


Counting in Qualitative Research: Why to Conduct it, When to Avoid it, and When to Closet it

Journal of Management Inquiry
20(1) 14–22
© The Author(s) 2011
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1056492610375988
http://jmi.sagepub.com


David R. Hannah¹ and Brenda A. Lautsch¹

Abstract

In this essay we discuss the issue of counting: the process of assigning numbers to data that are in nonnumerical form. We review why counting is a controversial issue in qualitative research, and explain how this controversy creates what we call the “multiple audience problem” for qualitative researchers. We then identify the purposes that can be served by four different types of counting, explore when counting should be avoided entirely, and discuss when the results of counting should be concealed, or as Sutton put it, kept in the closet.

Keywords

research methods, qualitative research, qualitative data analysis, counting

In 1979, Miles, when lauding the strengths of qualitative data as “full, rich, {and} earthy” (p. 590) argued that it was an “attractive nuisance” because there was so little agreement, even among qualitative researchers, over how to analyze and write up qualitative data. Over 30 years have passed since he posed that argument, but recent work suggests that not much has changed. Although qualitative (and quantitative) researchers may desire to produce research that is well written, methodologically appropriate, and that makes a theoretical contribution, there is still no “boilerplate” for how to achieve these aims for qualitative researchers (Easterby-Smith, Golden-Biddle, & Locke, 2008; Pratt, 2009). One outcome of this lack of consensus is that qualitative researchers often face uncertainty about (a) how to conduct good research and (b) how to convince reviewers, editors, and the broader audience of readers of the quality of their work.

Ultimately, this general uncertainty manifests itself in more specific arguments about the merits of particular philosophical and methodological approaches. One such debate has arisen on the topic of counting: the process of assigning numbers to data that are in nonnumerical form. Some experts on qualitative research have written about the benefits that can come from counting, including its ability to help qualitative researchers produce more valid, persuasive arguments (e.g., Lee, 1999; Miles & Huberman, 1994; Silverman, 2000). Other scholars have emphasized the risks of counting, noting that it often will be inconsistent with the underlying assumptions held and the goals pursued by qualitative researchers, and therefore can reduce the quality of qualitative research (e.g., Fineman & Mangham, 1983; Gephart, 2004; Suddaby,

2006). As a result, qualitative researchers, as well as those who read and review qualitative research, can have a difficult time deciding when counting is and is not an appropriate approach to qualitative data analysis. The goal of this article is to provide qualitative researchers with advice to help them navigate this contested terrain.

Our approach was informed (and in some senses provoked) by Robert Sutton’s (1997) essay entitled “The Virtues of Closet Qualitative Research.” His central argument was that there are “. . . times when it is best to conceal or downplay the role that qualitative data played in developing an author’s ideas” (p. 97). In concluding his essay, he noted that it might be possible, though difficult, to similarly explore situations where closeted quantitative research might occur. In the present article we have explored one aspect of closeted quantitative research by examining instances where the counting or quantification of qualitative data is hidden. We have also gone somewhat further in two ways. First, we identify four different forms of counting and explore the purpose and potential benefits of each one. In addition, we suggest that there are times that the quantification of qualitative data should be avoided entirely, rather than conducted and then hidden.

¹Simon Fraser University, Burnaby, B.C., Canada

Corresponding Author:

David R. Hannah, Simon Fraser University, Faculty of Business Administration, 8888 University Drive, Burnaby, B.C., Canada
Email: dhannah@sfu.ca

To be clear, our focus is on qualitative research that aims to formulate new theory. In such research the primary aim is typically to develop rich and accurate understandings of an issue or phenomenon (Cook & Campbell, 1979). Although some have argued that qualitative research may be designed to test theory (Lee, 1999), such research is likely to place a higher priority on statistical means of assessing validity and generalizability, and as such the use of counting will be more often simply a requirement. We focus on theory-advancing qualitative research because we believe it is more likely to present qualitative researchers with difficult choices about quantitative methods and we believe that we can offer useful advice about those choices.

