DEPARTMENT OF POLITICAL SCIENCE

PECOS4021 - Research Methods
28 September 2021 at 09:00 – 12:00 (3 hours)

3 hours in duration

Short-answer questions:

Please answer <mark>all 5 questions</mark> below. Suggested use of time, in total, for all short-answer questions is 1,5 hours (weight: 50% of exam).

S1)
Explain what <u>reliability</u> implies, and how it can be assessed, in relation to (a) quantitative content analysis and (b) discourse analysis. *(5p)*

Reliability refers to the quality and integrity of data collection; in essence, the extent to which repeated measurement produces the same result. Halperin and Heath (p. 149) refer to consistency and repeatability: "A research design is reliable if other researchers can perform exactly the same results. A measure is reliable if it gives us the same result over and over again (assuming that what we are measuring isn't changing." With quantitative content analysis reliability can be assessed through intra-coder and inter-coder consistency in coding, if manually conducted. If similar words and phrases are given the same coding, it indicates high reliability. Discourse analysis principally allows for a form of intra- or inter-coder reliability as well, but is much more circumspect as coding is not straightforward and inferences are drawn on the basis of interpretations in light of context rather than observations as such.

S2)
Interviews in peace and conflict research pose a series of ethical issues, especially if they are conducted in conflict-affected settings. Establish two ethical challenges a researcher may encounter when doing interviews in conflict settings, and discuss what strategies (s)he could adopt to mitigate these. *(5p)*

The core ethical obligation of any researcher, but in particular in the area of peace and conflict research, is to do no harm. That means that we have to take great care to ensure that we do not endanger the safety and well-being of the interviewees (as well as any research assistants or translators) through their participation in our interviews.
The challenges students could discuss here are 1) re-traumatization, i.e. re-activating the trauma of prior experience of violence or armed conflict triggered by the interview, 2) exposing interviewees to security risks, e.g. by conducting interviews in locations that are not safe enough or by drawing the attention of state or armed actors to the interviewee who is seen meeting a (foreign researcher), 3) exposing interviewees to social risks. The example discussed in the lecture was of an interviewee who is a victim of conflict violence (in particular stigmatized forms of violence, like sexual violence) but has not publicly disclosed their victim status. If this individual is seen speaking to a researcher who is known to carry out research on conflict violence, there is a risk of "outing" that individual as a victim, which

can have social repercussions, such as being stigmatized, marginalized, discriminated against etc.

This is a non-exhaustive list, however, and the students could also discuss ethical considerations as they pertain to data security, ensuring the anonymity of research participants, or how more broadly different ways in which their own presence in the field may affect conflict dynamics on the ground. They should link these specifically to interviews, however.

Strategies students may, for example but non-exhaustively, discuss to overcome the above risks are 1) designing an interview guide that asks questions about war experiences broadly and in general terms but does not include specific questions on personal experiences of violence, unless absolutely necessary; or being very aware in the interview session, being empathetic, offering to take breaks or even ending the interview prematurely if necessary; and for 2) and 3) being careful and deliberate in selecting interview locations, taking into consideration also what is convenient and comfortable for the interviewee, but making privacy the decisive factor; or taking care to not disclose their researcher identity or the nature of their research project publicly.

On issues of data security and ensuring anonymity of research participants, student answers may include completing the NSD process and following NSD guidelines and instructions; storing ID keys for interviews separately from the interview transcripts; using pseudonyms and removing any other identifying information that could allow knowledgeable individuals to identify interviewees (such as organization they work for, the position they hold, where they live or grew up, detail on major life events or victimization events); and generally complying with UiO's regulations for yellow, red and black data when storing interview data and ID keys.

