**Assignment: collecting data using obtrusive and unobtrusive methods**

|  |  |
| --- | --- |
|  | **Unit 6:** In this assignment, you will learn about various methods of data collection, including surveys and observation. |
|  | [Types of data Collection Methods](https://vimeo.com/user29453510/review/136584709/98a0c4c08d)[Operationalization and Measurement](https://vimeo.com/user29453510/review/136585050/b1396c9bd9)[Data Collection - Survey](https://vimeo.com/user29453510/review/201096002/416a3c6416)Content analysis(see blackboard for links) |  | Concept aka termConceptualizationOperationalizationData collectionPrimary dataSecondary data(Un)obtrusive research(non)verbal measurementSurveyContent analysisCoding |
|  | Babbie Ch. 9 p. 247-261Babbie Ch. 10 p. 323- 334 |  |  |  | 90 minutes |
|  | Read this assignment carefully and answer the questions. Bring either a print or digital version to the lecture. |
|  | [feedback information]  |

**\*kan iemand die rare lijntjes weghalen uit de tekst hieronder? (en als dat gelukt is dit zinnetje ook ☺)**

1. A school wants to hire a new science teacher, and a panel of governors begins to look through the various candidates. They draw up a shortlist. All candidates are given a test, measuring their knowledge about physics. They decide to select the candidate with the highest score. Sadly, he proves to be an extremely poor science teacher.

1.1. Conceptualize the concept ‘quality of the science teacher’, what makes a person being a ‘good teacher’

Knowledge of the field (knows the field)
Able to explain well (able to explain)
Accessible and nice (kind)

1.2. Suppose you have to measure these aspects of being a good teacher. How would you do that?

example:

Knowledge of the field (knows the field): knowledge test
Able to explain well (able to explain): give a lesson assessed by students

Accessible and nice (kind): ask others? Meet a person privately?

1.3. How would you use these ‘traits’ to define the ‘quality of the teacher’?

Maybe having a minimum level of all aspects before someone is considered potentially good (so, at least knowing something about physics) and a simple index of all aspects to determine the quality of the remaining candidates? So, for example, rate all candidates on a five point scale, give all those which score 1-3 on any of the topics zero points (not suitable) and counting the total number of points for the remaining candidates.

2. Let’s examine the concept ‘feminism’.

2.1. Conceptualize the construct ‘feminism’ as a characteristic of a person.

For example: The belief that men and women should have equal rights and opportunities, and maybe also the extent to which that person is actively supporting women's rights on the grounds of political, social, and economic equality to men can be called ‘feminist’. If you conceptualize feminism as ‚support’ you would probably look for either opinions or behaviors, while i fit is conceptualized as a ‚belief’ you would probably look for opinions only (surveys or maybe content analysis of texts).

2.2. Write down (at least two) very different *methods of data collection* you could use to measure feminism of students or employees.

a. Survey with questions about attitudes

b. Observing behaviour in a meeting

c. Analysing texts written by the people you want to observe (content analysis)

2.3. Which of these methods is least obtrusive. Why?

Content analysis. People will probably give socially desirable answers to questions about their feminist attitude. Observing behavior is less obtrusive than surveys, but the presence of the researcher can still influence the behavior he/she is observing, so it very much depends on the exact mode of observation.

2.4. Suppose you want to measure the amount of feminism in a population, which of the methods you selected would you prefer? Why?

Content analysis and observing behavior is time-consuming, moreover, not all units have written texts or are observable. So a survey is most probably the more obtrusive method, but probably the best way to go forward.

2.5. In a recent survey[[1]](#footnote-1) feminism was measured with one single survey question: „Do you consider yourself to be a strong feminist, a feminist, not a feminist or an anti-feminist?” What do you think of this operationalization? How could you improve it?

The answering categories are not balanced (strong – somewhat – no – somewhat – strong would have been balanced). The survey question assumes everyone has the same understanding of the word ‚feminism’ (which is clearly not the case, see an article here: [https://www.bustle.com/articles/170721-7-things-the-word-feminist-does-not-mean)](https://www.bustle.com/articles/170721-7-things-the-word-feminist-does-not-mean%29). So maybe avoid the word ‚feminist’ in the survey question.

3. One way to collect data is by using texts from interviews, newspapers, tweets, business reports. Creating data using ‘texts’ is more and more often done by computers. However, in this assignment you do this by hand to get a better understanding of what ‘coding’ means. There are different approaches for using ‘texts’ to create ‘data’. One approach is sometimes called ‘deductive’: starting with clear units and variables and ‘finding’ indications of these variables in the texts. Another approach is more ‘inductive’: starting with the texts and ‘creating’ variables. In this assignment, we focus on the deductive approach only. For this assignment, we use a (very) small set of texts. This set of texts can be found in a separate document.

3.1. We want to study the extent to which Reuters gives attention to ‘the refugee crisis’ over time. What would be a clear (descriptive) empirical research question? Identify the units and the variable(s) of this question.

To what extent does the amount of attention Reuters gives to the refugee crisis change over time. Units: days (or any other time variable, weeks, months). One variable: amount of attention. You COULD argue that Reuters is the unit (of analysis).

3.2. How would you operationalize the variable of this question?

Count the number of newspaper articles per unit (so per day or week) and report the numbers of articles that mention the refugee crisis. You can even draw a graph with the units (days) ordered at the x-axis and the amount of attention on the y-axis.

3.3. We think that it is not just the amount of attention that changes, but also the content of the articles. It may be that there is a change from ‘focus on refugees’ to a ‘focus on policy’ (we call this variable ‘focus’). *Individually* code the articles using this dichotomy.

No comments here.

3.4. Compare your codes. Where do the differences come from? (This will be called later:”Inter coder reliability”).

We hope you become aware of the fact that the variable ‘focus’ is not clearly defined and operationalized. If an operationalization gives very different results, the operationalization is called ‘unreliable’. We will get back to ‘unreliability’ later. Inter coder reliability is one way to assess the reliability of an operationalization.

3.5. There are different ways to store these data in a data matrix. Describe your data matrix in terms of units and variables.

The simplest way would be to take each article as a unit (one row in the data matrix). The first *variable* could be the date, the second the dichotomy, the third all other variables you are interested in. This clarifies that ‘units’ are sometimes based on stored ‘variables’. Units are the objects you want to know something about (days), whereas the units of observation are the units stored in the most informative data set.

*End of assignment 6.*

1. (aka a poll) (see [https://www.washingtonpost.com/graphics/national/feminism-project/poll/)](https://www.washingtonpost.com/graphics/national/feminism-project/poll/%29) [↑](#footnote-ref-1)