

- MIT Press, 2005).
- Gerring, John, *Social Science Methodology: A Criterial Framework* (Cambridge University Press, 2001).
- Goertz, Gary, *Social Science Concepts: A User's Guide* (forthcoming).
- Goertz, Gary, and Harvey Starr, eds., *Necessary Conditions: Theory, Methodology, and Applications* (Rowman and Littlefield, 2002).
- Mahoney, James, "Nominal, Ordinal, and Narrative Appraisal in Macro Causal Analysis," *American Journal of Sociology*, Vol. 104, No.3 (January 1999).
- Mahoney, James, and Gary Goertz, "The Possibility Principle: Choosing Negative Cases in Comparative Research," *American Political Science Review*, Vol. 98, no. 4 (November, 2004).
- Munck, Gerardo, and Jay Verkuilen, "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices," *Comparative Political Studies* (February 2002).
- Ragin, Charles, *Fuzzy Set Social Science* (University of Chicago, 2000).
- Van Evera, Stephen, *Guide to Methodology for Students of Political Science* (Ithaca: Cornell University Press, 1997).

How Qualitative Research Really Counts

Kenneth R. Benoit

Trinity College, University of Dublin

kbenoit@tcd.ie

Polonius: What do you read, my lord?
 Hamlet (reading a book): Words, words, words.
 Polonius: What is the matter . . . that you read, my lord?
 Hamlet: Slanders, sir: for the satirical slave says here that old men have grey beards; that their faces are wrinkled; their eyes purging thick amber and plum-tree gum; and that they have a plentiful lack of wit, together with most weak hams: all which, sir, though I most powerfully and potently believe, yet I hold it not honesty to have it thus set down; for you yourself, sir, should be old as I am, if, like a crab, you could go backward.
 Polonius: Though this be madness, yet there is a method in't.
 (*Hamlet*, Act 2 Scene 2)

The main point of this essay is straightforward: The distinction between quantitative and qualitative research, when applied to empirical political analysis, is exaggerated and largely artificial. In fact, most political scientists can happily perform valid and useful research without being concerned about where they stand on the quantitative-qualitative divide. Furthermore, qualitative characterizations are often easily converted into quantitative characterizations, and many qualitative characterizations are implicitly quantitative to begin with. Finally, qualitative characterizations of the empirical world are almost always more useful when converted into quantitative ones.

In the spirit of a piece written for a newsletter on qualitative methods, I will at the outset fully acknowledge that my

essay is a discourse and should be treated as such. As a bit of background on the context of this text and its author, I describe myself as a comparative politics scholar, primarily quantitative but also familiar with field work, interviewing, and survey analysis. I teach Advanced Quantitative Methods to Ph.D. students but also a course in Research Design. My current research involves estimating political party positions on policy issues in numerous countries and in the European parliament, using surveys of expert judgments and computerized content analysis.¹

I will draw the distinction between quantitative and qualitative research in a deliberately simple manner, and then explore the implications of this distinction. The difference has to do with the use of numbers. *Quantitative* research characterizes observed phenomena using numbers, while *qualitative* research does not. A qualitative statement about voter attitudes toward political participation is that voters are disillusioned and apathetic, feeling that voting is a waste of time in the face of widely perceived corruption, ineffectiveness, and lack of meaningful policy content in party platforms. A quantitative statement would be that two-thirds of voters do not plan to vote, or that 45% report not feeling close to any particular party.

This simple distinction is normally confused by the unnecessary bundling of quantitative or qualitative research with other related yet logically separate issues. Such issues include the balance of cases to variables, whether research should be critical, normative, or positivist, and whether we can use case studies to prove causal propositions.

Let us deal with the first of these conflated issues, that the qualitative-quantitative distinction has to do with the number of cases, or more accurately, with the ratio of cases (call this n) to variables (call this k). Conventional accounts² of causal inference require that $n > k$, while qualitative researchers maintain that valuable knowledge, possibly even causal relationships, can be determined when $n < or = k$. Several points can be made on this issue. First, it is interesting that the identification of what is *qualitative* research in this framework rests on fundamentally *quantitative* grounds, namely the relationship of the quantities k and n . Second, to conceive of the qualitative-quantitative difference in terms of cases vs. variables makes it impossible to maintain that the two types of research are different in kind. Rather, it suggests that the difference is measured in degrees, even on a ratio scale, more precisely by the ratio of n to k . Finally, a focus on sample size shifts debate to other issues such as causal inference and case selection, obscuring the central issue of whether the empirical world consists of qualities or of things that can be counted.