S3)
According to Mahoney & Goertz (2006), quantitative and qualitative research often pursue causal analysis differently. Consider the assumed relationship between <u>state suppression</u> and <u>civil unrest</u>. How would research from each of 'the two cultures' proceed to investigate this? *(5p)*

Key elements from Mahoney & Goertz are summarised in their table on p. 229:

	Criterion	Qualitative	Quantitative
1	Approaches to explanation	Explain individual cases; causes of effects	Estimate average effect of independent variables; effects of causes
2	Conceptions of causation	Necessary and sufficient	Correlational; probability
3	Multivariate explanations	INUS causation; occasional individual effects	Addicitve causation; occasional interaction terms
4	Equifinality	Core concept	Absent concept

It is essential to note the difference between explaining outcome (qual) and measuring average effect of X on Y (quant); and the emphasis on causal narratives and INUS causation (qual) vs. causal effects for each variable (quant). The example of state suppression => civil unrest could be addressed through qualitative analysis by producing a precise causal narrative (combining necessary elements of a sufficient explanation) to account for a specific outcome of civil unrest. The emphasis would be on the mechanisms that generate the outcome; essentially, what makes state suppression cause civil unrest. Quantitative research would be disposed to calculate the average effect of state suppression on civil unrest (both quantitatively measured) alongside other variables.

S4)

Kreuzer (2019) refers to 'ontological calibration of evidence' as an important task for the historical researcher. Account in brief for the concept and discuss how it relates to generalisation in small-N research. *(5p)*

Ontological calibration concerns the status of observations made and inferences drawn in the analysis of a historical event or period. What features may be seen as case-specific and contextual, as opposed to features that are transferable to other times and places? The term thus relates to the historical or geographic boundary conditions and how clearly those conditions are spelled out. In Kreuzer's article, such calibration is a requisite task of descriptive inference in historical research. "Historians may detach pieces of evidence from their chronological or geographic context, re-order them, and make them more uniform and hence comparable," argues Kreuzer, but such detachment must be done with utmost care (and historians are typically very aware of this – more so, perhaps, than social scientists dealing with historical analysis and disposed to generalise). Kreuzer exemplifies with Skocpol's comparative analysis of revolutions, wherein she seeks to describe, classify and compare (and explain the outcome of!) cases from highly different contexts.

This discussion is in a sense the historical discipline's mirror image of discussions in social science methodology about the uniqueness of individual cases vs. the aspects that may be abstracted and generalised. This particular challenge of descriptive inference as discussed by Kreuzer is endemic to all small-N research, with the additional aspect of jumps in time when it comes to historical generalisation.

S5)

Discuss what <u>saturation</u> concerning (a) sampling and (b) analysis of data could imply in a research project based on semi-structured interviews. What ambiguity may occur when assessing that saturation has been obtained? *(5p)*

Saturation is often used as criterion for when either data collection or analysis is sufficiently extended; to conclude that enough is enough. According to Saunders et al. it is implicitly also an inference, a form of generalisation: «a statement about the unobserved based on the observed». As suggested by the question, an essential distinction is drawn between saturation referring to (a) sampling as opposed to (b) the analysis of data that you have already collected. Faced with interviews, (a) would refer to the decision to include (or not) more interviews, while (b) translated to sufficient analysis of interviews already conducted. Deciding when saturation has been obtained is never a rock-solid assessment. Concerning sampling, one criterion would be a form of representative sample. More often, however, the researcher would be interested in whether more interviews would provide alternative perspectives, more ideas or information. Is, for example, triangulation an important concern? Or might there be voices and ideas left in the dark? In analysing existing data (typically in the form of transcribed interviews), there will always be the question of whether more analysis would help develop or saturate further theoretical categories and illustrations thereof.

Long-answer question:

Suggested use of time for long-answer question is 1,5 hours (weight: 50% of exam).

L1)
The fall of Kabul on 15 August 2021 signified the Taliban's take-over of state institutions and accelerated the withdrawal of Western forces from the country.

You are asked to contribute to a research group analysing recent events in Afghanistan in light of academic literature on US foreign policy, insurgency and regime change.

