

Quantitative data acquisition methods (e.g. surveys and tests)

(version 0.7, 1/4/05)

Code: data-quant

Daniel K. Schneider, TECFA, University of Geneva



Menu

1. Basics	2
2. Survey design	3
3. Experiments	9
4. Sampling	10

1. Basics

1. Make a list of concepts (theoretical variables) in your research questions for which you need data.
2. For each of these concepts make sure that you identify its dimensions (or make sure that they are not multi-dimensional)
 - Consult the literature
 - Discuss with domain experts
 - Read the literature or discuss with researchers if you can
 - Make a list of conceptual variables (either simple concepts or dimensions of complex concepts)
3. For each conceptual variable think about how you plan to measure it
 - First of all (!!) go through the literature and find out if and how other people went about it
 - It is much better to use a suitable published instrument than to build your own. You can then compare your results and you will have much less explanations and justifications to produce !
4. Think about redundancy and triangulation
 - Do not measure a conceptual variable with just one question or observation
 - Maybe get data of different nature, e.g. combine surveys with objective data and observations
5. Rather ask people how they behave instead of how they think they behave
 - E.g. don't ask: "Do you use socio-constructivist pedagogies ?"
 - Ask several questions about typical tasks assigned to students.
6. Do not ask people to confirm your research questions
 - E.g. Don't ask: Did you manage to make your teaching more socio-constructivist with this new tech.
 - (Again) Ask what the person really does

2. Survey design

2.1 The basics of question and response item design

- Only ask questions that your target population understands (test your questionnaire with at least 2 people)
- Questions should avoid addressing 2 issues in 1 question!
- Make questions short
- Ask several questions that measure the same concept
 - Try be all means to find sets of published items (questions) in the literature that you can reuse
- Avoid open-ended answers (these will give a lot of coding work)
- Use scales that have at least a range of 5 response options
 - otherwise people will have a tendency to drift to the "middle" and you will have no variance.
 - e.g. avoid:
agree () neither/or () disagree ()
- Response options should ideally be consistent across items measuring a same concept
- If you feel that most people will check a "middle" value, use a large "paired" scale without a middle point
e.g. 1=totally disagree, 10=totally agree
1 2 3 4 5 6 7 8 9 10

2.2 Examples

Exemple 2-1: Social presence

Gunawerda, C.N., & Zittle, F.J. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *The American journal of distance education*, 11(3), 8-26.

The GlobalEd questionnaire was developed to evaluate a virtual conference. Participants (n=50) of the conference filled out the questionnaire. Internal consistency of the social presence scale was $\alpha=0.88$. Social presence was found to be a strong predictor of user satisfaction.

url: <http://www.presence-research.org/Overview.html> (Compendium of Presence Measures)

- Messages in GlobalEd were impersonal
- CMC is an excellent medium for social interaction
- I felt comfortable conversing through this text-based medium
- I felt comfortable introducing myself on GlobalEd
- The introduction enabled me to form a sense of online community
- I felt comfortable participating in GlobalEd discussions
- The moderators created a feeling of online community
- The moderators facilitated discussions in the GlobalEd conference
- Discussions using the medium of CMC tend to be more impersonal than face-to-face discussion
- CMC discussions are more impersonal than audio conference discussions
- CMC discussions are more impersonal than video teleconference discussions
- I felt comfortable interacting with other participants in the conference
- I felt that my point of view was acknowledged by other participants in GlobalEd
- I was able to form distinct individual impressions of some GlobalEd participants even though we communicated only via a text-based medium.

A 5-point rating scale was used for each question

Exemple 2-2: Socio-constructivist teachers

Class (and Schneider) 2005, PhD thesis project (scale based on Dolmans 2004)

The problem was how to identify socio-constructivist elements in a distance teaching course for interpreter trainers.