Yet it is this issue of counting that is central in distinguishing quantitative from qualitative research. The essence of the matter boils down to measurement and the type of information we can feasibly use in characterizing the observed world. In the language of measurement, in fact, the distinction is more sophisticated than a simple dichotomous difference as implied by quality vs. quantity. Observations can also be measured according to different levels of *scale*, typically de-

scribed as:

Nominal: observations are distinguished from one another in a purely qualitative fashion, such as parties, states, or ethnic groups.

Ordinal: observations contain an inherent ranking, such as Fail, Poor, Good, Very Good, and Excellent. Ordinal measures are qualitative but can easily be converted into quantitative measures, such as assigning 0-4 to the previous example (perhaps to compute a grade-point average). This sort of conversion is also carried out by Likert scales, for instance.

Interval: implies that observations can be measured on a scale where increments have a constant distance, such as when we measure temperature on the Fahrenheit scale. All interval scales are quantitative.

Ratio: this is a purely quantitative scale that takes interval measurement further by having a meaningful zero point, permitting ratios to be taken. For instance, we might measure the number of cases in one's research design on a ratio scale, with 100 cases being 100 times greater than a single-case study, and zero cases representing only a theory with no data.

The move from qualitative to quantitative measurement occurs as more information is incorporated. It is also a natural consequence of any effort to *compare* observations. Comparison implies ordering, whether on a qualitative or quantitative dimension. Ordering implies by nature that one quality is stronger or greater in one observation than in another. And relations such as "stronger" or "greater" imply, whether this is made explicit or not, a relative degree of *quantity*, even if the characteristic being compared is discussed in purely qualitative terms. The act of comparison, therefore, naturally and readily lends itself to quantification.

I will take this reasoning a step further, to make a strong claim for the innate superiority of characterizations of the empirical world based on quantitative research. Our understanding of the empirical world rests on a system of statements supported through evidence. One of the primary objectives of empirical research is to establish this evidence. I contend that when it comes to establishing and defending such statements, quantitative evidence is superior to purely qualitative evidence. Evidence based on numbers is easier to compare, easier to verify, and easier to refute than that based on qualitative evidence. Even purely qualitative evidence, such as expert opinion, is elevated in reliability when it is expressed the implicitly quantitative framework of a consensus or experts.

Now at this point you may strongly disagree with my views, or you may disagree, or you may neither agree nor disagree, or you may even agree or strongly agree. I suggest that if you do not agree with any rating other than strongly disagree, then you have a logically inconsistent position if you call yourself a qualitative researcher. (A true qualitative position would permit only either categorical agreement or

disagreement with the proposition that quantitative measurement is innately superior to purely qualitative representations.)

In the discussion to this point I have assumed that our enterprise was to characterize the empirical world. This returns to the second of the "conflated problems" I discussed above, which is the mode of inquiry. By restricting ourselves to characterizing the empirical world, we remove from the quantitative-qualitative research discussion not only formal theory, political philosophy, normative political argument, but also interpretative approaches such as discourse analysis, constructivism, social constructivism, post-positivist neo-feminist critical constructivism, and so on. These latter approaches share not only an inevitably strong qualitative element, but also a different basic objective from empirical ("positivist") research. At the extreme of these are interpretivist approaches which deconstruct reality as if it were a text, where the reader interacts with the text and its social context and attempts not just to uncover but also to construct meaning. Critical literary analysis, whether deconstructive or not, is typically interpretative, where the goal is to uncover meaning for the purpose of understanding a text, its story, the social world it represents, etc.

To draw on a more quotidian form of literary analysis consider film reviews. Film critics compare and evaluate, but with the goal of aiding the reader to understand and appreciate a film, in addition to knowing whether it is worth seeing. A reader of film reviews will typically know something about the critic's tastes based on a contextual knowledge of the critic's previous reviews, and will therefore be able to interpret the review accordingly. In this way, for instance, a reader of the *Times of London* might read between the lines of a one-star thumbs down from a culturally elitist British reviewer and, despite an unfavorable review, disregard the reviewer's suggestions and nonetheless go to see a perfectly good film like *X-Men*.

