The ofness and aboutness of survey questions

Improved indexing of social science data

Tanja Friedrich and Pascal Siegers, Data Archive for the Social Sciences at GESIS
Outline

• Introduction: Indexing survey data
• Ofness and aboutness according to Sara Shatford
• The ofness and aboutness of survey data
• How to index ofness and aboutness
• Ofness and aboutness in retrieval
• Conclusion
Outline

• Introduction: Indexing survey data
• Ofness and aboutness according to Sara Shatford
• The ofness and aboutness of survey data
• How to index ofness and aboutness
• Ofness and aboutness in retrieval
• Conclusion
Introduction: Indexing survey data

• GESIS is documenting and distributing survey data for reuse

• Holdings are described in the data catalogue
ZA5800: International Social Survey Programme: Health and Health Care - ISSP 2011

Abstract


Themes: satisfaction with life (happiness); confidence in the educational system and the health system of the country; changes of health care system is needed; justification of better medical supply and better education for people with higher incomes; assessment of the health care system of the country (scale: estimation of improvement of the health care system, usage of health care services more than necessary, government should provide only basic health care services, inefficient health care system); willingness to pay higher taxes to improve the level of health care for all people in the country; attitude towards public funding of: preventive medical checkups, treatment of HIV/AIDS, programs to prevent obesity and conduct organ transplants; attitude towards the access to publicly funded health care for people without citizenship of the country and even if they behave in ways that damage their health; estimated part of people without access to the health care system; causes of severe health problems (behavior that damages health, environment, genes, poverty); evaluation of patients for smoking habits, age and the presence of young children for a needed heart operation; attitude towards alternative (traditional or folk) medicine (provides better solutions for health problems than...
Introduction: Indexing survey data

• Study descriptions are very detailed, but not indexed with a controlled vocabulary
Introduction: Indexing survey data

- Study descriptions are very detailed, but not indexed with a controlled vocabulary
- Our goal is to index our data with a social science thesaurus for improved retrieval
Introduction: Indexing survey data

• Our approach: user-centred indexing

• Point of departure: How are users looking for survey data?
Introduction: Indexing survey data

- Users are not reusing entire surveys
Introduction: Indexing survey data

• Users are not reusing entire surveys
• Users are looking for data on the social science constructs that they are interested in
Introduction: Indexing survey data

- Users are not reusing entire surveys
- Users are looking for data on the social science constructs that they are interested in
Introduction: Indexing survey data

• Thus, users are not looking for entire surveys, but for results that fit their own research questions

• They find them in particular measurements of constructs that have been used in a survey

• These measurements are reflected in the survey questions that have been asked
Introduction: Indexing survey data

- For indexing, that means:
  - the indexing level is the question (or variable) level
  - the primary information source is the questionnaire
Introduction: Indexing survey data

• For indexing, that means:
  – the indexing level is the question (or variable) level
  – the primary information source is the questionnaire
Introduction: Indexing survey data

• Main challenge for indexing social science constructs from questionnaires:
  – While the measurement can be derived from the questionnaire, the measured construct oftentimes cannot (latent constructs)…
  – … because constructs themselves are not measurable but have to be broken down into measurable properties (operationalization)
Introduction: Indexing survey data

• For Example, a construct such as *education* is not in itself *measurable*, it has to be broken down into its measurable properties

• “After all, no-one is measuring ‘a table’ […], but only a specific aspect: e.g. ‘the height of the table’ […]” (Schnell et al. 2011, 119, original in German)
Introduction: Indexing survey data

• Creating measurements from constructs is done by operationalization

• Operationalization is the process of specifying the range of observable objects representing the theoretical construct. It defines the measurement of the latent constructs. (cf. Schnell et al. 2011, 121)
Introduction: Indexing survey data

→ Conclusion: due to operationalization we have two subject layers in surveys: constructs and their measurements.

→ How to capture them both in indexing?
Outline

• Introduction: Indexing survey data
• Ofness and aboutness according to Sara Shatford
  • The ofness and aboutness of survey data
  • How to index ofness and aboutness
  • Ofness and aboutness in retrieval
• Conclusion
Ofness and aboutness according to Sara Shatford

- To analytically capture the two subject levels of survey data, we draw on Sara Shatfords theory on indexing pictures from 1986.
Ofness and aboutness according to Sara Shatford

• Sara Shatford (Shatford 1986, Shatford Layne 1994) developed a theory of indexing pictures, drawing on different aspects and attributes that are particularly important for the indexing of pictures.