We begin by discussing why decisions about counting have been so challenging for qualitative researchers. Next, we try to provide some clarity to that decision process by exploring in detail the various types of counting. We then consider when counting should be avoided entirely, and when it should be closeted.

Counting and the Multiple Audience Problem

As noted in the introduction, there is little consensus about what metrics are appropriate for assessing the quality of qualitative research. Lee (1999) argued that the traditional concepts of reliability and validity are compatible with both quantitative and qualitative research. Marshall and Rossman (1995) highlighted credibility, transferability, replicability, and confirmability as alternate standards that should be used to judge qualitative research. Golden-Biddle and Locke (1993) pointed instead to criteria such as authenticity, plausibility, and criticality. Given the multiplicity of metrics to choose from, it is possible that different researchers will approach their own research, and the specific decision about whether or not to count their qualitative data, having come to different conclusions about what “good” qualitative research entails. Importantly, so may potential readers of their work, including editors and reviewers.

The lack of consensus on how to go about evaluating qualitative research creates what we call the “multiple audience problem” for qualitative researchers. The problem can be summarized in the following terms. Different audiences for qualitative research may have different beliefs about what constitutes good qualitative research. On occasion, those beliefs will contradict one another. Therefore, if qualitative researchers try to satisfy the preferences of one audience, they risk not meeting the preferences of another audience. Even if researchers placate their editorial audience and succeed in getting an article published, they may still find that other audiences in the broader community of organizational scholars come to negative assessments about the merits of their methodological choices.

We believe that counting presents a particularly tricky multiple audience problem due to its contentiousness. On the one hand, many qualitative researchers are highly skeptical of the usefulness of counting. This skepticism can be attributed to at least two causes. First, as many qualitative researchers are aware, counting is not an appropriate strategy for all forms of qualitative inquiry. If counting is combined with an interpretivist approach, for example, it may result in an inconsistency between the assumptions of that approach and the methods used to test its research question (Suddaby, 2006).¹ Second, many qualitative researchers have had disagreements with reviewers or editors over whether counting was warranted in a given article (Pratt, 2008). As a result, some qualitative researchers—out of principled objections to counting in their research and perhaps out of resistance to forced counting—are strongly predisposed against any form of counting in qualitative research.

At the same time, our field continues to be dominated by scholars who subscribe to economic, quantification-based perspectives (Amis & Silk, 2008; Ashforth, 2005; Bartunek & Seo, 2002; Pfeffer, 1995; Symon, Buehring, Johnson, & Cassell, 2008). As a result, many of the scholars who read, review, and edit our articles are likely to be steeped in those perspectives. They may be more likely to value counting because it can provide evidence, in the form of numbers such as interrater reliability statistics (e.g., Miles & Huberman, 1994) or frequency counts, that the findings were derived by means of a rigorous, objective analysis of the qualitative data.

Qualitative researchers therefore find themselves in a dilemma when it comes to counting, because choices over counting are likely to satisfy some audience members while alienating others. The existence of this dilemma is what initially motivated us to write on this topic. We felt that the field could benefit from a set of guidelines that both authors and reviewers could consult to decide when counting is an appropriate approach to qualitative data analysis. As our work proceeded, we realized that authors’ decisions over counting might evolve during the research and publishing process; a researcher may collect and count qualitative data, only to conclude that the qualitative data alone is stronger, or more interesting, or likely to be preferred by the audience at a particular journal targeted for publication. We concluded that at times qualitative researchers should conceal their counting; or, as Sutton (1997) put it, they should closet their counted data.

Next, we continue our exploration of these issues by considering why researchers might choose to count their data. We suggest that there are four types of counting, each of which serves a specific purpose: autonomous, supplementary, corroborative, and credentialing. Later, we consider when counting should be avoided entirely and when counting should be closeted.

