What Are Standards Of Rigor For Qualitative Research?
Gery W. Ryan
RAND Corporation

In anticipation of spending two days discussing the question above, I think it would be helpful to make explicit some of my basic underlying assumptions and definitions regarding: (1) What is qualitative research? (2) What are the goals and objectives of qualitative researchers? and (3) What is the range and scope of qualitative data collection and analysis methods? Answering these questions first makes it easier to address issues of standards of rigor and how they might be best applied to the broad range of social science research methods and investigations.

What is Qualitative Research?

The terms "qualitative data analysis" and "qualitative research" are mischievously ambiguous. Does qualitative data analysis mean "the analysis of qualitative data" or "the qualitative analysis of data"? And what specific aspects of "research" does the word "qualitative" modify? The confusion can be eliminated by clearly distinguishing between data and analysis. Figure 1 lays out the possibilities.

Cell *a* is the *qualitative analysis of qualitative data*. Interpretive studies of texts, like transcriptions of interviews, are of this kind. Investigators focus on and name themes in texts. They tell the story, as they see it, of how the themes are related to one another and how characteristics of the speaker or speakers account for the existence of certain themes and the absence of others. Researchers may deconstruct a text, look for hidden subtexts, and try to let their audience know—using the power of good rhetoric—the deeper meaning or the multiple meanings in it.

Cell *d* refers to *numerical or statistical analysis of numerical data*. Lots of useful data about human behavior come to us as numbers. Closed-ended questions in surveys produce numerical data. So do national censuses. Organizations, from businesses to charities to zoos, produce numerical data, too—data about the socioeconomic characteristics of people who use their products or services, data about how often they have to replace managers and secretaries, and on and on.

Cell *b* is the *qualitative analysis of quantitative data*. This can involve the search for patterns using visualization methods, like multidimensional scaling, correspondence analysis, and clustering. Cell *b* is also about the search for, and the presentation of, *meaning* in the results of quantitative data processing. It's what quantitative analysts do after they get through doing the work in cell *d.* It includes everything from the finding of regularities in a scatter plot to the interpretation of meaning and substantive significance of statistical tests. Without the work in cell *b*, cell *d* studies are sterile and vacuous.

Which leaves cell *c*, the quantitative analysis of qualitative data. This involves turning words, images, sounds, or objects into numbers. Scholars in communications, for example, tag a set of television ads from Mexico and the U.S. to test differences in how older people are portrayed in the two countries. Political scientists code the rhetoric of a presidential debate to look for patterns and predictors of policies. Archeologists code a set of artifacts to produce emergent categories or styles, or to test whether some intrusive artifacts can be traced to a source.

Given the dichotomies above, workshop participants may want to consider the standards for data collection as well as for standards of data analysis.

What Are the Goals and Objectives of Qualitative Researchers?

Whether conducting quantitative or qualitative research, social scientists typically address at least one of four fundamental research objectives. The general questions associated with each are shown in Table 1.

1. In exploratory mode, the goal is to discover themes and pattern and to build initial models of how complex systems work. Whether they are investigative journalists tracking a story, archaeologists looking for new sites, ethnographers studying cultural groups, or grounded theorists studying how the elderly experience chronic illness, researchers doing exploratory work follow leads and hunches. They take a step forward, and then they backtrack, trying to uncover what is there, to experience the phenomenon they are studying as fully as possible, and to identify what is common and what is unique.

2. There are four basic types of descriptions: *thematic, case*, *group*, and *cultural*. *Thematic descriptions* present the concepts and themes identified in a corpus of text. *Case descriptions* involve a single case, and include the listing of typical events as well as the listing of idiosyncrasies and exceptions. In *group descriptions*, researchers describe a set of cases (a set of individual people, a set of churches, a set of rituals), noting how individuals are both similar to and different from each other and how the differences are distributed. Group descriptions can be qualitative or quantitative, or both, and vary widely in the amount of precision involved. Some cases call for a broad, sweeping description of a phenomenon. In other cases, we want to know simply whether something is present or not, and if so, how much of it is there? In group descriptions, the amount of precision required influences sampling strategies. Finally, in *cultural descriptions*, researchers describe the culture in which the cases reside. Unlike group descriptions, which focus on the members of a group, cultural descriptions focus on the beliefs and practices *shared* by members of a group.