Decomposition of “socio-constructivist design” in (1) active or constructive learning, (2) self-directed learning, (3) contextual learning and (4) collaborative learning, (5) teacher’s interpersonal behavior (according to Dolmans et al., 1993)

Note that headers regarding these dimensions (e.g. "Constructive/active learning" are not shown to the subjects. We do not want them to reflect about theory, but just to answer the questions ... So they are just shown below to help your understanding

Statements: Teachers stimulated us ...		Totally disagree	Disagree	Somewhat agree	Agree	Totally agree
		1	2	3	4	5
(Constructive/active learning)						
4	... to search for explanations during discussion	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	... to summarize what we had learnt in our own words	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6	... to search for links between issues discussed in the tutorial group	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7	... to understand underlying mechanisms/theories	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8	... to pay attention to contradictory explanations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
(Self-directed learning)						

	Statements: Teachers stimulated us ...	Totally disagree	Disagree	Somewhat agree	Agree	Totally agree
9	... to generate clear learning issues by ourselves unclear	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10	... to evaluate our understanding of the subject matter by ourselves	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
(Contextual learning)						
11	... to apply knowledge to the problem discussed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12	... to apply knowledge to other situations/problems	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13	... to ask sophisticated questions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14	... to reconsider earlier explanations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
(Collaborative learning)						
15	... to think about our strengths and weaknesses concerning our functioning in the tutorial group	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16	... to give constructive feedback about our group work	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17	... to evaluate our group cooperation regularly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18	... to arrange meetings with him/her to discuss how to improve our functioning as a group	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2.3 The questionnaire

The Introduction in written questionnaires (on paper or on-line)

- Write a short introduction that states the purpose of this questionnaire
 - Example (in french): The purpose of this questionnaire is to help us understand how well you liked the on-line delivery of the first blended edition of the Certificate for Interpreter Trainers. Each one of the 117 statements below asks about your experience in the on-line part of the Certificate. Data will be processed and published only statistically. The following questions will be dealt with: personal information, teachers' behavior, learning environment, tutoring support structure, tools and skills. Filling in this questionnaire will take you about 20 minutes. Please be assured that your responses will be treated confidentially, and that they will not affect your assessment. Thank you very much for your cooperation.
- guarantee that you only will publish statistical data (no names !)
- specify how long it will take to fill it in

Coding information for the researcher

- Assign a code (e.g. number) to each question item (variable) and assign a number (code) to each response item
 - see example on previous and next slide
- This will help you when you transcribe data or analyze data
- use "small fonts" (this information is for you)

Ergonomics

- Do not include anything else than questions and response items (besides the introduction)
- Make sure that people understand where to "tick".

Exemple 2-3: Example of a question set (one question with several sub questions):**Teachers' behavior**

Below you will find general statements about teachers' behavior. Please indicate to what extent you agree or disagree with them? Please tick the appropriate circle on the scale (totally disagree - totally agree) for each statement.

Statements: Teachers stimulated us ...		Totally disagree	Disagree	Some-what agree	Agree	Totally agree
		1	2	3	4	5
4	... to search for explanations during discussion	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	... to summarize what we had learnt in our own words	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14	... to reconsider earlier explanations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. Experiments

- Designing a true experiment needs advice from some expert. Typically, a beginner makes the mistake to differentiate 2 experimental conditions by more than one variable !!

There are many kinds of experimental measures

- observations (e.g. Video, or recordings of computer input)
- tests (similar to surveys)
- tests (similar to examination questions)
- tests (performance in seconds)
- tests (similar to IQ tests)

Consider all the variables you want to measure

- most times the dependant variables (to explain) are measured with tests
- usually the independent (explanatory) variables are defined by the experimental conditions (so you don't need to measure anything, just remember to which experimental group the subject belonged)

See the literature !!

- First of all, read articles about similar research !
- Consult test psychologists if you need to measure intellectual performance, personality traits, etc.
- Use typical school tests if you want to measure typical learning achievement

4. Sampling

4.1 The ground rules

The number of cases you have to take into account is rather an absolute number

- therefore not dependent on the size of the total "population" you study

The best sampling is random, because:

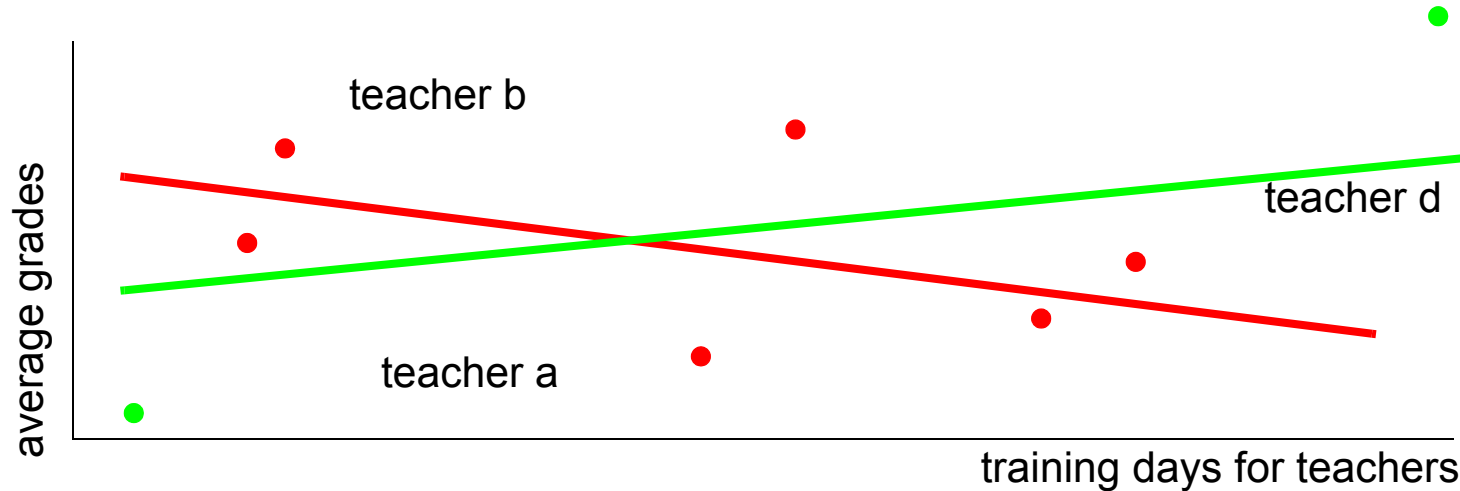
- you have a likely chance to find representatives of each "kind" in your sample
- you avoid auto-selection (i.e. that only "interested" persons will answer your survey or submit to experiments)

When you work with small samples, you may use a quota system

- e.g. make sure that you have both "experts" and "novices" in a usability study of some software
- e.g. make sure that you (a) both interview teachers who are enthusiastic users and the contrary, (b) schools that are well equipped and the contrary in a study on classroom use of new technologies.

4.2 A first look at significance

- Significance of results depend both on strength of correlations and size of samples
 - therefore: the more cases you have got, the more likely your results will be interpretable !



- Let's assume you have data from only 6 teachers (the red dots):
 - your data suggest a negative correlation: more training days lead to worse averages
- By only adding 2 teachers (the 2 green dots) this relation will switch from negative to positive
 - data suggests a (weak) positive correlation.
 - So: doing a statistical analysis on very small data sets is like gambling. If your data set included 20 teachers, adding these 2 more wouldn't have changed the relationship

4.3 Typical sampling for experiments

- **preferably** 20 subjects / experimental condition
- **at least** 10 / experimental condition (but expect most relations to be non-significant)

Exemple 4-1: Study the effect of multimedia on retention

Explanatory (independent) variable X: **Static diagrams** vs. **animation** vs. **interactive animation**

Dependant (to be explained) variable Y1: Short term recall

Dependant (to be explained) variable Y2: Long-term recall

- Both dependant variables (Y1 and Y2) can be measured by recall tests
- For variable X we have three conditions
- Therefore we need $3 * 20 = 60$ subjects
 - If you expect very strong relations (don't for this type of research !) you can get away with $3 * 15$
 - Note: we can not administer the three different conditions to each individual (because by moving from one experiment to another they will learn). You may consider building $3 * 3 = 9$ different kinds of experimental materials however and have each individual do each experiment in a different condition. However, they may get tired or show other experimentation effects ... and producing good material is more expensive than finding subjects.

4.4 Typical samples for survey research

- As much as you can if you use written or on-line surveys
- 50 or more, 40 is a minimum, 100 is good and 200 is excellent for a MSc thesis.
 - otherwise you can't do any sort of interesting data analysis, because your significance levels will be too high (i.e. bad) when you analyze detailed or complex relationships.

4.5 Typical samples for aggregate data

- e.g. schools, districts, countries etc.
- Since these data reflect real "realities" you can do with less (however talk to an expert, a lot depends on the kinds of analysis you plan to do).