Text analysis is in fact an excellent field on which to pitch this battle, since text analysis involves fundamentally qualitative matter that may be analyzed either qualitatively or quantitatively. Let us assert, for instance, that George W. Bush is more of a conservative internationalist than a liberal internationalist like Woodrow Wilson. This statement about the empirical world may be considered an accurate characterization by many scholars of foreign policy. But ultimately such claims must rest on evidence. We might analyze a number of George Bush's speeches to provide this evidence. (Note the use of the term *number*.) Even if we only analyzed one speech, we could seek evidence in quantitative measures of certain words whose use would imply a particular foreign policy orientation. A key feature of conservative internationalism, according to Professor Henry Nau, is an emphasis on freedom over democracy. We might note then that Bush mentioned freedom 27 times in his inaugural address and 21 times in his State of the Union Address and not once stability.³ This form of evidence is easier to compare—say to speeches by other U.S. presidents or other world leaders—easier to verify, and easier to refute, perhaps on the grounds that use of these words in these speeches is not an appropriate indicator of foreign policy orientation.

Table 1

<i>Goal</i>	<i>Quantitative</i>	<i>Qualitative</i>	<i>Examples</i>
Interpretation	No special value	Critical	Appreciate <i>Hamlet</i> Construct discourse on winking
Understanding	Useful	Critical	Understand <i>Hamlet</i> Understand meaning of winking
Description	Critical	Useful	Compare <i>Hamlet</i> to other Shakespeare plays Record observed winking ritual Describe a wine Award an Olympic gold-medal winner Characterize a country's democraticness Express or record a political preference Classify research as quantitative or qualitative
Explanation	Critical	No special value	Judge authenticity of <i>Hamlet</i> authorship Identify states likely to fail Determine whether campaign spending affects electoral success Determine the factors that influence ticket splitting

Consider another text. The preface to this essay quotes *Hamlet*.⁴ The question of what to make of this text is not unlike the question facing researchers confronted with the political world. Are we concerned with whether Hamlet is really mad or merely faking it? What should we conclude about the book Hamlet is reading, based on Hamlet's description? Are we to understand that in his oblique invention of a text about nasty old men in his reply to Polonius, and in mistaking Polonius for a lowly fishmonger just before this exchange, Hamlet is violating strongly held cultural norms regarding politeness and the display of respect for one's elders?

Those are goals of interpretation. We might also attempt to establish falsifiable, empirical statements about Shakespeare's texts. For instance, we could examine *Hamlet* as a whole to determine stylistic evolutions between this and later plays written by Shakespeare. Or, we might attempt to determine the authenticity of authorship based on cryptographic clues possibly left by Frances Bacon or some other ghost-writing impersonator. As it turns out, such debates actually occupy a great deal of space in the literature on Shakespearean literature.

An example: In 1985 a new poem was discovered in the Bodleian Library of Oxford by Gary Taylor and attributed to Shakespeare. On what evidentiary basis would we consider it authentic? Authoritative declarations from Oxford Shakespeare experts? Considered more definitive was quantitative evidence established by two statisticians at Stanford University, Bradley Efron and Ronald Thisted, who statistically analyzed the Bard's entire 900,000-word vocabulary in order to establish usage patterns. Efron and Thisted tested this distribution against the writings of Shakespeare's contemporaries, using

the rules of statistical inference, and found that the test distinguished clearly the writing of Shakespeare from Donne, Marlowe, and Jonson. They then used the observed distribution to predict similar patterns in the 430-word mystery poem and concluded that it perfectly fit the profile for Shakespeare's work.⁵

In having brought literary analysis into this discussion at all, however, we find ourselves at the brink of a dangerous pit regarding the notion of science and whether the author or investigator does or should matter in research, and whether political science research should be closer to the stereotype of hard science or whether it can or should share elements of film reviews.