3. Comparisons can be made at the case and at the group level, and can be qualitative or quantitative or both. They can also be more precise or less precise. In some cases, a statement like "X is bigger than Y" is enough precision. At other times, we need to know exactly how much bigger something is and whether the difference between X and Y could be expected by chance.

4. Model testing includes the systematic assessment of a single hypothesis or a set of multiple, related hypotheses (i.e., a model). This can take place at both the case and the group levels. In fact, in many ways, the testing of models is a special type of comparison where cases and groups of cases are compared to an ideal set of relationships. For example, an investigator using grounded theory might use negative case analysis to assess how well a model fits a single case study. Another investigator, using classic content analysis might test a single hypothesis across many cases.

These basic research objectives are not inimical to one another. In fact, many projects involve exploration, description, comparison, and model testing. Some scholars rely on qualitative data for exploration and discovery, and then use quantitative data for testing models. Increasingly, though, research across the social sciences relies on a balanced, commonsensical mix of both kinds of data.

Given the degree to which research objectives and research methods are intricately intertwined, workshop participants may want to consider whether there should be standards for achieving specific types of research objectives or standards for conducting specific data collection and analysis techniques.

What Are the Techniques for Collecting Qualitative Data?
Over the last century, social scientists have invented hundreds of ways to collect qualitative and quantitative data. Figure 2 shows a rough taxonomy of data collection techniques. Data collection can be broken into three fundamental categories—techniques for indirect observation, techniques for direct observation, and techniques for elicitation—based on the degree to which we interact with the people being studied. The more interaction we have with people whom we study, the more we have to be concerned about reactivity, or response effects. A fourth category, mixed methods, includes combinations of the other three. For example, ethnography, participant observation, case studies and site visits are likely to involve indirect and direct observation, as well a variety of elicitation techniques.

Elicitation techniques can further be categorized into four fundamental types of interviews: unstructured, semi-structured, structured and mixed elicitation tasks. Unstructured interviewing can be either informal or ethnographic. Informal interviews resemble casual conversations and are characterized by a total lack of structure or control. Though ethnographic interviews often look and feel like casual conversations, both the researcher and the informant know they are conducting an interview. Unlike more structured interviewing, an ethnographic interview allows for longer questions and more in-depth probing.

Semi-structured and structured interviews are designed so that each informant is asked a set of similar questions. This is particularly important if the goal is to make comparisons across individuals or groups of individuals. The interviewer initiates the conversation, presents each topic by means of specific questions, and decides when the conversation on a topic has satisfied the research objectives. The respondent in the research interview is led to restrict his or her discussion to the specific questions posed by the interviewer.

Structured interviews are used to measure the magnitude of a phenomenon or to make more precise comparisons within and across groups. The power of such interviews is based on the assumption that if we systematically administer a standardized instrument to a group of people, then variations in their answers are more likely caused by differences among *them* rather by differences in the instrument to which they responded. Structured interviews include surveys and questionnaires, as well as domain elicitation techniques such as free lists, pile sorts, frame elicitation tasks, triad tests, and paired comparisons. Note that structured interviews may include open-ended and closed-ended questions that elicit both long and short qualitative and quantitative responses. Mixed strategies involve a combination of structured, semi-structured, and unstructured techniques. Each method has its advantages and disadvantages. For example, more structured techniques are better for making comparisons and less structured interviews may be more appropriate for early exploratory phases of research.

Workshop participants will need to decide whether to concentrate some of their effort on setting standards for how qualitative data are collected, or to focus on how such data are best analyzed once they are collected.

What Are the Techniques for Analyzing Qualitative Data?
Once the data have been collected, investigators want to: (1) identify themes and subthemes, (2) build and apply codebooks, (3) describe phenomenon, (4) make comparisons, and (5) build, display, test and validate models. The analytic options available to investigators for accomplishing these goals is staggering.

