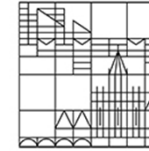


Survey Experiments (Day 1)

GESIS Summer School 2014
Kyiv, September 15-19

Prof. Dr. Katrin Auspurg, Goethe-University Frankfurt Main
Prof. Dr. Thomas Hinz, University of Konstanz

Instructors and Organization



- **Instructors**

- Prof. Dr. Katrin Auspurg (Goethe University Frankfurt Main)
- Prof. Dr. Thomas Hinz (University of Konstanz)

- **Collaboration**

- The material of this seminar (slides etc.) was developed together with Dr. Carsten Sauer, Bielefeld University.

- **Organization**

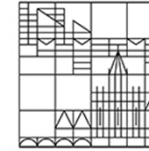
- Prof. Andrii Gorbachyk

- **Funding**

- German Academic Exchange Service (DAAD)

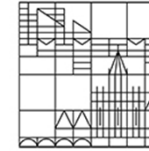
- **Please introduce yourself:** research fields, background, motivation for the seminar, experience with survey experiments and Stata; do you have an idea for a project you want to work on during the seminar?

Course Contents and Objectives



- Theoretical and practical overview on survey methods (in particular factorial survey methods).
- Some information on further experimental survey methods (conjoint analysis, choice experiments).
- Practical insights into steps that are needed to design factorial survey experiments:
 - specification of dimensions and levels,
 - selection of an experimental design,
 - drafting and programming of questionnaires,
 - special methods needed for data analyses (like multilevel regression analyses, estimation of willingness to pay).
- Participants should select a research question related to their own thesis or field of research for these practical exercises.

Agenda



Monday, 15th

- Survey experiments: introduction, selection of dimensions and levels
- Working with the statistical software Stata

Tuesday, 16th

- Vignette and respondent samples, experimental designs
- Building up vignette texts

Wednesday, 17th

- Survey modes, answer scales, data collection; generation of questionnaires
- Presentations and discussion of research ideas of participants I

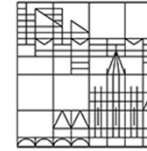
Thursday, 18th

- Data preparation and data analyses
- Presentations and discussion of research ideas of participants II

Friday, 19th

- Methodological research, related methods
- Final discussion and additional advice for participants' research projects

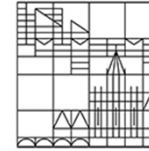
Organization of the Course



Literature and course material

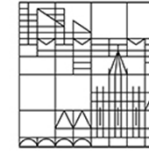
- All course materials (slides, data, assignments, etc. are provided at Dropbox (you will be invited by an e-mail).
- Material will be available in the morning hours at the latest.

Presentations of Participants



- You should present first ideas during the course, in particular if you are interested in research collaborations or staying abroad in Germany .
- Presentations should cover:
 1. **Research question and motivation:** What is the research question you want to work on? Why to use a survey experiments like FS for this research question? (The research question may be methodological or substantive.) Are there already other studies working with an experimental design?
 2. **Theory and hypotheses:** Which theory or hypotheses do you want to test? You should specify at least one hypothesis.
 3. **Research Design:** Which dimensions and levels will you use? What could be a task respondents should do in your experiments? (→ draft of a vignettes).
 - [4. Outlook: What kind of analyses do you plan?]
- Presentation format: ~10 minutes; you might prepare slides or hand-outs.

Agenda



Monday, 15th

- **Survey experiments: introduction**, selection of dimensions and levels
- Working with the statistical software Stata

Tuesday, 16th

- Vignette and respondent samples, experimental designs
- Building up vignette texts

Wednesday, 17th

- Survey modes, answer scales, data collection; generation of questionnaires
- Presentations and discussion of research ideas of participants I

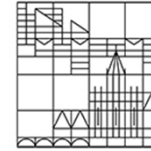
Thursday, 18th

- Data preparation and data analyses
- Presentations and discussion of research ideas of participants II

Friday, 19th

- Methodological research, related methods
- Final discussion and additional advice for participants' research projects

Factorial Surveys: First Steps

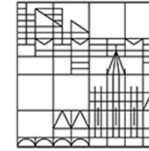


Please fill in the sample questionnaire!

And memorize your ID!

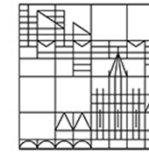
(You will need the ID on Thursday for data analyses)

Experimental Survey Methods



- Experience with filling in the sample questionnaire?
- Features of experiments?
 - Stimulus (experimental variation) deliberately set by researcher
 - At least two different groups of participants (experimental and control)
 - Random assignment of stimulus to groups of participants
- Advantages, disadvantages?

Factorial Surveys: Idea

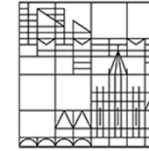


Age Sex Education Occupation

A 35-year old man with vocational training is working as a programmer.
His gross income amounts € 2.500 per month.

According to your opinion, is the income of the described person fair or is it unfairly too high or too low?

