palgrave pivot

Counting as a Qualitative Method

Grappling with the Reliability Issue in Ethnographic Research

Wayne Fife

palgrave macmillan

Counting as a Qualitative Method

Wayne Fife

Counting as a Qualitative Method

Grappling with the Reliability Issue in Ethnographic Research

palgrave
macmillan

Wayne Fife
Department of Anthropology
Memorial University of Newfoundland
St. John's, NL, Canada

ISBN 978-3-030-34802-1 ISBN 978-3-030-34803-8 (eBook)
https://doi.org/10.1007/978-3-030-34803-8

This Palgrave Pivot imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Acknowledgments

Earlier field and archival research work in Papua New Guinea, London, England, and on the island of Newfoundland was supported by a Social Sciences and Humanities Research Council of Canada Doctoral Fellowship, a Social Sciences and Humanities Research Council of Canada Postdoctoral Fellowship, or a Social Sciences and Humanities Research Council of Canada Standard Research Grant. I thank them for it. I also wish to offer my thanks to the anonymous reviewers of my manuscript and note that I have made use of some of their insightful comments for the improvement of the book. Manuscript review is often a thankless task, but our research fields would be less rich without the selfless work that so many do on behalf of all of us. Special thanks to Mary Al-Sayed, Commissioning Editor for Sociology and Anthropology at Palgrave Macmillan. She has been astonishingly helpful throughout the entire process and displayed unfailing good humor while dealing with what can sometimes be a complicated and frustrating endeavor. My gratitude also to Madison Allums, Editorial Assistant for Sociology, Anthropology, and Psychology at Palgrave Macmillan, who provided thoughtful and very helpful suggestions and guidance while seeing me through the tricky task of final manuscript preparation for publication. And finally, to my wife, the anthropologist Sharon Roseman. Sharon has always been my first reader and is a professional quality editor in her own right. Never afraid to let me know when the content of my writing might be improved and always willing to correct my less than stellar grammar, I rely on her to tell me the truth about my work. This book is dedicated to her, the love of my life.

Contents

1	The Reliability Issue	1
2	Creating a Counting Schedule	23
3	Success, Failure, and a Missed Opportunity	33
4	Counting Qualitative Results	47
5	Counting in the Archives	63
6	Tourism: Counting the Overlooked	81
7	Making Fiction Count	99
8	The Importance of Counting for Qualitative Research	121
	References	129
	Index	137

List of Tables

Table 2.1 Disciplinary actions in the classroom (raw numbers) 27
Table 2.2 Disciplinary actions per teaching hour 28
Table 3.1 Teacher interactions with male/female students (raw numbers) 35
Table 3.2 Raw percentage of teacher questions for male/female students for all classes combined 36
Table 4.1 My future work (Kimbe Community School) 51
Table 4.2 My future work (Bialla Community School) 52
Table 4.3 My future work (Ewasse Community School) 52
Table 4.4 Good versus bad students (interview material) 58

CHAPTER 1

The Reliability Issue

Abstract This chapter outlines the problem of reliability in ethnographic or qualitative approaches, strongly suggesting that incorporating counting in both the research and analysis stages of projects can help deal with the issue. After discussing several standard ways to check on the reliable gathering of information (e.g. length of study, reflexivity, saturation, and triangulation), the author describes examples from several decades of ethnographic research and makes the case for adding counting as a standard component for all qualitative fieldwork. The author briefly describes why he decided to create counting schedules in order to check on the results of more traditional ethnographic methods, such as interviewing, participant-observation, and self-reporting techniques. An explanation is given for why counting is not to be confused with standard forms of statistical or quantitative research. A case is also made for learning to count for both research and analytical purposes manually, rather than uncritically relying on various forms of Qualitative Data Analysis programs.

Keywords Qualitative reliability • Ethnographic methods • Validity • Counting

W. Fife, *Counting as a Qualitative Method*,
https://doi.org/10.1007/978-3-030-34803-8_1

The Problem

A key concern for any kind of ethnographic research is the question of reliability. How can we know when the information we gather through participant-observation, interviews, self-reporting, and other ethnographic research methods[1] is reliable enough to form the basis for a well-tempered portrait of a group of people? How do we take care not to create caricatures out of human beings, or skew our results by relying too heavily on either the words of a few insiders or on our own preconceived notions?