I am going to go circumvent this pit, however, by arguing that even film reviews and interpretive accounts can be enhanced by using quantitative information. For film reviews, a "thumbs up" or "thumbs down" may be qualitative, but a rating from one to five stars⁶ converts the qualitative, highly subjective, interpretative measure into quantitative information. Wine and cigar reviews do the same, such as not only describing the Carlos Toraño Signature Collection as having an "earthy core with hints of leather and sweet spice," but also as having received the ("astounding and nearly unprecedented") rating of 4.7 from *Smoke Magazine*.⁷

In the summer of 2004 we probably all watched some of the summer Olympic events. Judging Olympic events involves making highly subjective judgments of observable behavior, where the past experience, artistic and athletic context, and personal orientation of the judges all play an important and acknowledged role. Many competition events have formal guidelines (such as the mandatory components of gymnastic

routines), but winners are determined ultimately according to intangible qualities. In other words, the excellence of a performance is an intrinsically qualitative characteristic. But once we need to *compare* performances, such as determining three top-ranked winners, then this qualitative performance *must* be measured quantitatively. Of all the events in the Olympics, there is not a single event—whether based on solo artistic or technical performance, or on scoring goals—whose outcome is not determined on the basis of a quantitative score.

So, to summarize, in order to enhance their usefulness qualitative measurements are generally converted into quantitative information. We need reach no further than our bookshelf or latest journal copy to find abundant examples from our own discipline: location of political party positions on a left-right scale; public attitudes toward post-materialist values; levels of democratic governance or corruption; the levels of conflict in international environment; the relationship between electoral systems and the expected number of political parties. Goals that we value in political science, such as the ability to make meaningful comparisons, the manageability of data, the ability to replicate analyses, the capacity to characterize confidence or uncertainty, and the potential of our propositions to be falsified are all enhanced when research rests on quantitative evidence. Purely qualitative approaches are most useful if we wish to interpret or understand an observed phenomenon (event, idea, text, etc.) but if we wish to compare interpretations, we are likely to need numbers. (See Table 1 above for an attempt I have made at classification in order to maximize comparison and potential refutability of my claims.) Purely qualitative description is possible, but comparison implies relative quantities, meaning that this form of qualitative research really *counts*. Explanations of the sort in which we are likely to have confidence will involve some form of quantitative statements 100% of the time.

Notes

¹ A description of and data from the expert survey project is available from <http://www.politics.tcd.ie/ppmd/>. A full description, research papers, and software for the computerized content analysis project are available from <http://www.politics.tcd.ie/wordscores/>.

² See Gary King, Robert Keohane, and Sidney Verba (1994) *Designing Social Inquiry* (Princeton University Press), known in many circles simply as “KKV.”

³ Henry R. Nau, “Bush’s classic conservatism.” *International Herald Tribune*, March 29, 2005, p7.

⁴ William Shakespeare, 1603. *Hamlet* is a play whose namesake is a Wolverine-type character (*X-Men*) in the sense that the has parental issues, those close to him worry about his stability, and he inadvertently stabs a good guy. (And on what basis can you judge this literary interpretation worse than any other?)

⁵ Bunce, Nigel, and Jim Hunt, “The Statistics of Shakespeare,” *The Science Corner*, May 13, 1968, University of Guelph, <http://www.physics.uoguelph.ca/summer/scor/articles/scor109.htm>.

⁶ Or tomatoes; see for instance <http://www.rotten-tomatoes.com>.

⁷ See https://www.cigarsforless.com/Cigars_in_US/Carlos_Torano_Signature.htm for the full description and ordering details.

The American Political Methodology Debate: Where is the Battlefield?

Bernhard Kittel

University of Amsterdam

b.e.a.kittel@uva.nl

Like many other realms of political science, the debate over political methodology in Europe has been influenced by the American discussion up to the point of being reduced to a perhaps somewhat belated commentary of debates deemed terminated in the American context. An analysis of references to methodological contributions in political science articles would probably reveal that American authors are far more frequently cited by Europeans than vice versa. It is even hard to discern something like a European methodology debate worth of its name.

Recent observers of the differences between American and European practice have in particular highlighted the stronger emphasis on and disciplinary status of comparative politics in Europe as compared to the United States (Lijphart 1997) and the more systematic empirical-analytical approach using quantitative data for rigorously testing hypotheses in the United States as compared to more institutionalist, descriptive, constructivist, and more generally qualitative approaches in Europe (Marsh and Savigny 2004; Moses, Rihoux, and Kittel 2005; Norris 1997). These stereotypes, which can already be found in David Lodge’s characters of Philip Swallow, the worrisome British academic, and Professor Zapp, the jovial American versed in the ways of the world, in his novel *Changing Places*, contain, like all of such generalizations, some elements of truth. But reality is always more complex and we can find practitioners of all denominations in both academic communities.