(*NOTE: None of the questions below requires in-depth knowledge on the conflict in Afghanistan.*)

.
One strand of research seeks to analyse US withdrawal from Afghanistan in light of previous military interventions. The exit from Vietnam in 1975 is singled out for comparison, with the fall of Saigon that year seen as broadly equivalent to the fall of Kabul. The hypothesis is that both were caused by a weakening of American resolve and a radicalisation of the enemy.

In response to this proposal, a reviewer states that comparing the two raises serious validity concerns.

  a) Account for the concepts of <u>measurement validity</u> and <u>measurement equivalence</u>, then discuss why equivalence may be difficult to obtain when comparing two separate cases. Feel free to illustrate with the variables mentioned above, without going into detail. *(7p)*

According to Adcock & Collier (p. 530), "valid measurement is achieved when scores (including the results of qualitative classification) meaningfully capture the ideas contained in the corresponding concept". Measurement validity assesses the correspondence between concept and indicator. Equivalence addresses whether measurement across different contexts is capable of covering the same concept empirically. Such equivalence may be difficult to obtain, as exemplified by Afghanistan vs. Vietnam. Is American resolve empirically measurable at all (weakening? strengthening?), and does it mean the same thing in the two cases? What would resolve imply in the "Cold War, fighting a proxy enemy" context of the mid-1970s as opposed to the occupation of and withdrawal from Afghanistan? Is radicalisation sufficiently generic to "travel" across cases, and can it be measured and assessed in a similar way in Vietnam and Afghanistan? It might well be that comparison does yield essential insight about both cases. But there is reason for caution as to what can be learnt from this comparison, as low contextual equivalence may blur any clear line of argument cutting across the two cases.

Students may well exemplify strategies to strengthen measurement equivalence through Context-specific domains of observation, Context-specific indicators and adjusted common indicators. These are not explicitly asked for, but follow naturally from an elaboration on how to ensure measurement equivalence.

Engaging a broader literature on military withdrawal and regime change, the reviewer also argues that the proposal should avoid "selecting on the dependent variable".

b) Why would selecting on the dependent variable be a problem when seeking to contribute to theoretical knowledge about causes of regime change? What could the researcher do to address this concern? *(5p)*

Selecting on the dependent variable is problematic for the simple reason that we need variation in order to explore causal relationships. Following Lijphart as well as KKV, Lijphart (1971:683), all empirical resarch draws upon a shared logic of inference: the verification of a (causal) relationship between single variables while isolating the effect of others. "The **experimental method***, in its simplest form, uses two equivalent groups, one of which (the experimental group) is exposed to a stimulus while the other (the control group) is not. The two groups are then compared, and any difference can be attributed to the stimulus. Where experiments are not feasible, equivalence (ceteris paribus condition) is pursued through randomisation, comparing scores on the dependent variable among cases with various scores on the independent variables in question. This is not possible in small-N research, which opens for a discussion on how to select cases strategically. Conventional methodology would have that maximising variation is an inherent good in order to explore causality. Other comparative designs, especially most similar systems design, may be of help in isolating an assumed causal relationship to analyse further in depth. This could certainly be pursued in the example above, by including comparative cases where regime change did not occur, despite various circumstances pointing in that direction. It is also worth mentioning that "within-case evidence" and the wide extent of observations enabled by in-depth case study means that a case is always more than one data point, allowing for causality to be explored by comparing across time and potentially space within the bounds of the case.*

Distancing herself from general theories of regime change, another reviewer claims that INUS causation would be a more fruitful approach to analysing the path towards regime change in Afghanistan.

c) What are the principles of INUS causation, and what does it seek to accomplish? *(6p)*

Small-N research typically stresses "how and why" something occurs, rather than "how often" or "with what probability". Thus, causal analysis typically addresses why an event or outcome occurred in a specific case. Process tracing is a way of specifying how to go about such research. Focus is directed towards causal process observations - that is, evidence which helps explain and substantiate how X led to Y.