Over the years, a number of qualitative checks have been developed. This is not the place to review them, but I will mention a few for illustrative purposes. One way is to make sure that each important interpretation (e.g. about a marriage pattern, economic practice, or cosmological point of view) has made use of two or more specific methods for gathering information for analysis. That is, we can ask ourselves whether or not the information gathered from interviews aligns with that from observations, researcher participation with self-reporting? The assumption is that analysis based on two or more sets of data at the same time has a better chance of being reliable than something that is based solely on one method of information gathering. This is often referred to as triangulation (e.g. Fetterman 1989: 91; Holliday 2002: 43; DeWalt and DeWalt 2011: 128; Glesne 2011: 47; Brynman and Bell 2016: 306–307; Scott and Garner 2013: 185). In some circumstances, this can even involve multiple researchers recording the same event so that notes can be compared later (e.g. DeWalt and DeWalt 2011: 113).

An alternative to this is the saturation approach (e.g. Glaser and Strauss 1967), in which the researcher constantly compares inductively gathered information within the same category or subject in order to ensure that our explanations reach a level of internal reliability (e.g. Maykut and Morehouse 1994: 126–149; Brynman and Bell 2016: 270). Saturation is reached by building up the bits and pieces of information for a given topic (e.g. naming practices, joking behavior, or a particular religious ritual) until no new information about that practice or belief is forthcoming. Saturation is an ideal and is never fully achieved, but there will be a moment when diminishing returns inform the researcher that it is time to move on to a new topic.

When using either saturation or methodological comparison we have to be sure to include notes about the contradictions and disagreements we find among the various members of a group. Only in this fashion can we properly construct nuanced depictions of disparate ways of life—even as they occur within more or less common economic, social, and cultural patterns (e.g. Fetterman 1989: 35).

The long-term nature of ethnographic research projects can also help provide checks and balances, increasing the trustworthiness of our findings (for illustrative examples of how this can be done, see Pelto and Pelto 1978; Angrosino 2007). As David Fetterman (1989: 46) suggests: "Working with people, day in and day out, for long periods of time is what gives ethnographic research its validity and vitality."

A form of methodological reflexivity can also serve as a check on our work. In this situation, researchers strive to produce and retain very clear records about what they do throughout the research project, so that others (or even themselves) might 'audit' their work at a later date to check for consistency and other reliability issues (e.g. DeWalt and DeWalt 2011: 184–185; Scott and Garner 2013: 243; Brynman and Bell 2016: 169). The idea here is that others should be able to follow the way we gathered our evidence and critically assess it. Not everyone agrees with this idea. Brinberg and McGrath (1985, 13), for example, suggest that this unnecessarily implies that something can ever achieve 'full validity,' rather than being an ideal toward which we work. Roger Sanjek has noted that any suggestion that others would be able to actually replicate a singular fieldwork experience is spurious. As he puts it (Sanjek 1990. 394): "In ethnography, 'reliability' verges on affectation." Reflexivity, on the other hand, need not be aimed at replication. For most researchers, it is about letting the reader or viewer know enough about who they are and how they conducted the research so that they might better judge the validity of the work being offered to them (e.g. Sanjek 1990; Emerson et al. 2011). The authors just cited suggest that this kind of useful reflexivity can and should be embedded within our field notes, which we can then draw upon as required.

My own position is that reliability or validity do not have to be seen as an either/or proposition. I am in agreement with David Brinberg and Joseph McGrath, who have considered the issue from almost every angle. What they have ended up concluding is that "validity is like integrity, character, or quality, to be assessed *relative to purposes and circumstances*" (Brinberg and McGrath 1985: 14). What I am suggesting, therefore, is that incorporating more counting into our standard qualitative projects at both the research and analytical levels will give us one more way to help assess our work *relative to purpose*. At the same time, there is nothing privileged about numbers.

In the way that numbers are being used here, they are neither better nor worse than prose or other forms of investigation or analysis. As noted in the book numerous times, qualitative counting can be used effectively only *in conjunction* with other qualitative methods.

