforschung
Die Forschungseinrichtung der
Bundesagentur für Arbeit



Welchen höchsten allgemeinbildenden Schulabschluss haben Sie?

- Sonderschulabschluss, Abschluss der Förderschule
- Hauptschulabschluss, Volksschulabschluss
- Polytechnische Oberschule (POS) Abschluss 8. Klasse (DDR Abschluss)
- Polytechnische Oberschule (POS) Abschluss 10. Klasse (DDR Abschluss)
- Realschulabschluss, Mittlere Reife, Fachoberschulreife
- Fachhochschulreife (z.B. Abschluss einer Fachoberschule)
- Abitur / allgemeine oder fachgebundene Hochschulreife
- Abschluss einer Erweiterten Oberschule (EOS) oder Berufsausbildung mit Abitur (DDR-Abschlüsse)
- Anderer deutscher Schulabschluss, und zwar
- Ausländischer Schulabschluss und zwar
- kein Abschluss

Weiter

Unordered Educational Attainment

UNIVERSITÄT
MANNHEIM

Institut für Arbeitsmarkt-
und Berufsforschung
Die Forschungseinrichtung der
Bundesagentur für Arbeit



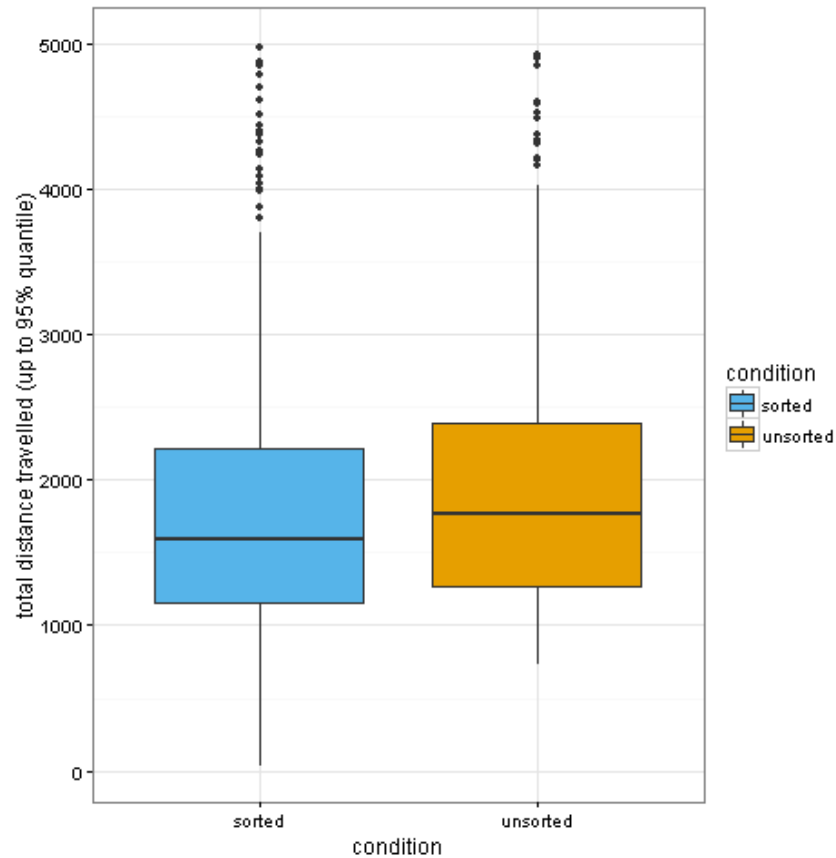
Welchen höchsten allgemeinbildenden Schulabschluss haben Sie?

- Abschluss einer Erweiterten Oberschule (EOS) oder Berufsausbildung mit Abitur (DDR-Abschlüsse)
- Sonderschulabschluss, Abschluss der Förderschule
- Fachhochschulreife (z.B. Abschluss einer Fachoberschule)
- Polytechnische Oberschule (POS) Abschluss 8. Klasse (DDR Abschluss)
- Abitur / allgemeine oder fachgebundene Hochschulreife
- Hauptschulabschluss, Volksschulabschluss
- Realschulabschluss, Mittlere Reife, Fachoberschulreife
- Polytechnische Oberschule (POS) Abschluss 10. Klasse (DDR Abschluss)
- Anderer deutscher Schulabschluss, und zwar
- Ausländischer Schulabschluss und zwar
- kein Abschluss