• Of particular importance are the aspects of ofness and aboutness.
Ofness and aboutness according to Sara Shatford

• The subject of a picture according to Shatford:
  – Ofness: referring to the concrete and objective subject (or factual meaning)
  – Aboutness: referring to the abstract and subjective subject (or expressional meaning)
Ofness and aboutness according to Sara Shatford

“For example, an allegorical image might be of a man and a lion, but be about pride [...]);
or an image of a person crying might be about sorrow.”

(Shatford Layne 1994, 584)
Ofness and aboutness according to Sara Shatford

To sum up …

… **ofness** refers to what is visibly depicted in the picture, while

… **aboutness** refers to an intended meaning that is not visible in the picture, but identifiable on the grounds of **world knowledge**
Outline

• Introduction: Indexing survey data
• Ofness and aboutness according to Sara Shatford
  • The ofness and aboutness of survey data
  • How to index ofness and aboutness
  • Ofness and aboutness in retrieval
• Conclusion
The ofness and aboutness of survey data

• Shatford’s theory can help us to identify the subjects of survey data in order to improve
  – indexing of data and
  – retrieval of data
The ofness and aboutness of survey data

• The aboutness of survey data refers to the construct that the primary investigator intended to investigate.
The ofness and aboutness of survey data

• The ofness of survey data is the literal manifestation of the operationalization of this construct, as it is found in the questionnaire.
The ofness and aboutness of survey data

• Example: *During the last week, how often have you been talking to your neighbour about politics?* (Klingemann/Mochmann 1975, 185, original in German)
The ofness and aboutness of survey data

– Ofness: respondents tell us *of* their talking about *politics* with their *neighbour*

– Aboutness: respondents tell us *about* their *political interest* and *about* their relationship to *neighbours*
The ofness and aboutness of survey data

• Summing up ofness and aboutness of survey data:
  – Ofness: concrete or objective aspects of the measurement found in the questionnaire wording
  – Aboutness: the abstract or subjective aspects of the measurement (even if they are not in the question wording)
The ofness and aboutness of survey data

• Why is it important to index aboutness?
• By indexing aboutness, we access the **construct level** of what is being measured
• Fits search behavior of (secondary) users
• Facilitates search if very different question wordings refer to the same construct
The ofness and aboutness of survey data

• Example for the ofness of survey data
  – A country’s cultural life is undermined by immigrants (European Values Study 2008, ZA4800)
  – Ofness: CULTURE; IMMIGRATION
The ofness and aboutness of survey data

• Same example for the aboutness of survey data:
  – A country’s cultural life is undermined by immigrants (European Values Study 2008; ZA4800)
  – Aboutness: ATTITUDE; MIGRANT
Q78 Please look at the following statements and indicate where you would place your views on this scale?

A. Immigrants take jobs away from natives in a country
   1 2 3 4 5 6 7 8 9 10
   Immigrants do not take jobs away from natives in a country
   DK NA
   88 99

B. A country’s cultural life is undermined by immigrants
   1 2 3 4 5 6 7 8 9 10
   A country’s cultural life is not undermined by immigrants
   DK NA
   88 99

C. Immigrants make crime problems worse
   1 2 3 4 5 6 7 8 9 10
   Immigrants do not make crime problems worse
   DK NA
   88 99

D. Immigrants are a strain on a country’s welfare system
   1 2 3 4 5 6 7 8 9 10
   Immigrants are not a strain on a country’s welfare system
   DK NA
   88 99
The ofness and aboutness of survey data

- Identifying aboutness of survey data, is not easy …
The ofness and aboutness of survey data

• Identifying aboutness requires …
The ofness and aboutness of survey data

• Identifying aboutness requires …

… to examine the context of the question,
The ofness and aboutness of survey data

• Identifying aboutness requires …
  … to examine the context of the question,
  … to have field knowledge
The ofness and aboutness of survey data

• Identifying aboutness requires …
  … to examine the context of the question,
  … to have field knowledge
  … and data literacy
The ofness and aboutness of survey data

• Identifying aboutness requires …
  … to examine the context of the question,
  … to have field knowledge
  … and data literacy