The issue of reliability plagued me from the very first time I did an extended ethnographic study. This occurred during my M.A. degree in anthropology, when I carried out eight months of fieldwork in a home for the elderly in Southwestern Ontario (Fife 1983). The research was conducted through lengthy visits to the home, broken up into extended periods of daily work, two or three times per week. At the time, I was heavily influenced by what became known as symbolic or interpretive anthropology. My main focus was on trying to figure out how rural people, who had previously overwhelmingly subscribed to a cultural belief in individual 'independence' and 'self-reliance,' coped while living within an institution that redefined them as dependent human beings. Using the primary research methods of participant-observation, event analysis (e.g. birthday parties, family days, special events), and semi-structured and unstructured interviews, I came to learn about the ways that many residents used to reappropriate this government-run institution and reconstitute it as a 'home away from home' and the workers in it as 'just like family.' In keeping with their pre-home understandings of life, family was defined as the people who were *supposed* to look after each other. Therefore, by symbolically reconstituting both paid workers and the other inhabitants as analogues to one's own family, residents were simultaneously recasting themselves as being 'entitled' to the help and care the government institution supplied. This obviated the need to view themselves as 'charity cases' who were receiving 'hand-outs,' or to accept negative connotations they would have previously associated with social and economic dependency. The to-and-fro that occurred between residents, actual family members, and both workers and volunteers (who did in fact see the residents as dependent human beings) was fascinating to undercover. Still, I worried about relying too heavily upon what people said about themselves and not enough on what people were actually doing. How did I really know if my analysis was correct? Event analysis and fairly intensive periods focusing on the observation portion of participant-observation, while following the 'saturation' and 'triangulation' methods of data checking noted above, helped give some confidence in my findings. But I remember wishing that I had another method to check some of my most important explanations and understandings. Although I recorded the number of residents, staff members, and volunteers—along with other similar kinds of information—I never really thought of counting as having anything substantial to do with ethnographic research. It was strictly for background material, just part of the context of 'real' ethnographic data collection and analysis.

Anxiety over the reliability issue intensified further when I went for a year (1986/1987) of fieldwork in the province of West New Britain in the country of Papua New Guinea. My goal was to conduct ethnographic research for my Ph.D. in anthropology. The focus of the study was on what I could find out about the kinds of connections that existed between formal education, social change, and cultural continuity (e.g. Fife 1992, 1995a, b). Having moved to a different university for the new degree, my theoretical interests had also shifted. Without losing interest in the more symbolic aspects of life, my work now took on a stronger materialist grounding. This means that economics and the structural constraints of social formations carried a new importance for me.

In the late 1980s, the majority of Papua New Guineans received six or less years of formal education in community schools. These schools followed an externally introduced (i.e. European model) of education. Only 'elite' students, as defined by rigid examinations (at grades 6, 8, and 10) *and* by the ability and willingness of a family and/or a larger kinship group (such as a lineage) to financially support additional years of education, could ensure a chance at finishing standard high school (grade 10) or potentially going on to attend one of the very few grade 12 schools in the country. These special grade 12 high schools served as direct conduits into tertiary forms of education (e.g. the University of Papua New Guinea or Teachers Colleges). My focus was on the first six years of this process and the impact community schools had on families and the larger society as a whole. How did it dovetail with employment opportunities? What impacts did it have on already existing cultural formations?

There were a lot of moving parts to contend with—far more than I had faced in my earlier work. This went along with a concomitant rise in concern about whether or not I was gathering useful information (consistent and pattern-producing forms of evidence) which could be used to effectively teach me about the role of community schools in contemporary life. I studied three schools (two 'urban' and one 'rural,' though distinctions among the student body were not always so clear-cut), involving 26 teachers, 3 headmasters, other local leaders, various educational officials, many kinds of civil servants, parents, and the structural constraints (both historic *and* contemporary) that informed the present situation. There were clashes at parent-teacher meetings, differences of opinion among educational officials, very disappointed school-leavers who almost never realized their earlier dreams of a life in the new cash-based and urban-defined social economy, and so many other things with which to contend.

After approximately six or seven months of research, I felt I was starting to get a handle on the most important social and educational patterns. For example, I was beginning to predict which issues parents, teachers, students, or government officials would care enough to fight about during meetings. And those they were willing to work on with a goal of reaching consensus—the widespread ideal used in decision-making situations in most local cultural formations.

Still, there were concerns. I distrusted some of my own biases and ethnocentric assumptions, and worried about the inevitable overinterpretation a researcher often engages in after becoming excited about beginning to understand specific patterns of behavior and thought. Having spent many months taking extensive and wide-ranging notes, maybe it was time to focus more tightly? I wanted to check some of my most important tentative conclusions—but how?

Because I was very interested in the theoretical concept of hidden curriculum, I had brought along an article written by the educational anthropologists Frederick Gearing and Paul Epstein (1982). This specific piece focused on a research method for doing a form of micro-analysis that could be used to reveal the details of hidden curriculum as it was actually having an impact.

The best way to understand hidden curriculum is to realize how it contrasts with overt curriculum. The latter involves the math, science, language, health, and other lessons that a student is expected to learn within a school. That is, it makes up the content of classroom instruction. When students learn that two plus two equals four, or that the capital of Papua New Guinea is Port Moresby, they are learning overt curriculum.

