

Sampling from a universe of items and the De-Machiavellization of questionnaire design

Mohler, Peter Ph.

Veröffentlichungsversion / Published Version

Sammelwerksbeitrag / collection article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Mohler, P. P. (2006). Sampling from a universe of items and the De-Machiavellization of questionnaire design. In M. Braun, & P. P. Mohler (Eds.), *Beyond the horizon of measurement: Festschrift in honor of Ingwer Borg* (pp. 9-14). Mannheim: GESIS-ZUMA. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-49166-2>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

SAMPLING FROM A UNIVERSE OF ITEMS AND THE DE-MACHIAVELLIZATION OF QUESTIONNAIRE DESIGN

PETER PH. MOHLER

***Abstract:** Up-to-date design and development of survey items is like a medieval rite rather than a quality-controlled scientific undertaking. This holds despite existing evidence-based protocols propagated for many years by methodologists. Most prominent is the creation process of items, which in general is not a systematic selection from a universe of items belonging to a specific construct. A case is made here, to give up the round-table ritual of item creation in favor of a well-defined selection process. A differentiation of work roles is advocated in the light of knowledge accrued in other disciplines to streamline the item design and selection process for future surveys.*

Sampling from a universe of items

Whenever survey researchers meet to design and develop a questionnaire they open their gathering with a ritual of sorts. Imagine a setting with a number of eminent survey researchers sitting around a table heading for a new survey instrument (questionnaire). Usually the chair sets out to declare that one will apply most rigorous methods and strive for highest standards. His opening statement is followed by a round of positive murmur from all sides accompanied by the culturally acceptable body language, for example, nodding heads into all directions. Then the importance of clear conceptual delineation is stated combined with a pledge to closely connect concepts with relevant theories.

A world-famous saying in Germany says “von nun an ging’s bergab” (Hammerschmidt & Knief 1967) (from now on it went downhill), because what follows will be hours, days, and often weeks of entrenched defense of chairs’ and other eminent members’ beloved ‘tried and tested’ items/questions. The sometimes heavy battlefield-like inter- and counteractions are ‘interluded’ by a relaxing, enjoyable activity called ‘item tinkering’. After a round of serious, sometimes heated debates whether to choose an item or not, no one can deny the fun of rephrasing items as a common effort to re-establish group cohesion. One

might add, tinkering is also the last resort for opponents to ridicule an item by absurdly rephrasing it. However, as it is the case with hierarchical academic structures, it is the chair who, inevitably, will claim to have the last word.

Apart from the fact that this Round Table Approach is neither in nor according to the books, tinkering or rephrasing points to the strong conviction that one could really improve the measurement property of items just by phrasing them more elegantly. According to the textbooks of item design one can, of course, not endorse this approach.

Rephrasing of items takes advantage of a language property we all know of, but hardly ever use consciously: language actually is infinite. Similar to the natural numbers where one always can add one more number (not countable but enumerable), there is, theoretically, no end to words. Thus there is nothing such as the 'last word'. Although, the universe of survey items appears to be limited compared to every-day language it is also infinite in theory. Just think of the rather straightforward concept AGE, i.e., the time span between today and a person's birth date. Just think of the multitude of ways to ask for AGE:

- How old are you?
- When were you born?
- What is your birth date?
- Could you please tell me the year you were born in?
- Would you like to tell me your age?
- Would you, please, like to tell me your age?
- May I, please, ask for your age?
- May I right now, please, ask for your age?

And so on and so forth, one could add more words and more words, change words, replace words until the end of all ages. However, this tinkering exercise will quickly become boring and tedious. While language in principle is infinite as is our universe, the actual usable universe of items could be fairly small as it is the case with AGE (see also Borg 1992: 16).

The task is then to delineate for all practical purposes a finite universe¹ similar to defining a population of respondents. The next step then is to select or sample items from the universe or population of items. The very notion of ‘sampling’ items from a universe of items seems to counter every single design principle in survey research. But this holds only, if one does not look over the fence to psychologists and their long tradition in instrument design. They take great efforts to collect all possible items, including variations of wordings. A selection is then made following strict rules of item quality testing. The resulting scales (item batteries with well-defined properties) may contain some hundred items as well as one item only. Other than in psychological testing, survey methodologist up-to-date did not opt for commonly agreed-upon protocols for questionnaire design (cf. international testing association). The listings of ‘rules for good question wording’ (Groves et al. 2004: 226-235) or textbooks full of ‘good and bad’ examples cannot count here as well-defined selection criteria as stated in the relevant psychological literature. It is timely that survey methodology adopts a more strict position, i.e., well-defined procedures and rules for the selection of items similar to sampling strategies for selecting respondents.

A scientific sample of respondents is said to be representative of the population under observation. Populations are groups of people living at a specific time in a defined region, belonging to a specific societal stratum, sharing specific characteristics, etc. Similarly, populations or universes of items can be defined as groups of items belonging to a specific theoretical concept (indicators measuring an unobservable construct which itself represents a theoretical concept). Having said this, the question may rise to what extent would such an approach be helpful in survey research (apart from the rather rare case where sociologists and political scientists borrow scales from psychologists)? One might also ask: Isn’t this simply a rather academic description what underlies the ritual sketched out above?

Such questions become obsolete, if one adds one more requirement, namely a strictly formal structure, best a predictive one. Items selected must correlate with other items of a scale or other characteristics of interest in a well-defined way, as is the case in Facet Theory.

1 The term ‘finite universe’ deviates from the contemporary concept of our universe, which is seen being infinite (Rees 2000).