*The causal arguments arising from such research often distinguish between necessary and sufficient causes. -**A necessary cause X must be present for Y to occur, yet cannot account for Y on its own**. Thus, economic growth could be be necessary for a rise in welfare among the working class, but it does not provide a sufficient explanation, as something else must be set in train too.*

A sufficient cause X means that Y will occur as a result of it being present - but Y could have been caused by other causes too. *An example would be that a financial breakdown (X) would be a sufficient cause for the deficit in the state budget (Y), but the breakdown was not necessary for the deficit to occur (it could have occurred for other reasons too).*

*Often in process tracing (and in any small-N research with the goal of developing or testing a causal explanation), what we would be looking for is so-called **INUS causation**, which combines necessary and sufficient conditions by identifying **necessary components of an explanation that is sufficient to account for the outcome**. The concept acknowledges that different sets of circumstances could lead to a specific outcome (outbreak of war; famine; civil unrest...). The quest for the researcher is to identify the causal pathway; a collection of causes that together constitute a sufficient explanation of the outcome. In practice, this means finding an explanation that is detailed enough to demonstrate why this outcome had to occur but which avoids including elements that are not necessary to the explanation.*

Another strand explores what she calls "the ambiguity of popular support to the outgoing regime", drawing on fieldwork and extensive interview data from recent years. The researcher presenting stresses the importance of <u>meta-data</u>, as conceptualized by Lee Ann Fujii.

d) Three of Fujii's examples of meta-data are *inventions or embellishments; denials; and silences*. Critically reflect on their importance as data. What potentials and challenges do you see relating to the role of meta-data in the research process? *(7p)*

***Inventions and embellishment**s: research participants may misrepresent or entirely invent (war-time) experiences; they may also embellish accounts when they do not recall details or specific events. **Denials**: about specific events, such as acts of violence perpetrated at one point or event, while the research participant draws attention to other times/ events when they experienced violence. **Silences and omissions**: research participants do not broach certain topics because these are stigmatized or attached with shame (e.g. sexual violence, pillaging in Rwanda)*

*Fujii discusses different ways in which these can be important data. **Inventions and embellishments** may illuminate dominant discourses and narratives, i.e. what it is acceptable to say, what it is expected to say or to have experienced (e.g. a story of hiding and escape during a genocide); as well as aspirations or desires. Thus, they may help illustrate, sharpen, contour the overall context of the research setting. **Denials** may direct the focus to other phenomena you might not have on your radar, e.g. Rwandan women denying violence happened in their area during the genocide and instead talking about their post-genocide victimization, which the researcher might otherwise have missed. **Silences and omissions**, similar to inventions and embellishments, may show what it is (un)acceptable to talk about, i.e. what dominant discourses are, what it is important not to deviate from and what not to disclose about yourself.*

Students can critically engage with these points Fujii raised or come up with their own, as long as they are plausible. A good answer goes beyond simply repeating Fujii's points and will discuss the difficulties of identifying such meta-data in the first place and what effect this may have on the analysis and the conclusions we draw. For example, how do we know whether a story a research participant tells is embellished or even an invention or whether there are silences? What if you never notice that a story an interviewee told you is likely an invention? And can you really know for sure anyway? How do we know that something happened during a war if research participants never mentioned it (silence or omission)? Strategies Fujii discusses that can help identify meta-data (and ensure our confidence in our interpretation of these) include looking for consistencies across multiple interviews with the same person, plausibility in light of interviews with other individuals, triangulation with other

data sources etc. – but what happens if you do not have sufficient data to do this? Other points that could be raised relate to the interpretation of meta-data and how they affect one's findings. Would two people doing similar interviews in the same context identify the same kind of meta-data and draw the same conclusions from these? This likely depends on experience, interpersonal skills and also the number of interviews you can do. What if the researcher never notices inventions, denials and silences and treats everything people tell them at face value? Again, this is a fairly open-ended question with many potential potentials and challenges for students to identify.