Hidden curriculum, conversely, is the context that both informs *and* contains its own (hidden) lessons. In the classroom, a good example might be found in the way a teacher favors male over female students when asking questions in mathematics or science classes, arranges desks or other physical space within the room (e.g. putting the 'well-behaved' students at the front of the class), hangs images or other visual clues on walls, disciplines some students while ignoring others, and so on. The forms (physical and behavioral) that human interactions occur within influence (and can even contradict) the lessons of overt curriculum. For example, students might be taught in a social studies class that males and females are social equals—the official lesson regarding gender egalitarianism as outlined in their schoolbooks. But what if the same teacher consistently calls upon male students to participate in classroom discussions involving lessons about the political structures of their country and suppresses female participation in these discussions? In this case, the hidden curriculum can have a much greater impact than the overt curriculum over what both male and female students come to believe about the 'real' politics of gender.

Unlike official kinds of overt educational curriculum, hidden curriculum can easily be seen at work outside of classrooms. The structure of teacher/parent meetings, the roles students are assigned in looking after school grounds, the varieties of play allowed at recess, how assembly proceeds prior to letting students into school buildings, and many other sets of interactions carry their own key effects. Even segregating a 'school' through the use of a fence carries a very strong message about the supposed naturalness of separating education off from the other concerns of life (e.g. making a living).

I felt I had identified several key trends in the hidden curriculum of the classrooms of West New Britain. But was I correct about them? Gearing and others (e.g. Gearing and Epstein 1982; for added context see Gearing 1979, 1984) developed a method by which they used movie cameras to record classroom behavior, so that they could do a micro-analysis of that behavior later in order to identify key trends in hidden curriculum. Here, the focus was squarely upon teacher-student interactions. My interests were broader than this. Besides, I did not have a video camera with me, or any way to obtain one. Simultaneously, I had always wanted my research methods to be intelligible and accessible (e.g. potentially replicable) to not only other researchers but also the teachers, educational officials, and other individuals who might be interested in conducting their own studies in the future. This mitigated against the effectiveness of more technical and detailed forms of micro-study. What I really wanted was to check on patterns of hidden curriculum that I believed I had already identified through many hours of classroom observations and the painstaking coding (for key themes) I had done on my own research notes (on coding, see Fife 2005: 74–83). Ideally, I wanted to be able to do this through a method that required no more than a pencil and paper to implement and which could be understood by anyone who wanted to check on the findings of my work (or their own future research). Serendipitously, I also happened to remember what the Canadian anthropologist Richard Salisbury often said to his audience at conferences. "When in doubt, count."

The Beginnings of a Solution

Brainstorming about the issue for several days, I decided to wed the sentiment of Salisbury with the spirit of Gearing and Epstein's methodology. I would do fine-grained analysis, focusing on the specific forms of hidden curriculum that I wanted to check for reliability. And I would do it with a pencil and a piece of paper.

With hindsight, I realize I was lucky that I tried the method out first in more tightly controlled classroom situations. Later, I would come to the understanding that the method and the basic principle behind it could be broadened to a much wider range of situations. But classrooms allowed me to work out several technical kinks without having to deal with too many distractions simultaneously.

The first thing I decided to check was a form of behavior I had witnessed repeatedly at every grade level and in each of the three schools in the study. I came to think of these behaviors as 'kinds of discipline.' I had extensive coding in my notes for many variations on the forms of discipline that were used by teachers to attempt to modify the words and actions of their students within classroom situations. I was starting to identify different types of discipline, but my notes were uneven. Some varieties showed up much more frequently than others, such as disciplining a specific individual versus the collective student body. Was this a true reflection of the relative usage of particular kinds of disciplinary actions, or was it simply what I happened to notice at the time while I attempted to fully record everything that was happening in the classrooms? It seemed to me that the pattern of discipline changed from the lower to the higher grades. Was this correct? I decided to see if I could count specific forms of disciplinary actions at work in the classroom. If so, would it allow me to be able to compare them, both within a given school and across the three schools? I will give the details about how I did this in the next chapter, along with the kind of analysis it allowed me to carry out. For now, I simply want to note that it not only answered many questions, it also helped me consider information I already had in new ways *and* recognize that the patterns that I saw were related to and sometimes replicated in other extra-classroom issues (such as assemblies and parent-teacher meetings). However, that was not the end of the benefits of this particular methodology. A few years later, while doing post-doctoral research in the archives at SOAS (School of Oriental and African Studies) in London (England) about the missionaries who had formed the initial educational systems in Papua New Guinea,