Looked at from a more long-term European perspective, the current American debate reiterates episodes which we have encountered in the European history of science. The battle between the nomothetic and the idiographic worldview has accompanied the social sciences since their first attempts to define their topics and approaches, and the relationship has never been one of great friendship. Among these, the debates between social philosophy and positivist social science in the early 19th century, the economic *Methodenstreit* of the late 19th century between the Austrian marginalist school and the German historical school, and the *Positivismusstreit* waging between the critical theorists of the Frankfurt school and the “positivists” during the 1960s are only the most notable. In this perspective, the current American debate may simply appear as the newest clash of academic civilizations. In comparison to these older debates, however, the current controversy is indeed astonishing in the extent to which the contending proponents seem to converge on fundamental issues. In this sense, perhaps, we could speak of a very American solution to the long-standing conflict.

Qualitative Methods

Newsletter of the
American Political Science Association
Organized Section on Qualitative Methods

Contents

Symposium I: The Quantitative/Qualitative Distinction

<i>Ten Fallacies About Qualitative Research</i> Gerardo L. Munck	2
<i>The Fallacy of Fallacies</i> Andrew Bennett	5
<i>How Qualitative Research Really Counts</i> Kenneth R. Benoit	9
<i>The American Political Methodology Debate: Where is the Battlefield?</i> Bernhard Kittel	12
<i>Clarifying Comparative-Historical Methodology</i> James Mahoney	19

Symposium II: Necessary Conditions

<i>Necessary Condition Hypotheses as Deterministic or Probabilistic: Does it Matter?</i> Gary Goertz	22
<i>It Ain't Necessarily So—Or Is It?</i> David Waldner	27
<i>Probability Tests Require Distribution</i> Jasjeet S. Sekhon	30

Announcements	31
----------------------------	----

APSA-QM Section Officers

President: Andrew Bennett, Georgetown University
 President-Elect: James Mahoney, Brown University
 Vice President: Elizabeth Kier, University of Washington
 Secretary-Treasurer: Colin Elman, Arizona State University
 Program Chairs: Gary Goertz, University of Arizona
 Ted Hopf, Ohio State University
 Newsletter Editor: John Gerring, Boston University
 Executive Committee: Kathleen Thelen, Northwestern Univ.
 Deborah Yashar, Princeton University
 Deborah Larson, Univ. Calif-Los Angeles
 Ted Hopf, Ohio State University

Letter from the Editor

John Gerring
 Boston University
 jgerring@bu.edu

In this issue, I implored authors to cast aside the usual norms of comity and good taste and, instead, to engage each other directly and unsparingly, leaving no reputation intact. Goaded them on in this unscrupulous fashion I fully anticipated making my reputation in the broader world as an editor provocateur, with attractive options in the commercial world of publishing and cable television (*Esquire?* *Crossfire?* *The O'Reilly Factor?*). I had my sights set.

Unfortunately, the authors insisted upon respecting each others' opinions, even as they thrust and parried. Readers of this issue will find that the two symposia resemble not mud-wrestling but rather equestrian battles, with all due norms of civility observed. Perhaps, in the end, it is more edifying, as well as more trenchant. I do not wish to de-fang the debates...

The first symposium is a wide-ranging (though by no means comprehensive) collection of views on the qualitative/quantitative distinction. Gerry Munck begins with a strong critique of qualitative methods, a field that "rests on a faulty methodological foundation." Ten fallacies inhibit the development of a consensus about what constitutes good methodological practice among qualitative researchers. Andy Bennett takes issue with each of Munck's criticisms, defending the progress that has been made within what might be called (here I resume my role of methodological provocateur) the "qualitative template." Ken Benoit pursues a line of argument that is, depending upon the reader's perspective, an extension of Munck's. Where it is possible to count things, Benoit argues, we ought to do so, for there are many methodological benefits to quantitative research. More important, there is no significant distinction between these two (supposed) forms of knowledge; words are incipient numbers. When we have several similar things we can—in addition to calling them by names—also count them. This offers distinct advantages, in addition to parsimony. Bernhard Kittel strikes out on many fronts, in an attempt to summarize various differences between American and European methodological perspectives. His piece offers a counterpoint to Benoit's, since Kittel—like Benoit, known mostly for quantitative work—is much more critical of the quanti-