- 5 -4 -3 -2 -1 0 1 2 3 4 5
- unfairly too low fair unfairly too high



Factorial Surveys: Research Aims

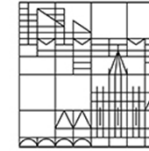
- “What information do respondents use to make judgments. How may such information be used? And how do individuals differ in the ways in which information of different sorts is combined?” (Rossi/Anderson 1982: 19) [1]
- Measuring influence of:
 - Single attributes (dimensions) on evaluations (e.g., attitudes, decisions)
 - respondents’ characteristics on vignette evaluations
 - amount of consensus across respondents.
- Testing theoretical models on the influence of single dimensions, wherefore additive models are used, e.g. (without assuming any interactions):

$$Y_{ij} = \beta_0 + \beta_1 \text{Dimension}_{1j} + \dots + \beta_k \text{Dimension}_{kj} + \varepsilon_{1j}$$

With:

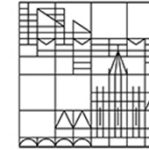
i:	index for single respondents	j:	index for single vignettes
Y_{ij} :	Vignette rating	β_0 :	Intercept/mean rating
β_k :	Impact/coeff. of dimension k	ε_{ij} :	random error in rating

Factorial Surveys: Advantages [2,3,4,5]

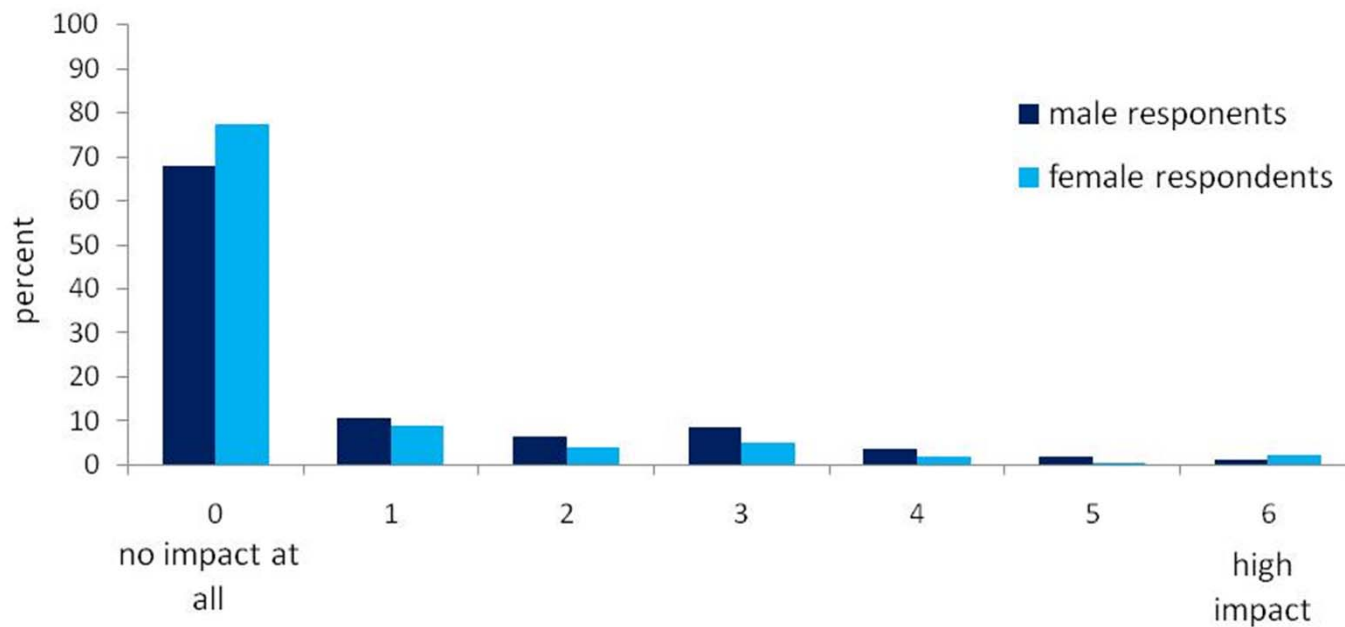


- Experimental design: enables disentangling of causal factors which are confounded in reality and enables evaluations of rare situations.
- Integration in survey: hypotheses can be tested on the basis of larger (random) samples of general population compared to the lab (→ higher external validity).
[Heinrich et al 2009: most experiments with WEIRDest People of the world: Western, Educated, Industrialized, Rich, and from Democratic societies.]
- Multi-factorial design (“conjoint”): simulation of the complexity of real world decisions by crossing a variety of factors (dimensions); respondents are forced to trade-off dimensions; stimuli are more standardized.
- Several vignettes per respondent: large numbers of cases can be achieved with few respondents, which means economizing research resources.
- Indirect evaluations: probably less prone for social desirability bias than direct questioning of sensitive items.^[6]

Example: Just Gender Pay Gap? Direct Questioning

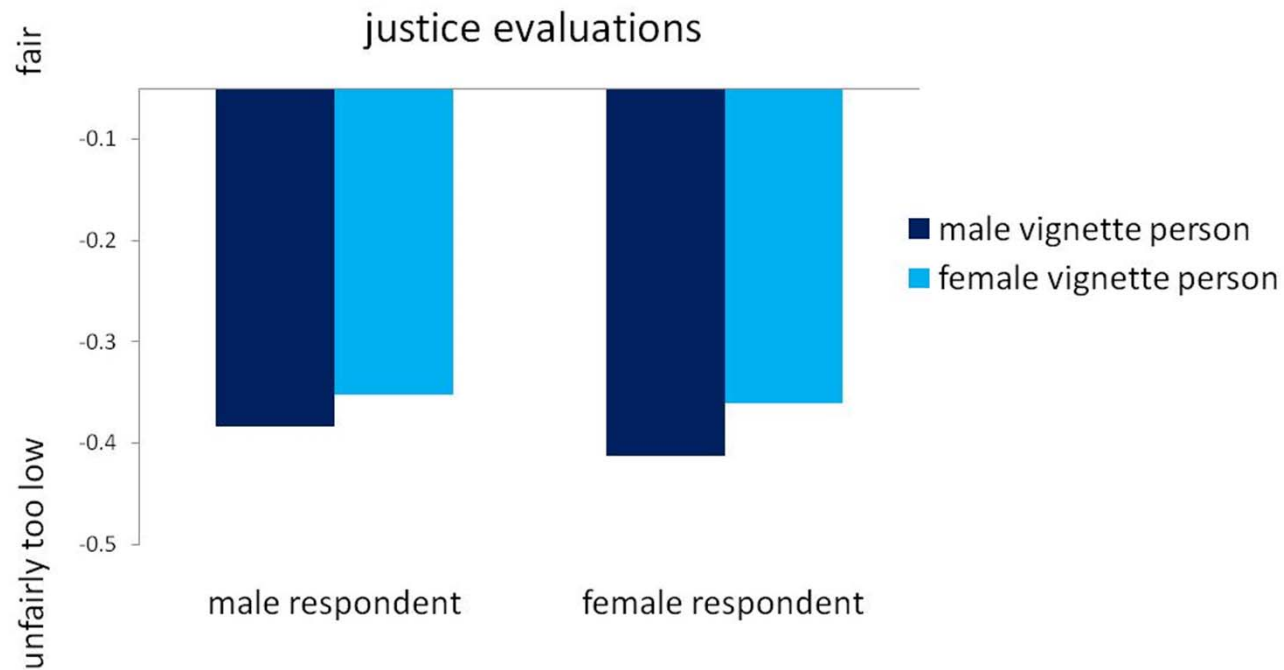
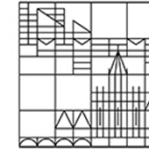