Weiter

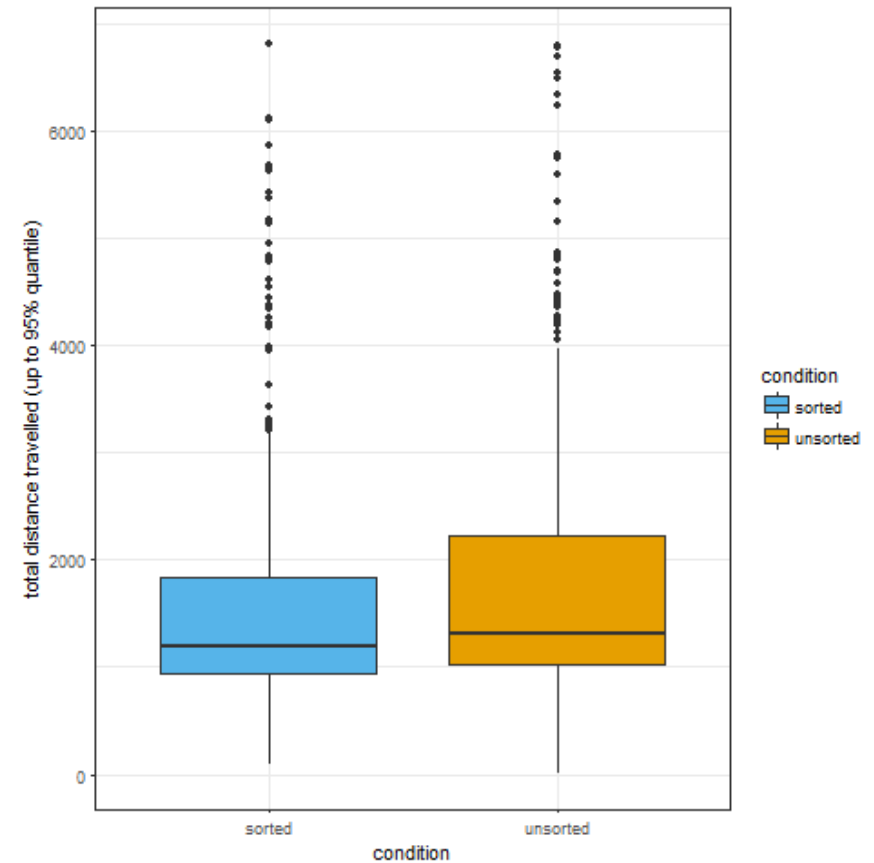
H1a: Greater Total Distance Traveled When Unordered