Figure 4 provides a general taxonomy of analysis techniques and research traditions. Like Tesch (1990), the taxonomy distinguishes between the linguistic tradition, which treats text as an object of analysis itself, and the sociological tradition, which treats text as a window into human experience. The linguistic tradition includes methods such as narrative analysis, conversation (or discourse) analysis, performance analysis, and formal linguistic analysis.

Within the sociological tradition, there are two kinds of written texts: 1) words or phrases generated by techniques for systematic domain elicitation; and 2) free-flowing texts, such as narratives, discourse, and responses to open-ended interview questions. Analysis techniques for words and phrases include componential analysis, taxonomies, and mental maps generated by such algorithms as multidimensional scaling, correspondence analysis and cluster analysis. The analysis of free-flowing text include: (1) word-based analyses such as key-words-in-context (KWIC), word counts, semantic network analysis; and (2) code-based techniques such as grounded theory, schema analysis, analytic induction, classic content analysis, content dictionaries, and ethnographic decision making to name a few. Like data collection

techniques, each of these methods of analysis has advantages and disadvantages. Some are appropriate for exploring data, others for making comparisons, and others for building and testing models. Nothing does it all.

The breadth of analytical techniques makes establishing common standards of rigor quite challenging. For example, classic content analysts who seek to test hypotheses, typically assess (and report) inter-coder reliability before moving forward with any additional analyses. In contrast, grounded theorist or schema analysts whose goal is to discover and describe social phenomenon rarely conduct systematic inter-coder reliability tests. To what degree should we expect different methods being used for different purposes to be held to the same standards?

What is meant by standards of rigor and what can we do to enhance them in the social sciences?
Below I offer a series of cautions and then a series of suggestions for how we might standards of rigor standards across such a broad array of goals and methods.

Cautions
First, we need to avoid confusing research rigor with concepts such as measurement precision, quantification, and generalizability. These latter concepts are choices that must be made by each investigator in determining how to best meet his or her research objectives and are not something that should be inherently desired in-and-of-itself.

Second, we need to be cautious about making claims that some data collection or analysis techniques are "more" rigorous than others. If techniques are tools in a researchers' toolbox, then this is like saying that "A saw is better than a hammer because it is sharper."

Third, we need to be careful that we do not overly prescribe "standards" for specific methodological techniques. Methodological techniques are a class of researcher behaviors that share a common set of core properties but include a range of variations and nuances. The power of a research technique lies in its ability to be adapted to multiple research situations. Truncating the variability around a technique will only make the tool less useful.

Fourth, we need to avoid trying to link specific techniques to specific research goals. As tools, methods are a means to an end. It is surprising how such means can be adapted to serve many different goals. For example, I could easily imagine scenarios where paired comparisons could be used to explore, describe, compare, or test hypotheses.

Fifth, we need to stop associating standards and rigor only with confirmatory and hypothesis-driven research. I see no reason why we cannot set standards of rigor for exploratory and descriptive research as well. I suspect that some of the criteria will vary based on specific research objectives, while some of the criteria will cut across all types of research. More on this below.

Suggestions
First, I would argue that rigorous research is research that applies the appropriate tools to meet the stated objectives of the investigation. For example, to determine if an exploratory investigation was rigorous, the investigator would need to answer a series of methodological questions such as: Do the data collection tools produce information that is appropriate for the level of precision required in the analysis? Do the tools maximize the chance of identifying the full range of phenomenon of interest? To what degree are the collection techniques likely to generate the appropriate level of detail needed for addressing the research question(s)? To what degree do the tools maximize the chance of producing data with discernable patterns? Once the data are collected, to what degree are the analytic techniques likely to

ensure the discovery of the full range of relevant and salient themes and topics? To what degree do the analytic strategies maximize the potential for finding relationships among themes and topics? What checks are in place to ensure that the discovery of patterns and models is not superfluous? Finally, what standards of evidence are required to ensure readers that results are supported by the data? The challenge for workshop participants will be to identify what questions are most important for establishing research rigor and to provide examples of how such questions could be answered for those using qualitative data.

Second, I would argue that rigorous research must be both transparent and explicit. In other words, researchers need to be able to describe to their colleagues and their audiences what they did (or plan to do) in clear, simple language. Much of the confusion that surrounds qualitative data collection and analysis techniques comes from practitioners who shroud their behaviors in mystery and jargon. For example, clearly describing how themes are identified, how codebooks are built and applied, how models were induced would help bring more rigor to qualitative research.