"According to your opinion, which impact should have the following aspects for the amount of fair earning? – sex of employees"



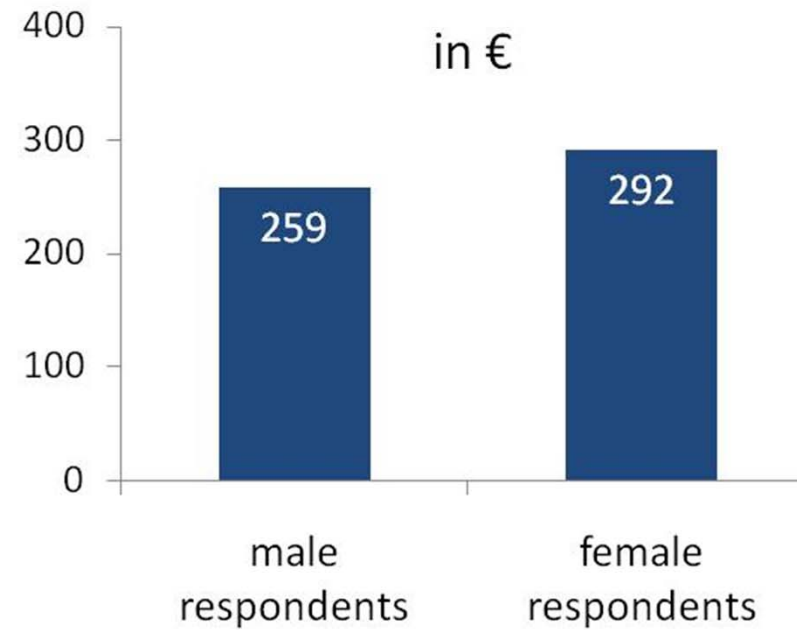
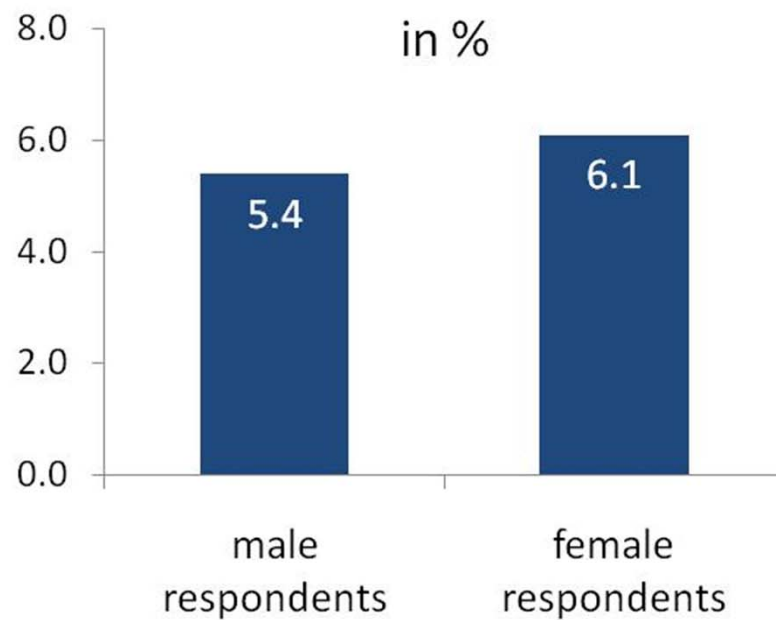
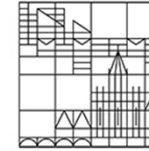
Source: Project „Factorial Survey Design“, general population sample (GSOEP)

Example: Just Gender Pay Gap? Vignettes



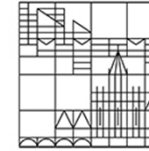
Source: Project „Factorial Survey Design“, general population sample (GSOEP)