Educational Attainment



n=824; Wilcoxon Rank Sum Test=74,524; p=0.002

Level of Employee

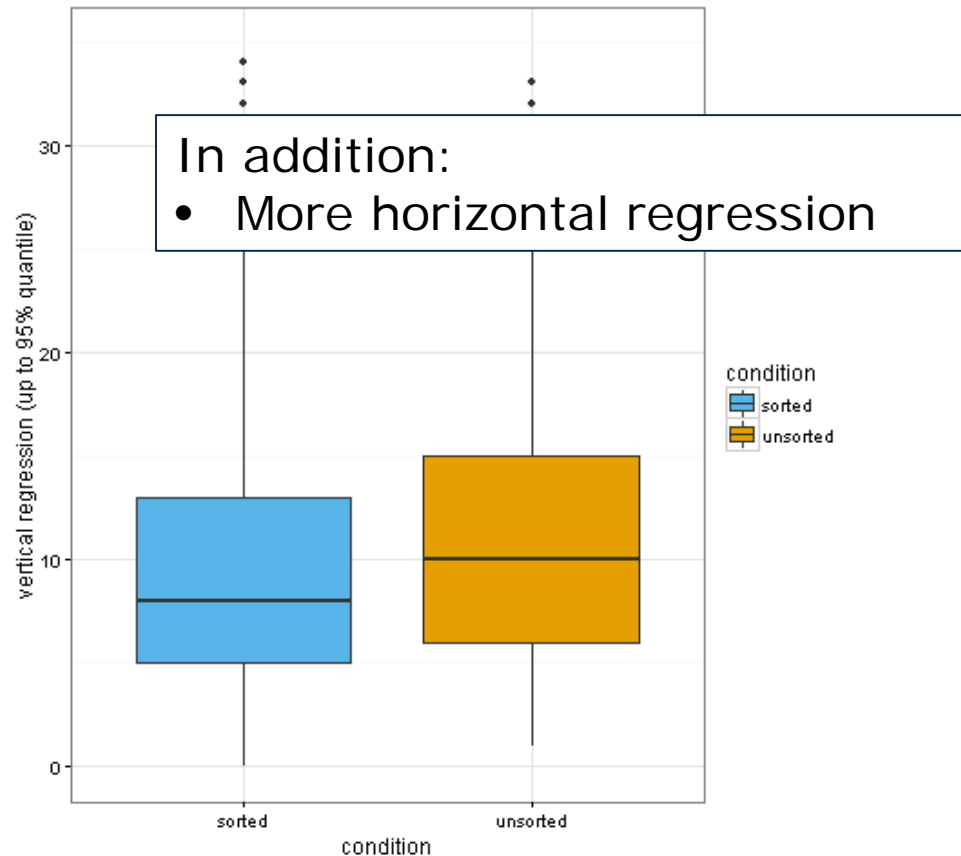


n=640; Wilcoxon Rank Sum Test=46,280; p=0.039

Figures remove top 5% of observations

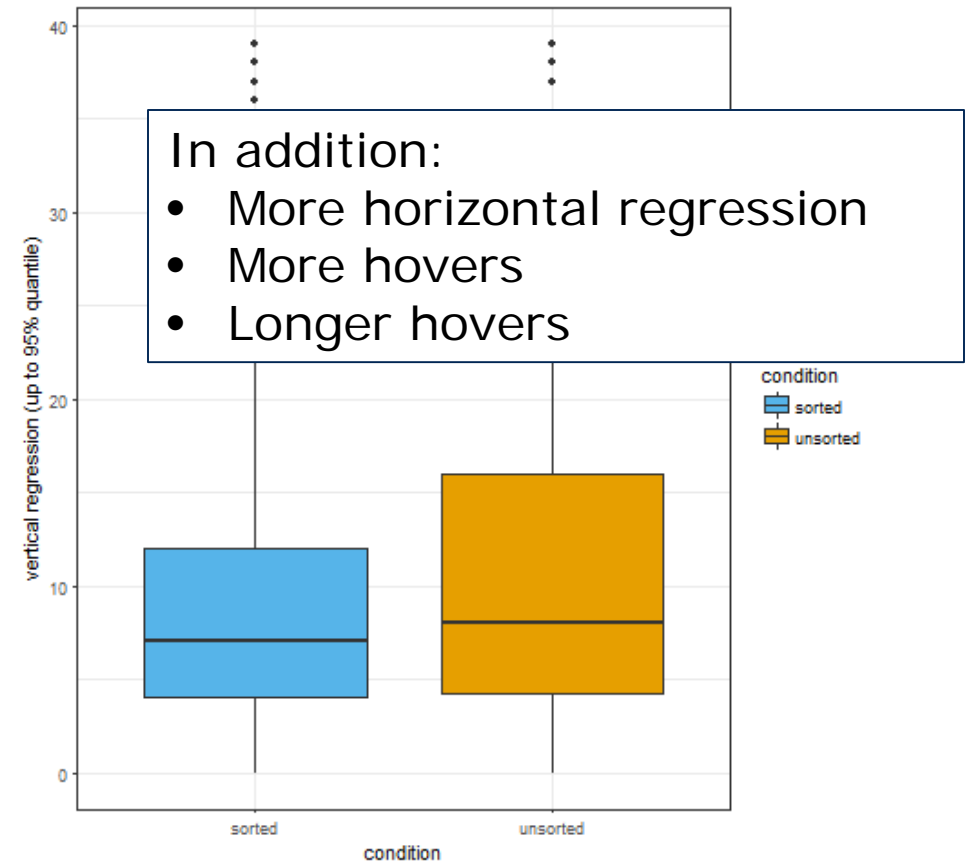
H1b: More Vertical Regression When Unordered

Educational Attainment



n=824; Wilcoxon Rank Sum Test=75,297; p=0.005

Level of Employee



n=640; Wilcoxon Rank Sum Test=45,740; p=0.021

Figures remove top 5% of observations

H2: Check-all-that-apply vs. Yes/No Grids

Aus welchen Gründen haben Sie Ihre vorherige Tätigkeit beendet?

Bitte benennen Sie alle zutreffenden Gründe.

- Ich war vorher bei keinem anderen Arbeitgeber beschäftigt
- Entlassung
- Beendigung eines befristeten Arbeitsvertrags
- Eigene Kündigung in der Erwartung eine bessere Arbeit zu bekommen
- Eigene Kündigung aus Unzufriedenheit mit dem Unternehmen
- Ruhestand, und zwar vorzeitig nach Vorruhestandsregelung oder nach Arbeitslosigkeit
- Ruhestand, und zwar aus gesundheitlichen Gründen
- Ruhestand, und zwar aus Alters- oder sonstigen Gründen
- Grundwehr-, Zivildienst
- Betreuung von Kindern, Pflegebedürftigen, Menschen mit Behinderung
- Sonstige persönliche und familiäre Verpflichtungen
- Beginn einer Ausbildung (auch Studium)
- Sonstige Gründe

Weiter

Aus welchen Gründen haben Sie Ihre vorherige Tätigkeit beendet?

Bitte benennen Sie alle zutreffenden Gründe.

- | | Ja | Nein |
|--|-----------------------|-----------------------|
| <input type="checkbox"/> Ich war vorher bei keinem anderen Arbeitgeber beschäftigt | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Entlassung | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Beendigung eines befristeten Arbeitsvertrags | <input type="radio"/> | <input type="radio"/> |
| <input checked="" type="checkbox"/> Eigene Kündigung in der Erwartung eine bessere Arbeit zu bekommen | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Eigene Kündigung aus Unzufriedenheit mit dem Unternehmen | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Ruhestand, und zwar vorzeitig nach Vorruhestandsregelung oder nach Arbeitslosigkeit | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Ruhestand, und zwar aus gesundheitlichen Gründen | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Ruhestand, und zwar aus Alters- oder sonstigen Gründen | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Grundwehr-, Zivildienst | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Betreuung von Kindern, Pflegebedürftigen, Menschen mit Behinderung | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Sonstige persönliche und familiäre Verpflichtungen | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Beginn einer Ausbildung (auch Studium) | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Sonstige Gründe | <input type="radio"/> | <input type="radio"/> |

Weiter

Aus welchen Gründen haben Sie Ihre vorherige Tätigkeit beendet?

