

King, Keohane, and Verba (KKV) aim to connect the traditions of quantitative and qualitative research through the use of logical inference. Although quantitative research, characterized by numerical measurements, statistical methods, and easily replicable methods, seems quite different from qualitative research, which uses small numbers of cases and investigates them through historical analysis and interviews, those differences are "methodologically and substantively unimportant" (p.3).

Both approaches are based on scientific research, which invokes the following goals and procedures: 1) the goal is inference, either causal inference where one learns about causal effects from the data observed, or descriptive inference, where one uses observations about the world to learn about other unobserved facts; 2) procedures are public, so that data collection techniques and data can be replicated and investigated by any interested party; 3) conclusions are uncertain, as inference is an imperfect process, and thus need uncertainty estimates; and 4) content is the method, as scientific research adheres "to a set of rules of inference on which its validity depends" (p. 9) and thus the content is the rules, not the subject itself, as anything can be researched scientifically.

The four components of scientific research design are the research question, the theory, data, and the usage of the data. KKV offer a series of suggestions on improving the quality of these pieces to ensure a sound, unbiased, and efficient research project. They recommend that research projects should pose questions which are important in the real world and should make a specific contribution by increasing our ability to create explanations of some aspect of the world.

The theory component should be a "reasoned and precise speculation about the answer to a research question. imply[ing] several more specific or descriptive hypothesis" (p.19). Researchers can select theories which could be wrong, but should always choose a theory capable of generating observable implications and a theory which is as concrete as possible.

To improve data quality, researchers should "record and report the process by which data are generated" to allow for the investigation of bias, confounding, etc. (p. 23). Further, to better evaluate their theories, researchers should collect data on as many of the theory's observable implications as possible. Finally, researchers should ensure that data collection methods are reliable and that they are replicable. KKV suggest that researchers look to "maximize leverage" with their theories, that is, explain as much as possible with as little as possible. If a few causal variables can explain a complicated phenomenon, the leverage is quite high; on the other hand, if it requires a number of variables to explain a simple event, leverage is low. In light of the concept of leverage, the authors recommend that researchers "routinely list all possible observable implications of their hypothesis that might be observed," (p. 30), as the more evidence, the more powerful the explanation becomes. The most serious problem with research in social sciences, they argue, is "the pervasive failure to provide reasonable estimates of the uncertainty of the investigator's inferences" (p. 32).

The authors lay out the differences between interpretation and inference. Interpretation seeks "accurate summaries of historical detail," placing the events "in an intelligible context within which the meaning of the actions becomes explicable" (p. 36). Using *Verstehen* (German for

empathy), researchers who employ interpretation seek to understand intentional aspects of human behavior. KKV view a useful approach coming out of this method, that of deep cultural immersion (e.g. "soaking and poking") prior to formulating research questions, but argue that interpretation is not transportable beyond the researcher herself; inference is a scientific, not emotional method which is more efficacious.

KKV then seek to understand the uses and techniques of inference in its descriptive and causal roles. Inference helps researchers to distinguish between events which are systematic [fundamental and predictable characteristics] and those which are nonsystematic [random events like changes in the weather or terrorist attacks]. If researchers fail to determine which events are systematic and which are not, "the lessons of history will be lost, and we will learn nothing about what aspects of our subject are likely to persist or to be relevant to future events or studies" (p. 63). KKV bring up three criteria for judging methods of creating inferences: unbiasedness, efficiency, and consistency.

Unbiased estimates that those that are replicated as the average when the same methods of inference are applied again and again to an event. If a measure shifts in the estimate in one way or the other, it is called bias. One major source of bias is that the informants who answer poll questions over under- or over-state their estimates to questions (p. 64). Efficiency, on the other hand, is a relative concept measured by calculating the variance of the estimator across hypothetical replications; the more observations there are, the better (i.e. smaller) the variability.

These self-checking methods are necessary because of the fundamental problem of causal inference: "no matter how perfect the research design, no matter how much data we collect, we will never know a causal inference for certain" (p. 79). [See p. 81- 82 for a mathematical definition.] KKV suggest two methods for avoiding this problem, and the problem of separating systemic from nonsystematic components: unit homogeneity and conditional independence (p. 91).

Homogeneity is achieved when "the expected values of the dependent variables from each unit are the same when our explanatory [i.e. independent] variable takes on a particular value" (p. 91). Homogeneity is the assumption that all units with the same explanatory variable value have the same expected value of the dependent variable. "We believe that the differences we assume in the values of the dep. variables are the results of the differences in the values of the explanatory variables" (p. 93). Conditional independence means that values are assigned to variables independently of the values taken by the dependent variables, thus avoiding the problem of endogeneity.