Example: Just Gender Pay Gap

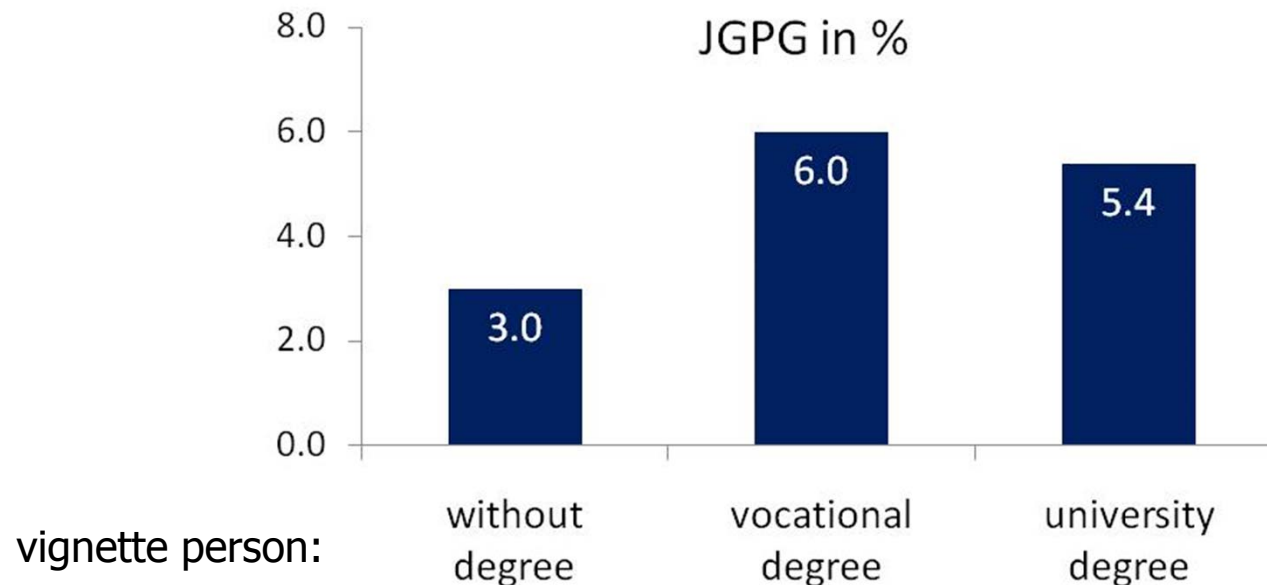


Source: Project „Factorial Survey Design“, general population sample (GSOEP)

Example: Interactions Between Dimensions

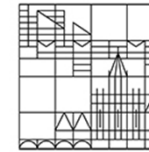


- Subgroup analyses and testing for interactions allow additional insights in judgement principles.
For instance, the mechanisms underlying the JGPG can be revealed: Gender specific income levels or gender specific returns to education?

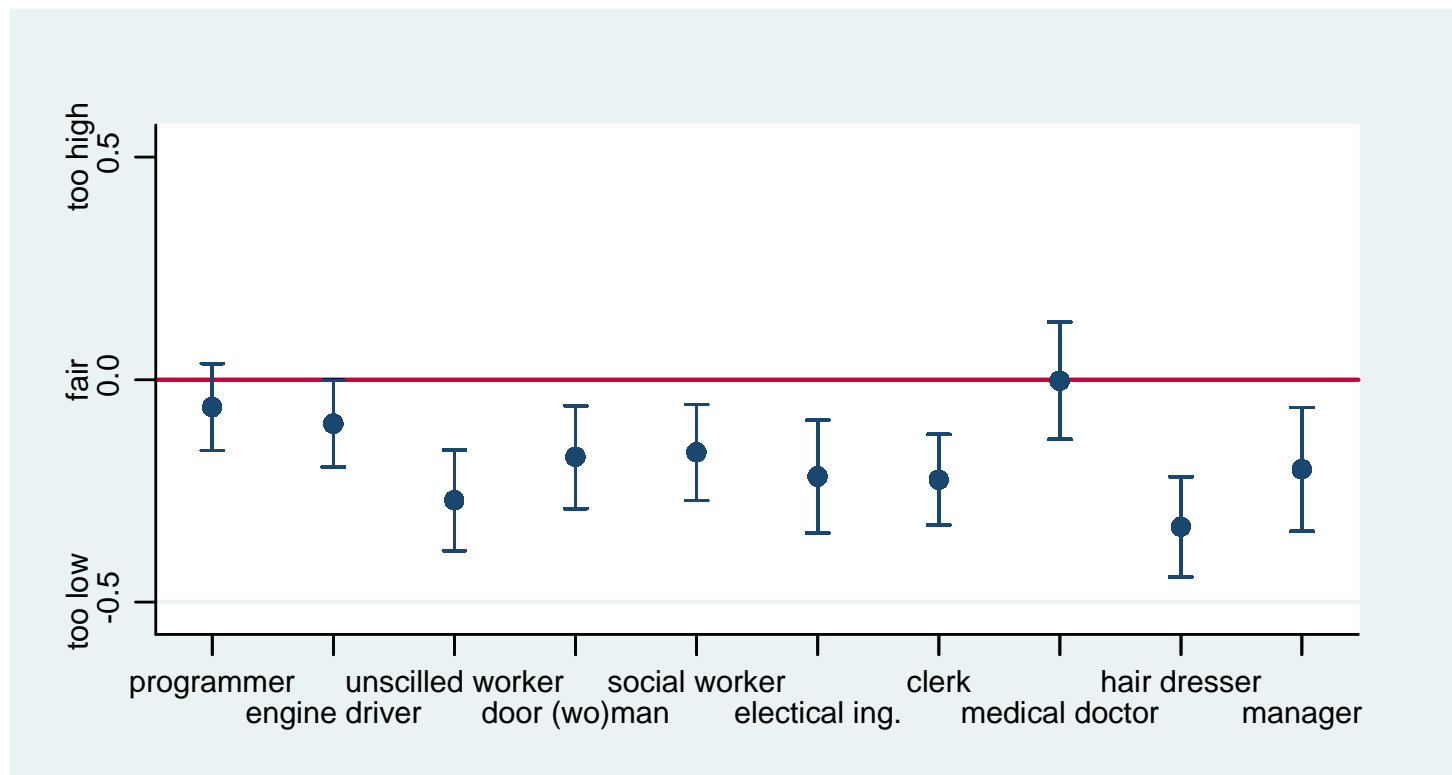


Source: Project „Factorial Survey Design“, general population sample (GSOEP)

Example: Cross-Level Interactions

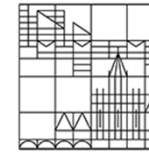


- Variation of JPGs with respondents' occupations: framed by real experience like status construction theory would predict? Framed by gender norms?