Bitte benennen Sie alle zutreffenden Gründe.

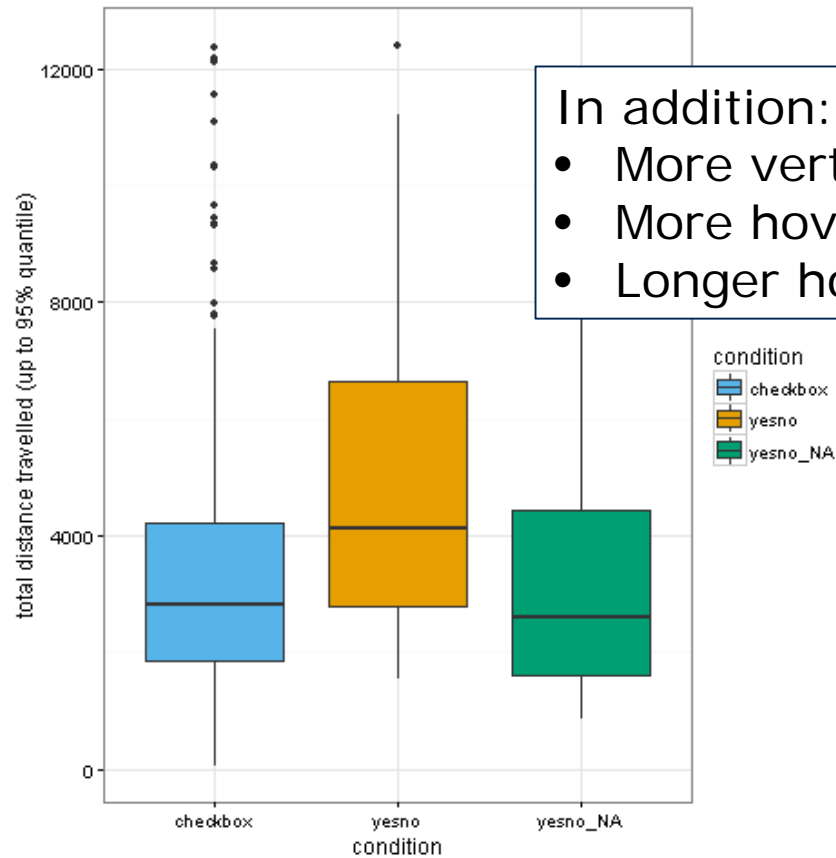
- | | Ja | Nein |
|--|-----------------------|-----------------------|
| <input checked="" type="checkbox"/> Ich war vorher bei keinem anderen Arbeitgeber beschäftigt | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Entlassung | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Beendigung eines befristeten Arbeitsvertrags | <input type="radio"/> | <input type="radio"/> |
| <input checked="" type="checkbox"/> Eigene Kündigung in der Erwartung eine bessere Arbeit zu bekommen | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Eigene Kündigung aus Unzufriedenheit mit dem Unternehmen | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Ruhestand, und zwar vorzeitig nach Vorruhestandsregelung oder nach Arbeitslosigkeit | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Ruhestand, und zwar aus gesundheitlichen Gründen | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Ruhestand, und zwar aus Alters- oder sonstigen Gründen | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Grundwehr-, Zivildienst | <input type="radio"/> | <input type="radio"/> |
| <input checked="" type="checkbox"/> Betreuung von Kindern, Pflegebedürftigen, Menschen mit Behinderung | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Sonstige persönliche und familiäre Verpflichtungen | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Beginn einer Ausbildung (auch Studium) | <input type="radio"/> | <input type="radio"/> |
| <input type="checkbox"/> Sonstige Gründe | <input type="radio"/> | <input type="radio"/> |