Third, we all need to become more familiar with the broad range of methodological techniques available to us. Social science has become methodologically parochial. Content analysts, grounded theorists, semantic network analysts, and analytic inductionists do not talk to each other. Cross-fertilization across methodological traditions, especially those that are dominated by a single discipline, is a rare event. Even more worrisome is the growing tendency for researchers to attack all problems with the same type of methodological hammer.

Fourth, reviewers of manuscripts and proposals need to be better selected, trained, and supervised. Reviewers should provide feedback as to whether the question or topic is of interest to the readership or funding agency and the degree to which objectives stated in the proposal or manuscript are met by the methods and data presented. Unfortunately there is a tendency for some reviewers to try to change the objectives of research to match their own methodological expertise. The classic example occurs when researchers conducting exploratory research are criticized for not using techniques that are more appropriate for hypotheses testing. The opposite, however, also occurs. On numerous occasions I have seen "qualitative" researchers insist that their colleagues use more unstructured data collection approaches even though these were less appropriate for their stated research objectives. Reviews would be more constructive if reviewers had had broader methodological experience and authors improved their ability to clearly express what they plan to do and why.

Fifth, there needs to be more methodological research on the basic techniques for identifying themes, building codebooks, marking texts, making comparisons and inducing models. There are many different methods for accomplishing these tasks and it is unclear to what degree they produce different results. For example there is debate among grounded theorists about whether to follow the different techniques exposed by Glazer or Strauss. To my knowledge, there is no empirical evidence suggesting one is better than the other. Likewise, there are many different techniques of identifying themes in texts but again little to no empirical evidence about what works best and under what conditions.

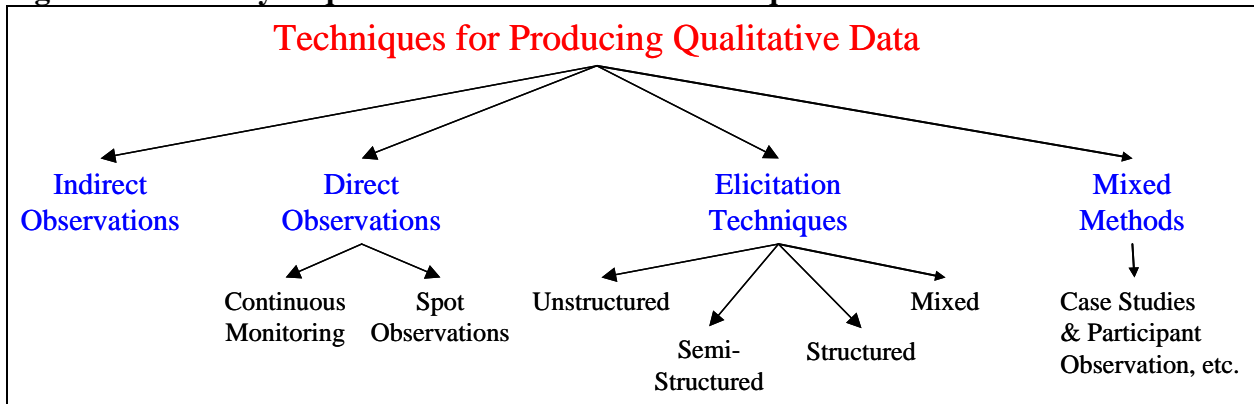**Figure 1 Key qualitative and quantitative distinctions**

	Data	
Analysis	**Qualitative**	**Quantitative**
Qualitative	A Interpretive text studies. E.g., Hermeneutics, Grounded Theory, Phenomenology	B Search for and presentation of meaning in results of quantitative processing
Quantitative	C Turning words into numbers. E.g., Classic Content Analysis, Word Counts, Free Lists, Pile Sorts, etc.	D Statistical & mathematical analysis of numeric data