→ This corresponds to the knowledge needed to index the aboutness of a picture (Shatford 1984)
Outline

• Introduction: Indexing survey data
• Ofness and aboutness according to Sara Shatford
• The ofness and aboutness of survey data
• How to index ofness and aboutness
• Ofness and aboutness in retrieval
• Conclusion
How to index ofness and aboutness

We apply a concept of syntactic indexing, based on the theory of ofness and aboutness:

• **Ofness** can be captured by a simple combination of thesaurus terms *(term linking)*
• **Aboutness** can be captured by combining directive and subject terms *(role operators)*
How to index ofness and aboutness

• **Term linking** and **role operators** are known methods of syntactic indexing (e.g. Lancaster 1998)

• We combine subject terms (e.g. **CORRUPTION**) with directive terms (e.g. **ATTITUDE, EXPERIENCE**).
How to index ofness and aboutness

• The term linking and role operators allow for identifying measurable constructs with the indexing, like
  – attitudes towards corruption
  – experience with corruption
How to index ofness and aboutness

Social Science Construct

Contents/topics (subject)
- any subject area relevant in social science
- E.g. work, family religion, education

Attributes (direction):
- Cognition
- Evaluation
- Affection
- Action
- [objective characteristics]

Measurable unit
How to index ofness and aboutness

**Subject terms**
- Specify the contents of the measurement
- As specific as possible
- Combinations of terms, if necessary

**Directive terms**
- Specify the attributes of the measurement
- Limited heterogeneity in directive terms to facilitate faceted retrieval
How to index ofness and aboutness

Examples for directive terms:

- **Cognition**
  - PERCEPTION
  - KNOWLEDGE
  - AWARENESS
  - INTEREST
  - BELIEF
  - ORIENTATION

- **Evaluation**
  - ATTITUDE
  - PREFERENCE
  - JUDGMENT
  - PREJUDICE
  - SATISFACTION
  - ACCEPTANCE/APPROVAL
  - REJECTION/REFUSAL

- **Affection**
  - MOOD
  - FEAR
  - ANGER/ANNOYANCE
  - HAPPINESS
  - HATE
  - LOVE

- **Action**
  - BEHAVIOR
  - USE/UTILIZATION
  - CHOICE
  - EXPERIENCE
  - INTERACTION
  - COMMUNICATION
  - CONTACT
How to index ofness and aboutness

- Linking *subject terms* captures ofness; linking them with *directive terms* reveals the aboutness
- Each *directive* and *subject* term combination represents a *measurable unit* of interest to the secondary researcher
How to index ofness and aboutness
How to index ofness and aboutness

Measurable Unit (e.g. survey question)
How to index ofness and aboutness

Measurable Unit (e.g. survey question)

subject term(s) (ST)  directive terms (DT)
How to index ofness and aboutness

Measurable Unit (e.g. survey question)

subject term(s) (ST)  

Precoordination/syntactic indexing  = linked terms that are specified by role operators

directive terms (DT)
How to index ofness and aboutness

Measurable Unit (e.g. survey question)

subject term(s) (ST)  directive terms (DT)

Precoordination/syntactic indexing
= linked terms that are specified by role operators

Aboutness of survey questions
How to index ofness and aboutness

„There is corruption in the national public institutions in Germany.“ (Eurobarometer 76.1; ZA5565)

Directive term: PERCEPTION
Subject term(s): CORRUPTION, PUBLIC INSTITUTIONS
Syntactic Indexing: PERCEPTION; CORRUPTION; PUBLIC INSTITUTIONS

„Are you personally affected by corruption in your daily activities?“ (Eurobarometer 76.1; ZA5565)

Directive term: EXPERIENCE
Subject term(s): CORRUPTION; EVERYDAY LIVE
Syntactic Indexing: EXPERIENCE; CORRUPTION; EVERYDAY LIVE
How to index ofness and aboutness

„There is corruption in the national public institutions in Germany.“ (Eurobarometer 76.1; ZA5565)

Directive term: PERCEPTION

Subject term(s): CORRUPTION, PUBLIC INSTITUTIONS

Syntactic Indexing: PERCEPTION, CORRUPTION, PUBLIC INSTITUTIONS

„Are you personally affected by corruption in your daily activities?“ (Eurobarometer 76.1; ZA5565)