Weiter

H2: Check-all-that-apply vs. Yes/No Grids

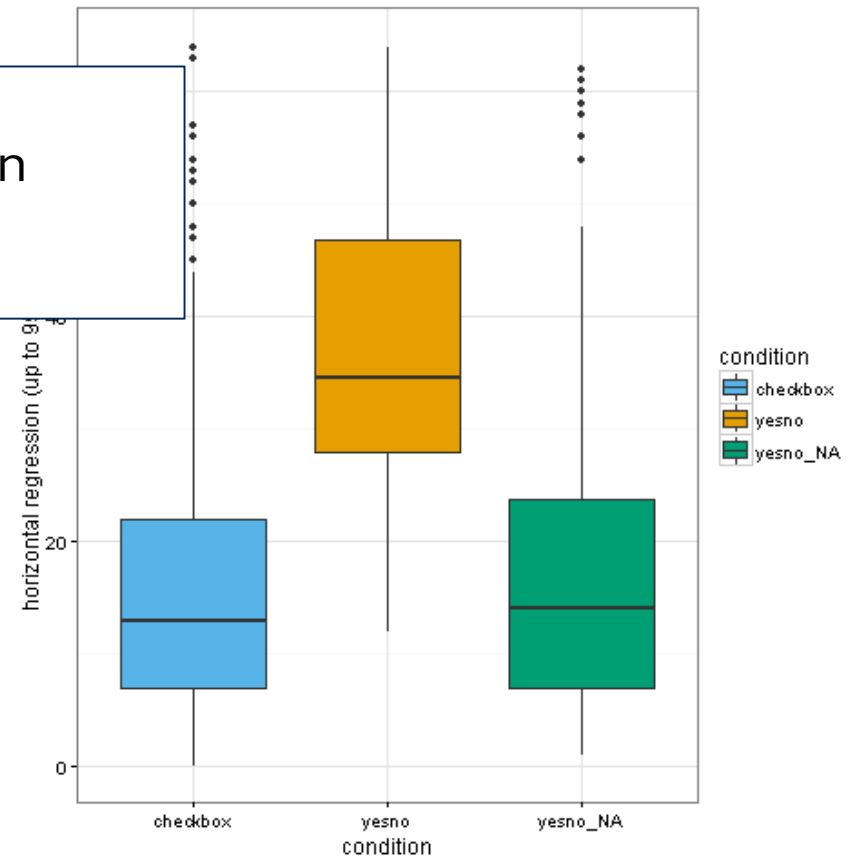
H2a: Greater Total Distance Traveled in Yes/No

H2b: More Horizontal Regression in Yes/No



In addition:

- More vertical regression
- More hovers
- Longer hovers



n=537; Kruskal-Wallis Rank Sum Test=27.91; p<0.001

n=537; Kruskal-Wallis Rank Sum Test=104.03; p<0.001

Figures remove top 5% of observations

H3: Straightforward vs. Difficult Response Options

Straightforward Response Options

UNIVERSITÄT
MANNHEIM

Institut für Arbeitsmarkt-
und Berufsforschung
Die Forschungseinrichtung der
Bundesagentur für Arbeit **IAB**

Sind Sie derzeit...

Falls Sie momentan in mehreren Beschäftigungsverhältnissen stehen, denken Sie bitte bei dieser und den folgenden Fragen an Ihre hauptsächliche Beschäftigung. Das ist diejenige Erwerbstätigkeit, in der Sie die meisten Stunden arbeiten.

- angestellt bei einem privaten, profitorientierten Unternehmen
- angestellt bei einem privaten, nicht profitorientierten Unternehmen
- angestellt im öffentlichen Dienst des Bundes
- angestellt im öffentlichen Dienst eines Bundeslandes
- angestellt im öffentlichen Dienst einer Kommune
- selbständig in einem freien Beruf
- selbständig in Handel, Gewerbe, Industrie, Dienstleistung, Landwirtschaft
- Freier Mitarbeiter/Freie Mitarbeiterin
- Mithelfende/r Familienangehörige/r

Weiter

Difficult Response Options

UNIVERSITÄT
MANNHEIM

Institut für Arbeitsmarkt-
und Berufsforschung
Die Forschungseinrichtung der
Bundesagentur für Arbeit **IAB**

Sind Sie derzeit...