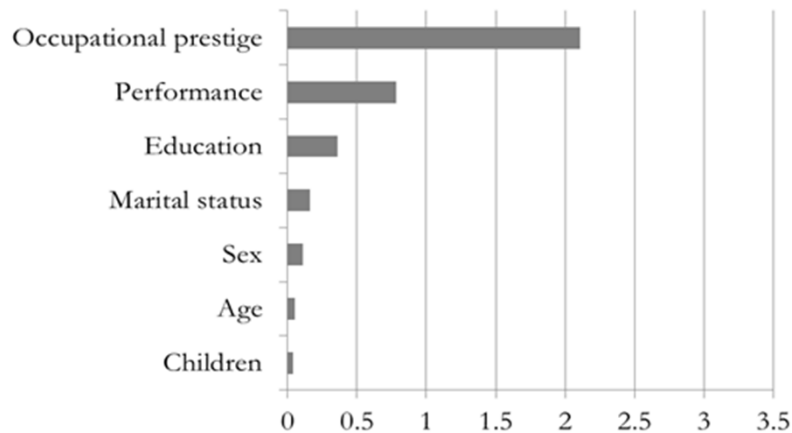
Source: Project „Factorial Survey Design“, general population sample and GSOEP data

Example: Cross-National Comparisons

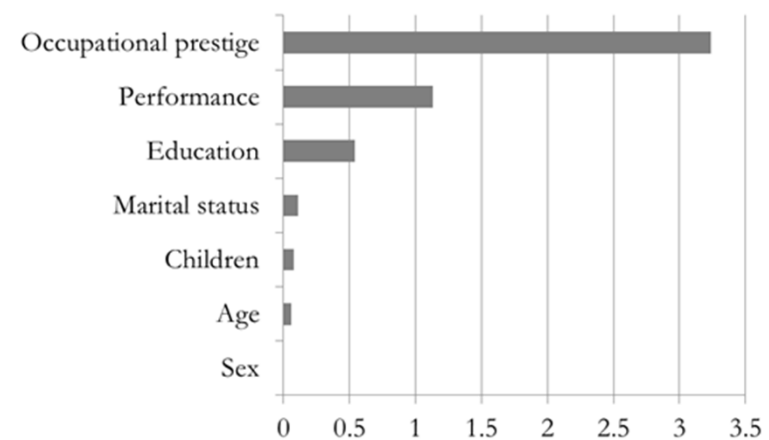


- Distribute practices perceived as just in Germany and Ukraine in 2009
(Auspurg/Hinz/Gatskova 2012): Relative importance of dimensions (semi-partial R^2)

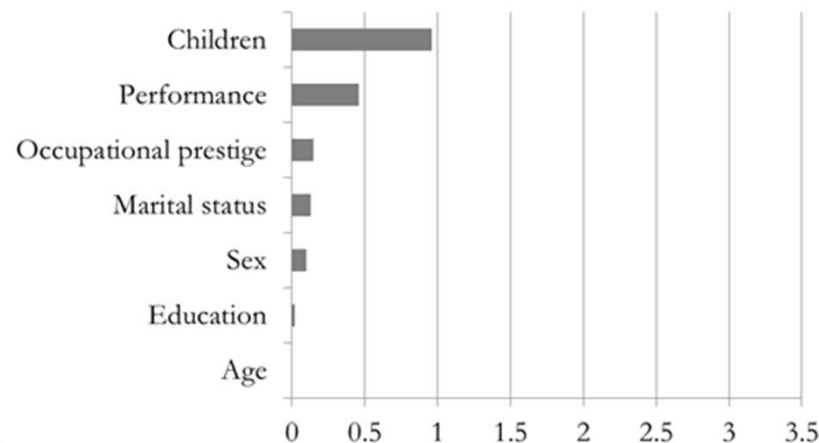
West Germany



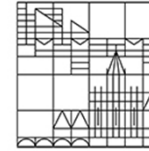
Eastern Germany



Ukraine



Factorial Surveys: Shortcomings

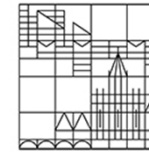


- High complexity of evaluation task; risk of cognitive overburden of respondents and methodological caused 'artefacts'.
- Repeated judgments per respondents: risk of learning effects (like use of heuristics) or fatigue effects.
- Measurement of decisions: only hypothetical, no real decisions.
- Very few research on methodological issues, lack of empirically tested guidelines and more effort needed for designing and administering surveys.

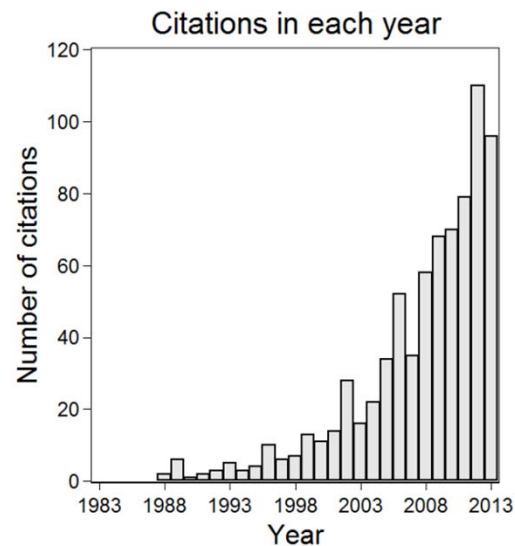
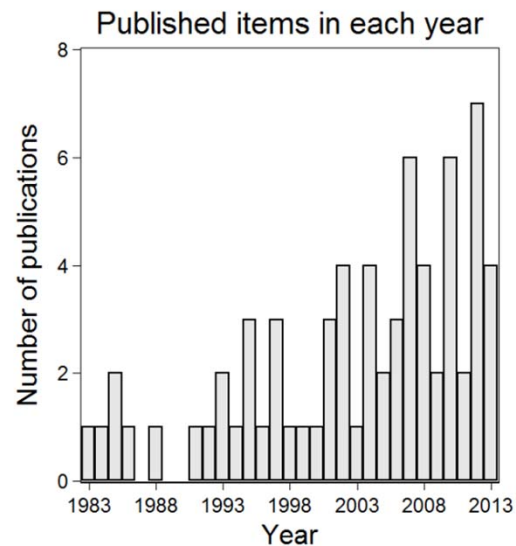
But nevertheless:

- The method is more and more regarded as a promising way of analyzing a broad range of substantive issues in social sciences.

Applications in Social Sciences^[3]

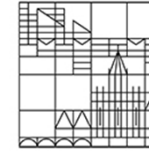


- 1951: Established by Peter H. Rossi in his dissertation
- 1982: First introductory text by Rossi and Nock
- 1982 – 2006: 106 studies published in core journals of sociology



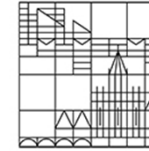
Source: Thompsons Web
of Science

Related Methods



- And there is also an increasing use of related experimental survey methods like conjoint analyses, choice experiments.
- Are you familiar with these methods?
- During the seminar, we will only cover **multifactorial survey experiments**.
(And not the scenario approach, simple split-ballot experiments, anchoring vignettes).
- The starting point will be factorial surveys (FSs), but we will discuss related methods on Friday.

Agenda



Monday, 15th

- Survey experiments: introduction, **selection of dimensions and levels**
- Working with the statistical software Stata

Tuesday, 16th

- Vignette and respondent samples, experimental designs
- Building up vignette texts

Wednesday, 17th

- Survey modes, answer scales, data collection; generation of questionnaires
- Presentations and discussion of research ideas of participants I

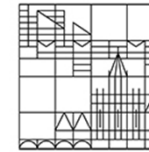
Thursday, 18th

- Data preparation and data analyses
- Presentations and discussion of research ideas of participants II

Friday, 19th

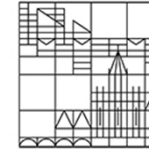
- Methodological research, related methods
- Final discussion and additional advice for participants' research projects

Main Terms and Components [1,2]



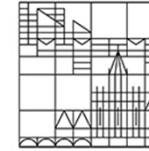
- Fictitious descriptions of situations or objects = *vignettes*.
- Characteristics varying across vignettes = *dimensions* [*attributes*].
- Values of these dimensions = *levels*.
- Subset of vignettes presented to single respondents = *deck* or *set* [*block*].
- All possible vignettes (combinations of levels) = *vignette universe* [*full factorial*].
- The vignette universe is obtained by orthogonally crossing all dimensions. The number of vignettes in the universe is the *Cartesian product* of all level dimensions.
- For example: 4 dimensions with 3, 4, 3 and 4 levels:
 $3 \cdot 4 \cdot 3 \cdot 4 = 3^2 \cdot 4^2 = 9 \cdot 16 = 144$ vignettes.

Main Terms and Components [1,2]



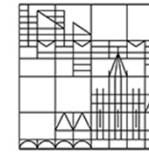
- Selection of vignettes drawn from the vignette universe = *vignette sample* [*fractional factoria*].
- The number of vignettes increases exponentially with the number of dimensions and levels. In most cases **random or systematically drawn samples** and **different decks** are used.
(For example: 15 decks à 10 vignettes: $15 \times 10 = 150$ vignettes sample overall)
- *Experimental design* or *experimental setup*: Plan for running an experiment. Kind of experimental factors and the way they are combined.
The experimental design determines the parameter identification, model flexibility, and the statistical efficiency of resulting estimates!
- The goodness of an experimental design is often evaluated by its *statistical efficiency*: precision of parameter estimates that can be reached.

Methodological Trade-Offs [3]



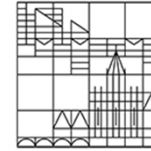
- There are several trade-offs between methodological decisions. For instance:
 - + The higher the number of levels and dimensions, the more information on judgement principles is gained.
 - But the higher is the cognitive burden for respondents; and
 - the higher is the number of vignettes in the universe and in turn the number of respondents required for meaningful data analyses.
- Similar trade-offs exist between the size of the vignette sample, the number of evaluations gained by different respondents, and statistical power.

Specification of Dimensions



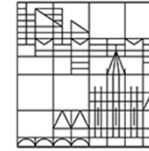
- Should be based on theories or prior empirical research.
- Ideally, one formalizes the assumed association between vignette evaluations and dimensions (and respondents' characteristics), including possible interactions and non-linear terms.
- **Number of dimensions:**
 - Especially if employing several vignettes per respondents: **~ 6-9 dimensions.**
 - Even more complex vignettes get evaluated (e.g.: 30 vignettes with 12 dimensions: very low missing values).
 - There is, however, evidence that in case of a high complexity (> 10 vignettes, ≥ 8 dimensions) older or less educated respondents show a lower consistency of responses and there is some evidence unwanted methodological effects like simplifying heuristics and order effects. ^[10]
 - **Information that is not varying across the vignettes: should be specified in the introductory part of FS modules / general instructions for respondents.**

Specification of Levels



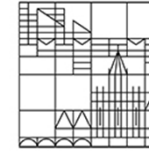
- Use similar numbers of levels across dimensions to avoid “number-of-levels-effects.”^[7]
- Use more than two levels only if it is necessary for theoretical or methodological reasons (e.g. to specify non-linear relationships).
- Avoid the occurrence of many illogical and implausible combinations.