Adapted from: Bernard, H. Russell. 1996. Qualitative data, quantitative analysis. *Cultural Anthropology Methods Journal* 8(1):9-11
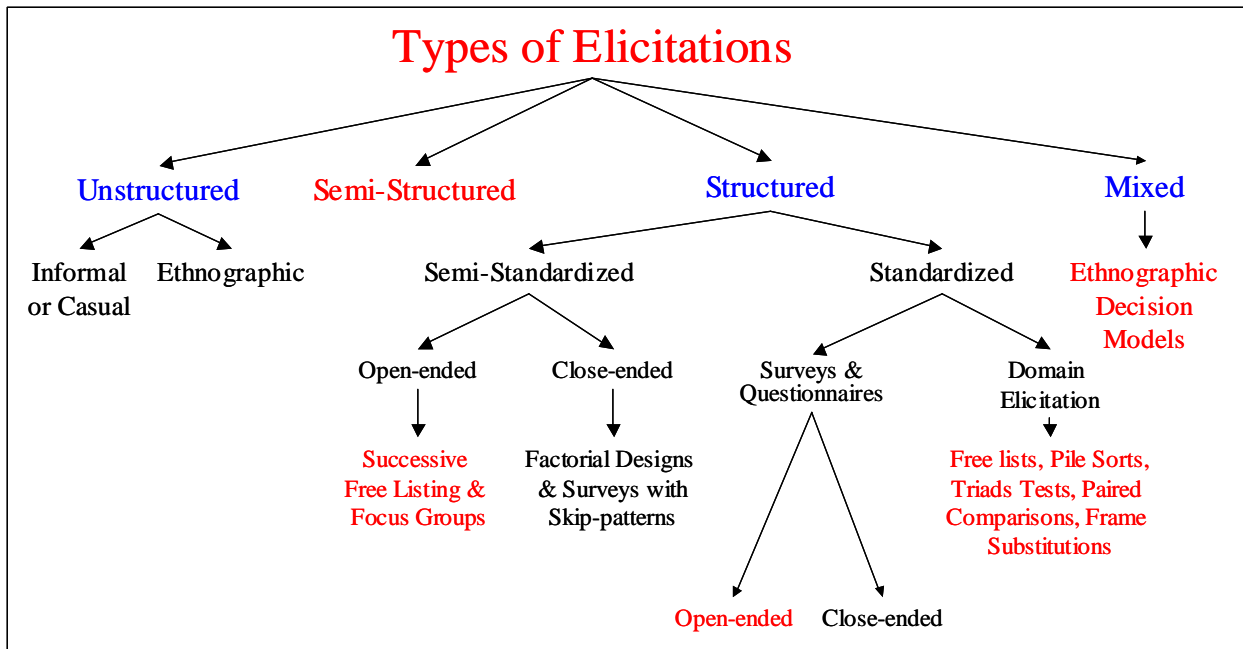
Table 1 Goals of qualitative research