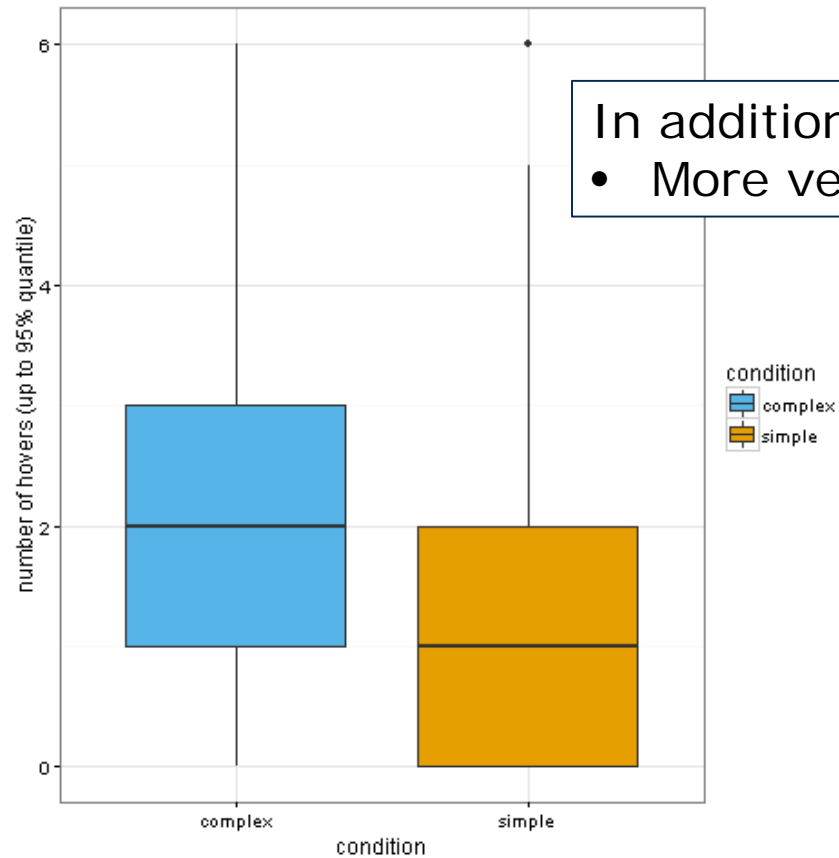
Falls Sie momentan in mehreren Beschäftigungsverhältnissen stehen, denken Sie bitte bei dieser und den folgenden Fragen an Ihre hauptsächliche Beschäftigung. Das ist diejenige Erwerbstätigkeit, in der Sie die meisten Stunden arbeiten.

- angestellt bei einem privaten, profitorientierten Unternehmen oder einer Privatperson, das/die regelmäßig Löhne oder Gehälter auszahlt
- angestellt bei einem privaten, nicht profitorientierten Unternehmen, welches gemeinnützig ist oder von Steuervergünstigungen profitiert
- angestellt im öffentlichen Dienst des Bundes (Arbeitgeber ist die Bundesrepublik)
- angestellt im öffentlichen Dienst eines Bundeslandes (Arbeitgeber ist ein Bundesland)
- angestellt im öffentlichen Dienst einer Kommune (Arbeitgeber ist zum Beispiel eine Gemeinde, eine Stadt, ein Landkreis, ...)
- selbständig in einem freien Beruf (z.B. Arzt/Ärztin, Rechtsanwalt/Rechtsanwältin oder Architekt/in)
- selbständig, jedoch nicht in einem freien Beruf im Bereich Handel, Gewerbe, Industrie, Dienstleistung, Landwirtschaft
- Freier Mitarbeiter/Freie Mitarbeiterin, der aufgrund eines Dienst- oder Werkvertrags für ein Unternehmen Aufträge persönlich ausführt
- Mithelfende/r Familienangehörige/r ohne Bezahlung

Weiter

H3: More Hovers When Response Options Are Difficult

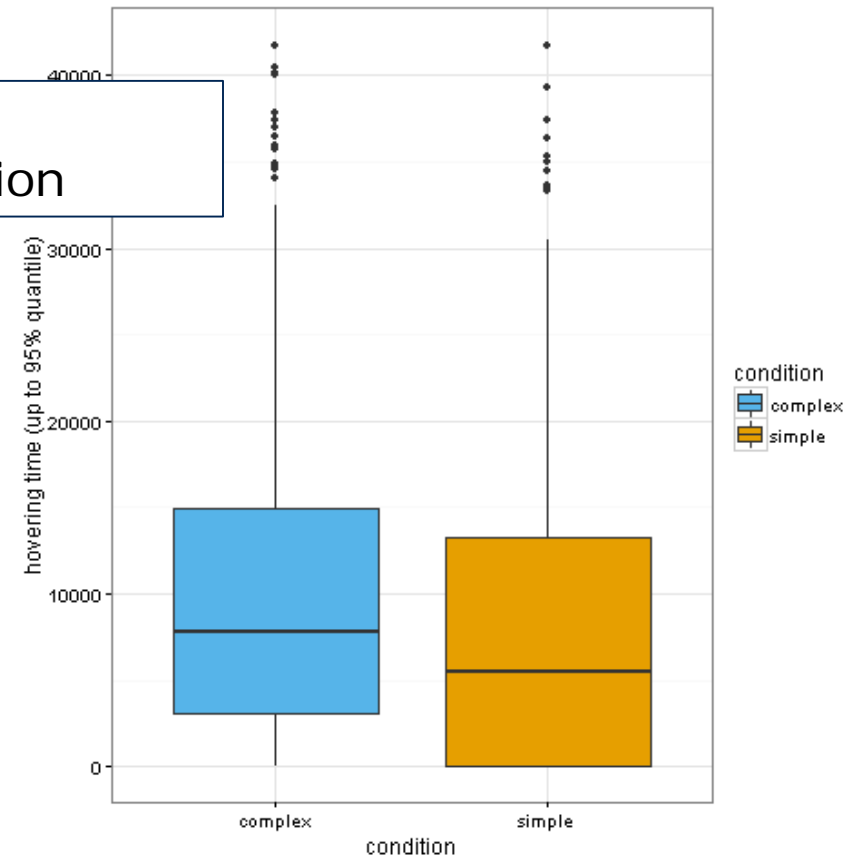
Number of Hovers



In addition:
• More vertical regression

n=716; Wilcoxon Rank Sum Test=77,306; p<0.001

Hover Time



n=716; Wilcoxon Rank Sum Test=75,430; p<0.001

Figures remove top 5% of observations

H4: General vs. Specific

General Economic Situation

UNIVERSITÄT
MANNHEIM

Institut für Arbeitsmarkt-
und Berufsforschung
Die Forschungseinrichtung der
Bundesagentur für Arbeit



Wie beurteilen Sie ganz allgemein die heutige wirtschaftliche Lage in Deutschland?