Directive term: EXPERIENCE

Subject term(s): CORRUPTION, EVERYDAY LIVE

Syntactic Indexing: EXPERIENCE, CORRUPTION, EVERYDAY LIVE
How to index ofness and aboutness

„There is corruption in the national public institutions in Germany.“ (Eurobarometer 76.1; ZA5565)

Directive term: PERCEPTION

Subject term(s): CORRUPTION, PUBLIC INSTITUTIONS

Syntactic Indexing: PERCEPTION; CORRUPTION; PUBLIC INSTITUTIONS

„Are you personally affected by corruption in your daily activities?“ (Eurobarometer 76.1; ZA5565)

Directive term: EXPERIENCE

Subject term(s): CORRUPTION; EVERYDAY LIVE

Syntactic Indexing: EXPERIENCE; CORRUPTION; EVERYDAY LIVE
How to index ofness and aboutness

„The practice of Islam should be restricted in Germany.“ (ALLBUS 2012; ZA4614)

Directive term: ATTITUDE
Subject term(s): ISLAM; RELIGIOUS FREEDOM
Syntactic Indexing: ATTITUDE; ISLAM; RELIGIOUS FREEDOM

„The presence of Muslims in Germany leads to conflicts.” (ALLBUS 2012; ZA4614)

Directive term: PERCEPTION / EXPECTATION
Subject term(s): MUSLIM; RELIGIOUS CONFLICT
Syntactic Indexing: PERCEPTION; ISLAM; RELIGIOUS CONFLICT
Outline

• Introduction: Indexing survey data
• Ofness and aboutness according to Sara Shatford
• The ofness and aboutness of survey data
• How to index ofness and aboutness
• Ofness and aboutness in retrieval
• Conclusion
Ofness and aboutness in retrieval

• The system of syntactic indexing allows for various faceting mechanisms in retrieval:
  – We can offer facets by our two role operators (subject and directive terms), e.g. as “Refine by subject” and “Refine by intention”
  – We can offer facets with linked terms that allow for retrieval of measurements, e.g. as “Refine by concept/construct”
### Refine by topic of study
- International politics (19)
- Conflict, security and peace (18)
- Society, culture (10)

### Refine by country
- USA (15)
- Israel (10)
- France (8)
- Australia (3)

### Refine by time period
- 2003 (10)
- 2012 (5)
- 2008 (5)
- 2010 (3)

### Refine by questions
- Middle East (20)
- Conflict (19)
- Israel (19)
- Peace (19)
- Palestinian State (19)
- USA (18)
- Egypt (17)

<table>
<thead>
<tr>
<th>Select</th>
<th>No.</th>
<th>Title</th>
<th>Collection Date</th>
<th>Order Download</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ZA4282</td>
<td>Transatlantic Trends 2005</td>
<td>30.06.2006 - 17.06.2005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZA4218</td>
<td>Worldviews 2001 (Transatlantic Trends)</td>
<td>06.2002 - 07.2002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZA2691</td>
<td>Panel of the University of the Arts in Berlin 1975 - 1995</td>
<td>08.1975 - 03.1995</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZA2446</td>
<td>Attitudes to Security Policy in the Federal Republic (June 1986)</td>
<td>06.1986 - 06.1986</td>
<td></td>
</tr>
</tbody>
</table>
Ofness and aboutness in retrieval

• Syntactic indexing at the variable level and faceting mechanisms in retrieval are useful for:
  • **question data bases**
    – Enabling search for specific measurements when designing questionnaires.
  • **multilingual documentation of questionnaires**
    – Enabling secondary use of foreign data if no English translation of the questionnaire is available.
Outline

• Introduction: Indexing survey data
• Ofness and aboutness according to Sara Shatford
• The ofness and aboutness of survey data
• How to index ofness and aboutness
• Ofness and aboutness in retrieval
• Conclusion
Conclusion

• Indexing survey data requires to consider different subject levels
• The theory of ofness and aboutness is suitable to capture the different levels
• Use of a syntax with term linking and role operators enables systematic indexing at the aboutness level
Conclusion

• Term linking and role operators enable us to make measurable units searchable for secondary researchers

• Syntactic indexing of survey data allows for sophisticated faceted searching
Thank you for your attention!

Tanja Friedrich
tanja.friedrich@gesis.org

Pascal Siegers
pascal.siegers@gesis.org
References