Assignment # 1a (Individual Learning)



- Design a Factorial Survey module for a your own research question. (Use [SurveyExp_vignettedesign_income.docx](#) and [SurveyExp_vignettedesign_template.docx](#)).
- You may work alone on your own thesis or in teams of 2-3 people.
 - Which research question do you want to resolve by means of the FS?
 - What are the most important theories and hypotheses? How can you operationalize them in a FS?
 - Specify dimensions and levels; draft the vignette text. That is, you should fill in sections 1, 2 and 5.
- We will discuss some of the drafts during the next days. Please be prepared to explain your research aims and specifications to the other course members.
- The other issues (answer scales, introductory information, illogical cases,) will be covered during the next days. You do not have to work on them yet.

Agenda



Monday, 15th

- Survey experiments: introduction, selection of dimensions and levels
- **Working with the statistical software Stata**

Tuesday, 16th

- Vignette and respondent samples, experimental designs
- Building up vignette texts

Wednesday, 17th

- Survey modes, answer scales, data collection; generation of questionnaires
- Presentations and discussion of research ideas of participants I

Thursday, 18th

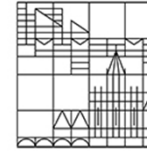
- Data preparation and data analyses
- Presentations and discussion of research ideas of participants II

Friday, 19th

- Methodological research, related methods
- Final discussion and additional advice for participants' research projects

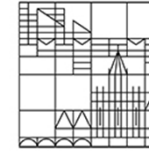
Assignment # 1b

Working with Stata



- The folder “day1/exercise” includes the following files:
 - an introductory manual with exercises.
 - an overview on Stata commands.
 - a factorial survey dataset.
 - two text files, so called do-files, that can be used as templates.
- For those who are new to Stata and have basic knowledge of statistics:
 - Go through this manual carefully and ask the instructors whenever something is unclear. At the end of the afternoon you should have at least completed exercise [9 of part A](#).
- For advanced Stata users:
 - Make sure that you know all commands in the manual.
 - Try to solve the advanced exercises!

Display Screens



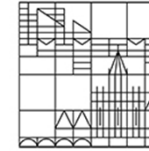
Four different windows:

- Commands
- Results
- Variables
- Review

Additional window
program syntax

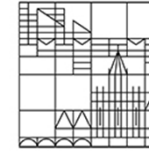
The screenshot displays the Stata software interface with four windows open:

- Review window:** Shows the command history and results. The command `do "D:\Users\Corinna\AppData\Local\Temp\STDG000000.tmp"` is listed. The results section shows the Stata perpetual license information for Universität Konstanz, including the serial number 40120525229 and the license to Universität Konstanz.
- Variables window:** A table with two columns: Variable and Label. The variables listed are `block`, `run`, `x1`, `x2`, `x3`, `x4`, `x5`, `x6`, `x7`, and `x8`.
- Command window:** Shows the command `tab block`.
- Do-file Editor window:** Shows the Stata program syntax for `SIEGEN-Sample Questionnaire_1.do`. The code includes comments in German and English, and Stata commands such as `od "D:\Users\Corinna\Desktop\Workshop Siegen\Corinna\Fragebogen\siegen"`, `*log using output, replace`, `use Blocked_120, replace`, `end of do-file`, `fre x1-x8`, `corr x1-x8`, and `gen x1x2 = x1*x2`.



Basic Information on Stata

- You should always work with do-files.
- All commands have to be written there within one row line (or use // to indicate that the commands continue in the following row line).
- Use the symbol for „execute selection“ in the task bar or strg + R to execute selected commands.
- Comments have to be characterized by asterisks (*).
- Some first commands:
 - . `tab varname` one-way table of frequencies
 - . `fre varname` one-way table of frequencies including value labels
 - . `sum varname` descriptive statistics
- Many commands can be abbreviated. For instance: you can use `sum` instead of `summarize`.
- Have a look on the help-files for additional information on commands and subcommands: `help command`.



Basic Information on Stata

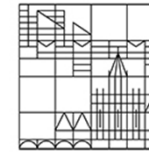
- You should always work with do-files.
- All commands have to be written there within one row line (or use // to indicate that the commands continue in the following row line).
- Use the symbol for „execute selection“ in the task bar or strg + R to execute selected commands.

- Comment commands with // These dots are only to mark commands in all Stata manuals. Do not write them in your do-file or into the command line!
- Some first commands are:

<code>. tab varname</code>	one-way table of frequencies
<code>. fre varname</code>	one-way table of frequencies including value labels
<code>. sum varname</code>	descriptive statistics

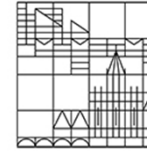
- Many commands can be abbreviated. For instance: you can use `sum` instead of `summarize`.
- Have a look on the help-files for additional information on commands and subcommands: `help command`.

