LECTURE NOTES

Last update: November 2 2014

SEMINAR 1A

Classical sociology and modernization

Classical sociology has very risen from understanding the process of modernization and industrialization. Take Marx as an example: industrialization has created factory work, in which ever more workers produce goods using means of production that the capitalists possess. The logic of modernization then creates a growing and empovering (?) proletariat, as opposed to a small class of owners (capitalists). However, the same conditions have brought the workers together in factories and labor movements can easily be formed. The workers will unite and overthrow the capitalist conditions. (Note that this is a theory about Social Inequality, Conflict and Social Participation). Weber: the rise of Protestantism (a religion without a personal god) has lead to / is part of the rise of rationalization, which is an important aspect of modernation. The protestant ethic leads to capital formation and unprecedented growth of wealth. It also allowed for the rise of scientific (rational) ideologies. Durkheim: Industrialization leads to division of labor and a decline of 'mechanical solidarity'. This may lead to condition of social disintegration (looser ties between members of societies) and anomia, and ultimately to (egoistic and anomic) suicide.

Inkeles on modernization

Inkeles (1969) is remarkably silent on what constitutes modernization as societal macro-level process. Smith & Inkeles (1966) list the following:

- Urbanization
- Higher levels of education
- Industrialization
- Extensive mechanization
- High rates of social mobility

Individual modernity then is values, attitudes and behavior either generated by or required for effective participation in modern society. Inkeles (1969) thinks the following are part of the modernity syndrome:

- Openness to new experiences

- Increasing independence from traditional authority (religion, parents) and shift to public leaders.
- Believe in efficacy of science and abandonment of passivity
- Ambition to reach high educational and occupational attainment
- Careful planning, being in time
- Interest and participation in public affairs and government
- Interest in global rather than local affairs.

The original measure contains 119 items. Since this a bit unpractical, they present a 14 item short hand scale (Smith & Inkeles, 1966: Chart 2).

SEMINAR 1B

Popper's falsifiability

Popper established the falsifiability criterium to distinguish ('demarcate') scientific statements from non-scientific statements. Scientific or testable statements are statements that can be shown to be not true by empirical data (observation). Popper's famous example is: "All swans are white". Observations of (many) white swans ('confirmations') do not show the truth of this statement, but observation of a single black swan shows that it is not true. This argument implies that there is no such thing as logical induction (as assumed by positivism), i.e. arguing from the data to general statements. At best induction is a psycho-logical operation. Science operates via the logical argument of deduction, from the theory to the data.

- "All swans are white" is a general of 'lawlike' statement, but it is not really a statement we would care for in science. However, it is a model of statements that are the core business of science, causal statements, that have the general format X → Y, or X increases Y, X produces Y, X causes Y etc. All causal statements are general or lawlike statements.

Later Popper

While the falsifiability criterium still stands as the demarcation of scientific and non-scientific statements, there are obvious problems with falsificationism as a scientific strategy. You cannot stop at testing your hypothesis – in the end everything will turn out to be not true. What can we do with a bunch of untrue statements? Poppers initial answer was to define several measures of truth approximation, or corroboration, but the real step forward here is to see growth of knowledge as an evolutionary process, in which multiple theories compete for survival. The crucial element here is that theories need not only be tested, but also or rather be compared. In this case it is not so important that all theories are imperfect accounts of the empirical observations, but rather the degree to which they are, relative to one another. Using two or competing theories a very useful way to set up any theoretical framework. This also leads to another important maxime: science does not start with observations (induction) but rather with an account of what has been found before. Science starts in the library!

Non-informative statements

There are many forms on non-informative statements and they can play an important role in our theorizing:

- Questions
- Definitions
- Existential question ("X exist")
- Orienting statements ("Gender plays an important role political participation").

Is a representative sample needed for scientific analysis?

Many of you raised the criticism against Inkeles and Verba that somehow their samples are "not representative". This is too easy a criticism. The core business of science is testing causal explanations, how things work. The golden standard of such a test is the randomized group experiment, in which you sample individuals randomly into a treatment and a control group. By randomization you make sure that control and experimental group differ in no other attribute than the experimental condition. The problem with experiments that they are hard to do in social science: e.g. we cannot sort individuals randomly in conditions of lower and higher education and then test which group is more modern in the end. To get closer to an (quasi-)experimental design we have to use control variables and keep these constant. Unlike with experiments we have to theorize and measure these control variables. And so the question of adequate design is not whether the sample is unrepresentative, but rather whether the right controls (confounding variables) are measured and controlled.

Random sampling may to some extent be useful to counter selectivity in composing the groups of comparison. E.g. Inkeles' comparison of high and low education seems to be a comparison of students and workers, which seems not to be the best way to get to low and high educated.

A way to include control variables in an observational design without measuring them is using panel data. By controlling the initial measurement (i.e. by studying change), you keep all variables constant that cause variation in the first measurement.

What are representative samples?

Despite the frequent use of the term, representativeness is in fact not a well-defined statistical term. Rather, sampling should be characterized as probability sampling, which means that we know the probability of each case in the population to be chosen is known. Simple random sampling (SRS) is something of a golden standard here, meaning that each unit has the same probability of being selected at the first step. This condition is assumed in standard statistical programs. Complex random sampling occurs when the probabilities are known (and may be even equal in the first step), but the units are chosen in steps, e.g. locations → streets → household → persons. Such samples need special statistical treatment (in SPSS: Complex Survey Sampling), because they are usually inefficient, i.e. the sampling variability is larger than calculated under SRS. With proper statistical treatment, any probability sample (simple of complex) will be unbiased: as the sample is larger we get closer to the population value. Non-probability samples (e.g. access panels, quota samples) are not necessarily biased (although the can be), but the big problem with it is that we cannot really estimate sampling variability.

When are representative samples needed?

Summary statistics of univariate distributions (such as a mean or a percentage) are most sensitive to sampling bias. Bias is much less likely for numbers that represent bivariate of multivariate distributions, such as correlations or regressions (differences between means). This is so because the biasing process would have to influence both distributions at the same time.