- Sehr gut
- Gut
- Teils gut/teils schlecht
- Schlecht
- Sehr schlecht

Weiter

Personal Economic Situation

UNIVERSITÄT
MANNHEIM

Institut für Arbeitsmarkt-
und Berufsforschung
Die Forschungseinrichtung der
Bundesagentur für Arbeit

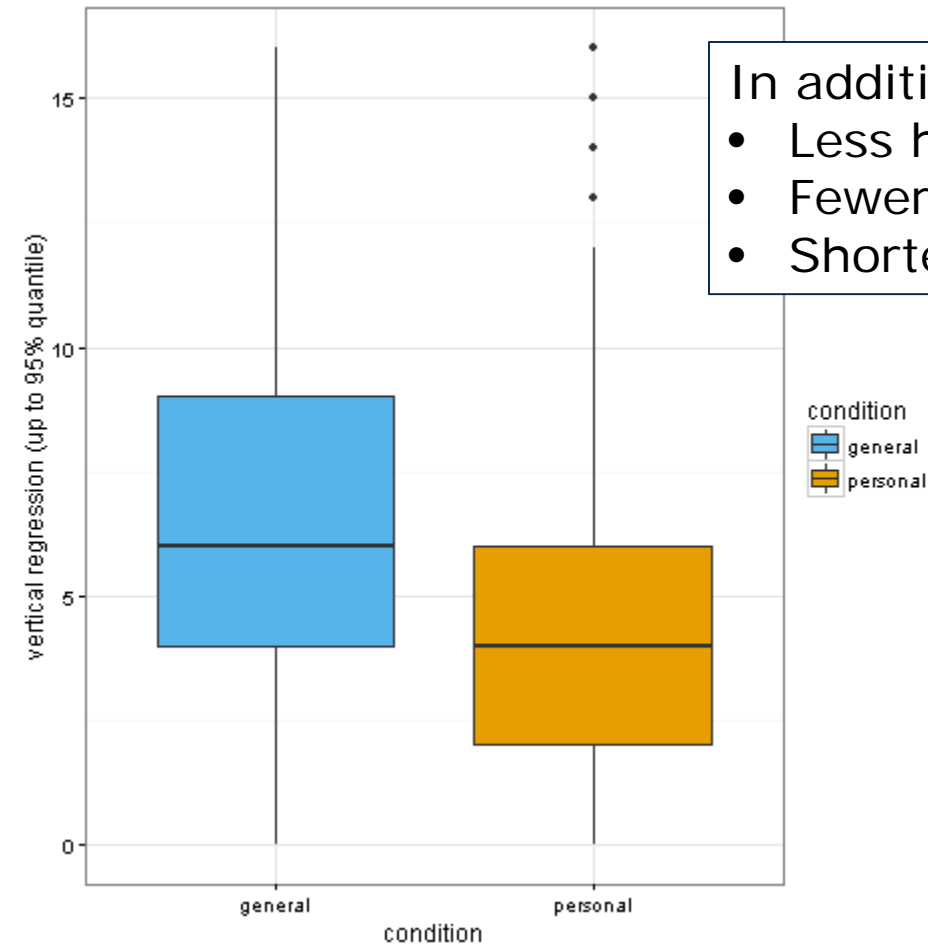


Wie beurteilen Sie Ihre eigene wirtschaftliche Lage heute?

- Sehr gut
- Gut
- Teils gut/teils schlecht
- Schlecht
- Sehr schlecht

Weiter

H4: More Vertical Regression in General Questions



In addition:

- Less horizontal regression
- Fewer hovers
- Shorter hovers

n=778; Wilcoxon Signed Rank Test=47,320; p<0.001

Figures remove top 5% of observations

Summary

- Stable indicators of difficulty based on mouse tracking measures
- Effects of manipulation were more general than expected
 - Each difficulty led to more vertical and horizontal regression and greater total distance traveled
- *Rs* seem to hover when response options not clear
 - Yes/No grids, difficult terms, more than one answer possible

Next Steps

1. Explain **measurement error** using mouse movements
 2. Explore of **privacy aspects** of tracking paradata
-
1. Develop interpretable regression models for continuous response with trajectories as predictors **based on semi-metrics**
 2. Develop interpretable regression models for continuous response with trajectories as predictors **based on shapelets**
 3. Development of **ensemble learners** for classification and regression (supervised learning) with trajectories

Turning the perspective around



YES

Do mouse movements provided sufficient predictive performance to recover the experimental groups, and which features of mouse movements were particularly important?