Basic Information on Stata



<code>recode varname (value = value)</code>	recoding of variable
<code>gen newvar = ...</code>	generation of a new variable
<code>command if condition</code>	selection of cases for which the if-condition is true
<code>cd path</code>	change working directory

- Within the do-file, you can add comments also by using "//", e.g. .
`fre vurteil // table of frequencies, vignette judgm.`



Basic Information on Stata

- Sub-commands have to be written after a comma and can be looked-up within the help files. They are marked there with brackets ([]). For instance:

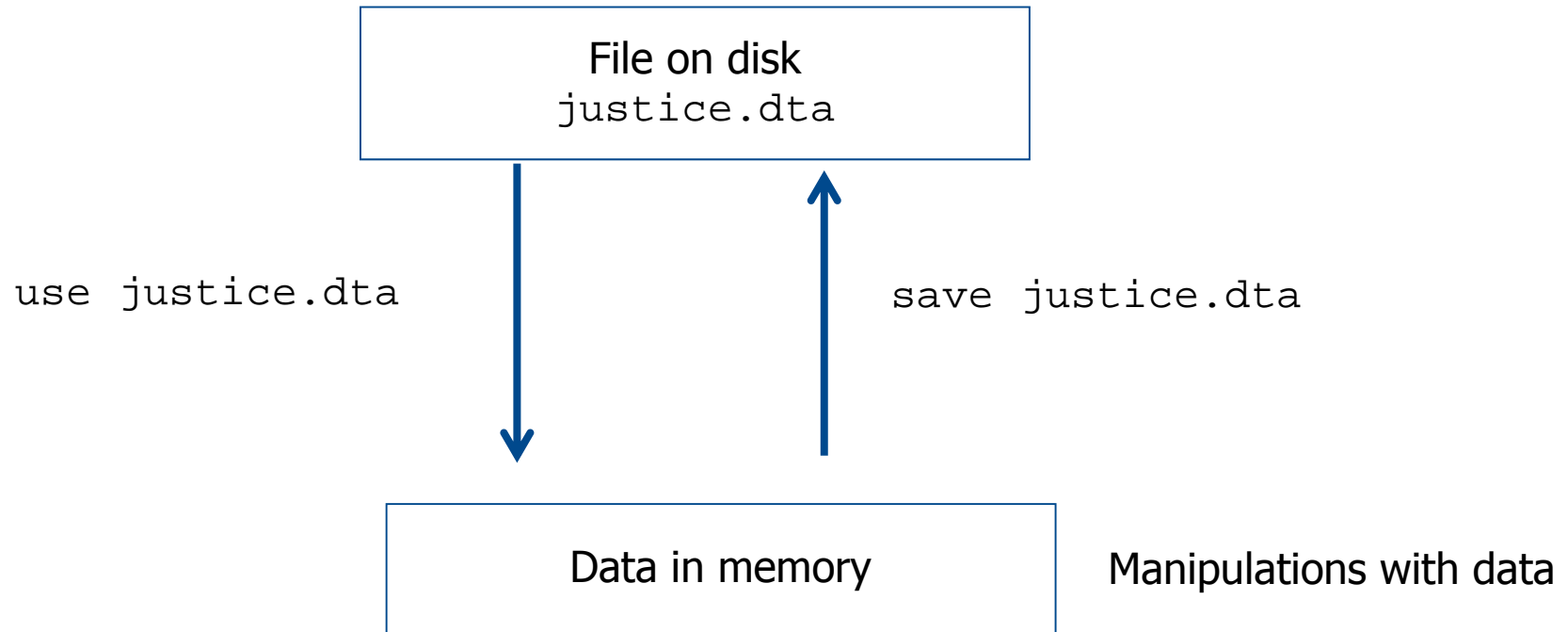
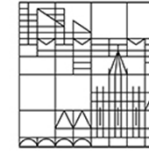
```
sum varname, detail // additional statistics
```

- But note: all if-conditions have to be written before the comma (as they belong to the main-command:

```
command if condition, subcommand
```

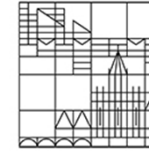
- Missing values are represented as as dot (.). These values are treated by Stata like infinite large numbers. This has to be considered when recoding variables!
- Upper- and lower-case letters make a difference!
- After making some changes, please store data sets under new names, and keep always some copies of your row data!

Do not destroy raw data!



- Risk to destroy good data by mistake!
- Though, change working directories, and work with do-files!

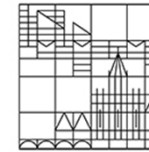
Source: Stata script by Svend Jooul



- Afternoon:
 - Independent work on exercises (FS design, Stata-introduction)
 - Individual support by Katrin Auspurg & Thomas Hinz: 14.00-16.00.
 - Please save your FS design and be prepared to present & discuss it with other course members during the next days.
- Tomorrow:
 - We will start at 9.00.
 - Morning session: Lecture on sampling of vignettes.
 - Individual learning/afternoon: Generation of vignette samples and vignette texts.

Any Questions? Comments? Suggestions?

References



- [1] Rossi, Peter H./Anderson, Andy B. (1982): The Factorial Survey Approach: An Introduction. In: Rossi, Peter H./Nock, Steve L. (Eds.): Measuring Social Judgments. The Factorial Survey Approach. Beverly Hills: 15-67 (p.19).
- [2] Jasso, G., 2006: Factorial-Survey Methods for Studying Beliefs and Judgments. Sociological Methods and Research 34: 334-423.
- [3] Beck, Michael/Opp, Karl-Dieter (2001): Der Faktorielle Survey und die Messung von Normen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 53: 283-306.
- [4] Wason, Kelly D./Polonsky, Michael J./Hyman, Michael R. (2002): Designing Vignette Studies in Marketing. Australasian Marketing Journal 10: 41-58.
- [5] Alexander, Cheryl S./ Becker, Henry Jay (1978): The Use of Vignettes in Survey Research. Public Opinion Quarterly 42: 93-104.
- [6] Armacost, Robert L./Hosseini, Jamshid C./Sara A. Morris, Sara A./Rehbein, Kathleen A. (1991): An Empirical Comparison of Direct Questioning, Scenario, and Randomized Response Methods for Obtaining Sensitive Business Information. Decision Science 22: 1073-1090.
- [7] Wittink, Dick R., Lakshman, Krishnamurthi/Reibstein, David J., 1989: The Effects of Differences in the Number of Attribute Levels on Conjoint Results. Marketing Letters 2: 113-123.