Thus three steps are involved for a scientific sampling of items:

1. Definition of the population of items (finite universe of items)
2. Census of the item population
3. Sampling of items from the population according to well-defined rules.

However, the concept of a universe of items and the necessity to sample items from that universe is only half the story about modern, evidence-based questionnaire design. Because almost all survey questionnaires are composites of items/questions selected from a number of universes, additional procurements must be taken. For instance, a survey may target political participation. Thus, it will contain a sample of items selected from the universe of political-participation items. In addition, researchers want to know about the social setting of respondents. Thus, they will need a sample of items from the universe of social settings (often called ‘background variables’). They also might be interested in values held by respondents and their personal traits. Thus, they will have to include item samples of the respective universes. Designing the questionnaire for a social survey is thus much more complex than to develop a ‘one universe’ test instrument. The increase in complexity, scope(s), and technology asks for a radical different approach than the Round Table mentioned before. Instead, an evidence-based professional route has to be taken. This in turn will change the Round Table Approach into a quality-controlled survey production process. In the following we will sketch out this new setting.

De-Machiavellization of questionnaire design

Let us come back to the introductory scene. What would be needed to turn the academic discourse and debate into a scientific procedure? The first new aspect would be that everyone in that room is aware that there are always several strategies at hand how to define item populations, create a good census of the item populations and the various sampling strategies possible. Secondly, the discussion will concentrate on which procedural approach will be optimal at each production step. After having decided upon this, the roles of the different players will be clearly assigned (see Noelle-Neumann & Petersen 2005: 110). Finally, the quality controls will be agreed, which allows to test whether the design process successfully leads to quality items.

Defining specific professional roles of participants and the implementation of a quality control process would be the major transition from an academic ritual to an up-to-date scientific endeavor. Because proper questionnaire design and item development requires different skills, talents, and knowledge on different stages we need at least nine professional roles:

1. excellent knowledge of the subject area (necessary for the definition of the item population),
2. institutional and cultural expertise (necessary to identify culturally or institutionally unacceptable item formulations),
3. linguistic knowledge (necessary to identify odd phrases),
4. texter skills (talent for creative writing like in marketing),
5. measurement and statistical expertise (to inform others involved in the design process about desirable measurement properties and the resulting analytical possibilities),
6. cognition & survey research know-how (to check for unintended effects),
7. information-technology know-how (searching and identifying of existing items),
8. documentation knowledge (for 'real time' production process documentation),
9. last but not least survey-process quality know-how (to inform developers and users about the level of quality achieved).

There is no single person who rightfully could claim to master all the knowledge, know-how, skills or talents for all roles simultaneously. Hence, the 'Round Table Approach' is obsolete in a modern Survey Methodology setting. It will be replaced by a Team Approach, where roles and decision power are clearly set out from the very beginning of the design process. Forming a quality team-production process is, of course, nothing new under the sun in areas outside of survey research (Deming 1986). It will be, however, a revolution to survey research. Instead of a strict hierarchical order with a Principal Investigator acting as Machiavelli's Principe, responsibilities and decisions are now distributed across the questionnaire production chain. In this setting, the Principal Investigator becomes a CEO who is responsible for the overall strategy, i.e., definition of the item population. She will delegate the production process to her team of specialists. As in every good industrial production process, no one can overrule a decision made by anyone in the production chain which is based on protocols agreed before. This will create a balance of power which will enforce Principal Investigators becoming informed by team members, instead of 'having the last word' as an authoritarian academic.

Team-Work Questionnaire Design will be started by a first team which looks into the subject area(s) of the questionnaire and the analyses targeted. Inevitably, this first team will observe a holistic view on the whole production process (i.e., link up with the study

design team (overall survey strategy)) and other teams such as the sample-design team or the data-collection team. The first team will hand definitions of relevant item universes to item-selection teams which will carry out the other eight tasks described above. The finalized instrument will then be handed to the data-collection team which in turn will deliver the data to data editors and finally to data analysts. On each step proper documentation will identify the quality level achieved.

What, hopefully, will not happen in Team-Work Questionnaire Design are the four deadly sins of the Round Table Approach:

1. Items are selected on the sole reason of previous usage, not because of specific measurement properties relevant for the current study.
2. Items are tinkered over and over again without any evidence of improved measurement quality.
3. Mis-designed or badly formulated items make it into pretests (for the sole purpose to convince il Principe).
4. Questions from different item universes are just 'glued', instead of the questionnaire being 'durchkomponiert' (i.e., completely designed).

Finally, one might ask whether all the fun of questionnaire development will be gone with tinkering in favor of a technically machinery? Of course not, it is just a change from a group 'catch as catch can' to a well orchestrated team effort and fun.

References

- Borg, I. (1992). *Grundlagen und Ergebnisse der Facettentheorie*. Bern: Huber.
- Deming, W. E. (1986). *Out of the Crisis*. Cambridge/MA: MIT-Press.
- Groves, R. M., Fowler, F. J., Couper, M., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey Methodology*. Hoboken/NJ: John Wiley and Sons.
- Hammerschmidt, H., & Knef, H. (1967). *Von nun an ging's bergab*. In: Halt mich fest (LP-Vinyl): Funkturm Verlag.
- Noelle-Neumann, E., & Petersen, T. (2005). *Alle, nicht jeder. Einführung in die Methoden der Demoskopie* (5th edition). Heidelberg: Springer.
- Rees, M. (2000). *Just Six Numbers. The Deep Forces that Shape the Universe*. New York: Basic Books.