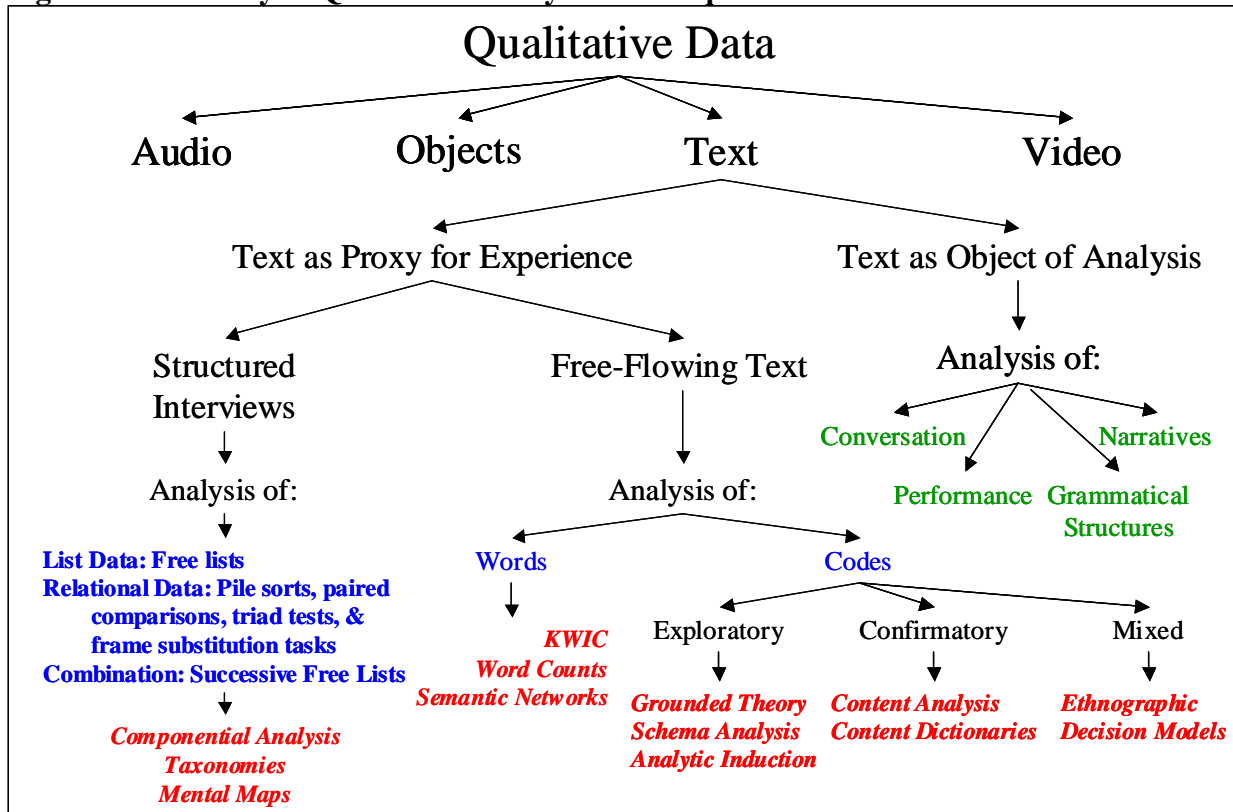| General Aim	Type	Questions
1. *Exploration*		What kind of things are present here?
		How might these things be related to each other?
2. *Description*		
	Thematic	What does a theme look like?
	Case	What does a case look like?
	Group	What does a set of cases look like?
		Is a particular kind of thing (A) present or not?
		How much of that kind of thing (A) is there?
	Cultural	What does the culture look like?
3. *Comparison*		
	Case	How is Case X different from Case Y?
	Group	How is Group Xs different from Group of Ys?
4. *Model Testing*		
	Case	To what degree does a particular case conform to the proposed model?
	Group	To what degree does a group of cases conform to the proposed model?

Figure 2 Taxonomy of qualitative data collection techniques

Techniques for Producing Qualitative Data

Indirect Observations

Direct Observations

Continuous Monitoring

Spot Observations

Elicitation Techniques

Unstructured

Semi-Structured

Structured

Mixed

Mixed Methods

Case Studies & Participant Observation, etc.

Figure 3 Taxonomy of elicitation techniques

Types of Elicitations

Unstructured

Informal or Casual

Ethnographic

Semi-Structured

Semi-Standardized

Open-ended

Successive Free Listing & Focus Groups

Close-ended

Factorial Designs & Surveys with Skip-patterns

Structured

Standardized

Surveys & Questionnaires

Open-ended

Close-ended

Domain Elicitation

Free lists, Pile Sorts, Triads Tests, Paired Comparisons, Frame Substitutions

Mixed

Ethnographic Decision Models

Figure 4 Taxonomy of Qualitative Analysis Techniques

Adapted from: Ryan & Bernard, 2000 Data Management and Analysis Methods. In: Handbook of Qualitative Research, 2nd ed., N. Denzin and Y. Lincoln eds., Thousand Oaks, CA: Sage Publications, 769-802..

Bibliography

Tesch, R. 1990. *Qualitative research: analysis types and software tools*. New York, The Falmer Press.

Bernard, H. Russell. 1996. "Qualitative data, quantitative analysis". *Cultural Anthropology Methods Journal* 8(1):9-11

Ryan & Bernard, 2000. *Data Management and Analysis Methods. In: Handbook of Qualitative Research*, 2nd ed., N. Denzin and Y. Lincoln eds., Thousand Oaks, CA: Sage Publications, 769-802.