Notably, the presence of a significant difference in certain mouse movements between easy and difficult versions of a question is, while an indicator, no guarantee for the predictive power of a specific feature.



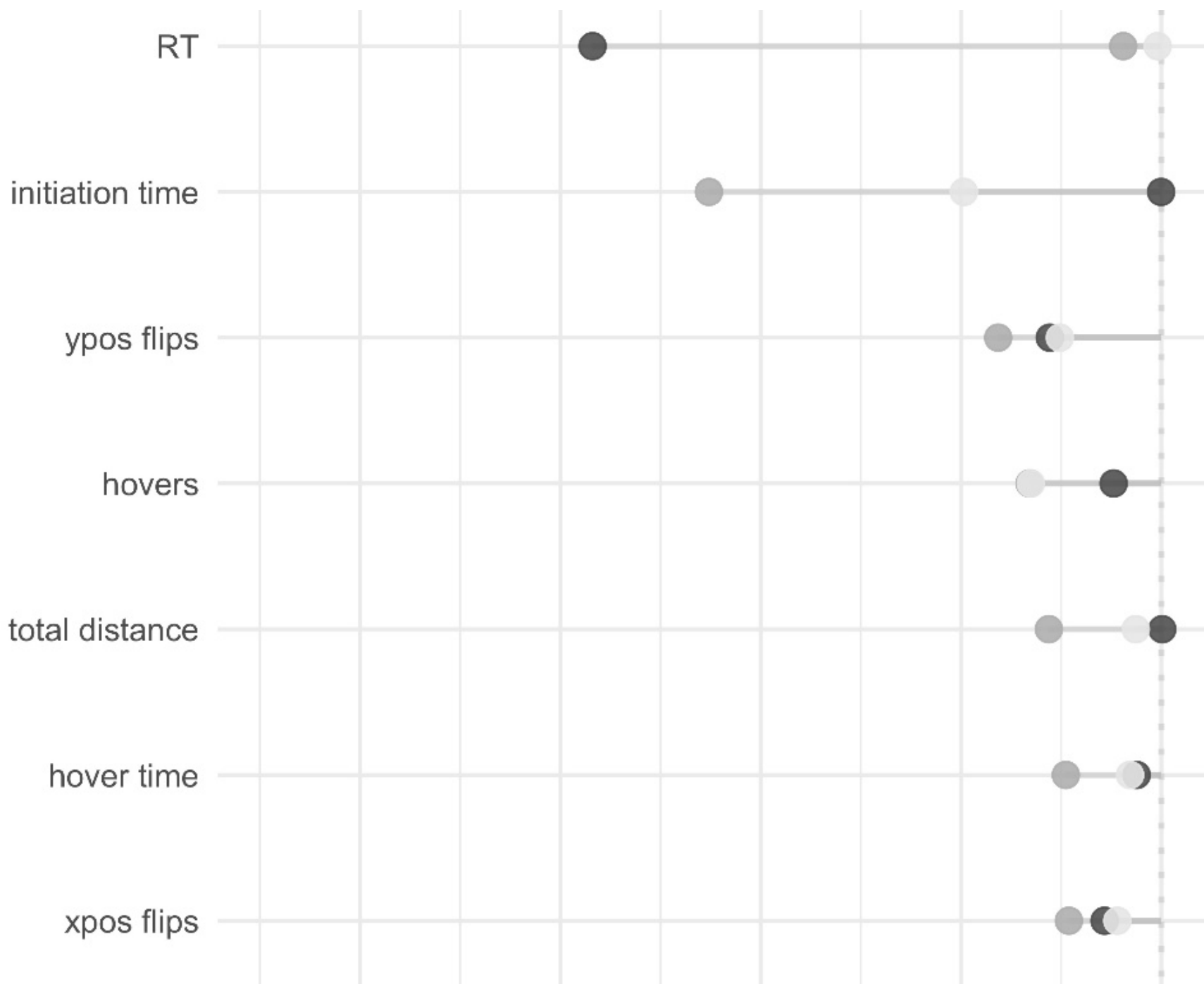
YES

Is the added predictive power of mouse movement features sufficiently large for their use in informing real-time responsive questionnaire design or in measurement error corrections?



When predicting response difficulty, what do we gain beyond response time through mouse-tracking features, and can we further improve prediction through personalization?

YES



Impact of measure is evaluated through accuracy reduction (randomly permuted).
The more negative the reduction, the more impact the measure has in the model.

- question
- employment detail
 - employee level
 - education level

Thank You!



Frauke Kreuter

University of Maryland, College Park

University of Mannheim, Germany

Institute for Employment Research

fkreuter@umd.edu