Autonomous Counting

The purpose of autonomous counting is to produce numbers that are intended to stand on their own as significant research findings. The fate of qualitative research that uses autonomous counting, in the sense of whether or not it will be accepted for publication and then recognized as making a contribution to the field, will depend substantially on the degree to which readers of the article believe that the numbers produced constitute, *in and of themselves*, valid and convincing evidence of the claims being made by the authors.

A key benefit of autonomous counting is that it can enable authors to develop a summary of the entire data set that authors can then scrutinize to discern patterns in the data (Miles & Huberman, 1994; Silverman, 2000). For example, Dutton, Ashford, O'Neill, and Lawrence (2001) used autonomous counting in their study of how managers attempted to influence organizational change by "selling" certain issues to top management. They identified a number of issue selling approaches, and for each one, they counted the number of times managers mentioned that approach in the context of successful issue selling episodes and unsuccessful episodes. When certain approaches appeared considerably more often in successful episodes, it was seen as evidence that managers believed that those approaches were important to successful issue selling.

Supplementary Counting

Unlike autonomous counting, supplementary counting is not intended to produce the central contribution of the research. Nor is it designed to confirm other findings like the next type of counting we will discuss, corroborative counting. Supplementary counting builds on other findings and adds to them, enabling researchers to develop new insights into their phenomena of interest.

In Barley's (1986) seminal work on how two hospitals restructured when a new technology, CT Scanners, was introduced into the hospitals, he used both supplementary and corroborative counting. The overall goal of his research was to investigate how the introduction of this new technology influenced the structure of the radiology departments of those hospitals. He found that "scripts," or role-defining interaction patterns between radiologists and technologists, had a key effect on how the structures of the departments evolved over time to accommodate the new technology. To develop a more in-depth understanding of how scripts influenced structuring, he counted the frequency with which each script occurred and tracked patterns in the use of different scripts across the phases of restructuring at the two hospitals. His use of supplementary counting allowed him to further develop his ideas regarding how technology can occasionally shift interaction scripts, which in turn reshapes organization structure.

Corroborative Counting

Corroborative counting is typically associated with a conventional triangulation approach involving a combination of qualitative and quantitative methods (Jick, 1979). In such an approach, counting is used to verify the conclusions reached by a purely qualitative analysis of the data. The logic underlying triangulation is that one can be more certain of a result if the same findings occur in two separate processes—purely qualitative analysis and quantitative analysis of counted qualitative data—each of which is subject to different biases and flaws.

An example of corroborative counting can be seen in the article written by Maguire, Hardy, and Lawrence (2004). They studied the legitimacy of actors in a particular institutional context and examined how two key "institutional entrepreneurs" were particularly influential in bringing about institutional change. To identify these two entrepreneurs, they used both qualitative analysis and counting. Qualitatively, they analyzed the interviews and concluded that there were two actors who were identified by others as being instrumental to the changes. They corroborated the decision to focus on those two actors by counting the number of legitimacy-conferring characteristics (including personal characteristics, characteristics of their roles, and characteristics of the organizations they belonged to) that each actor was perceived to have. As their two actors had a higher number of legitimacy-conferring characteristics than all the others, the authors were able to corroborate their decision to focus on them in more depth.

Credentialing Counting

We have called the fourth and final type of counting credentialing counting. The purpose of credentialing counting is to demonstrate why one should have confidence in the findings of a qualitative analysis. This type of counting typically does not produce findings of its own. Instead, it focuses on either (a) documenting counts of data sources, or (b) generating evidence of the analytical honesty of researchers.

It is common for qualitative researchers to provide counts of their data sources. Reay, Golden-Biddle, and Germann (2006, p. 982), for example, developed a table of the numbers of different types of respondents interviewed, the numbers of interviews and meetings held, and the amount of transcribed and archival material analyzed (1,405 pages total). Corley and Gioia (2004) presented two tables in their methods section, labeled "quantitative details of the interview data" (p. 182) and "quantitative details of documentation data" (p. 183). The presentation of data source counts can provide transparency about how the research was conducted, and they can also provide some insights into the amount of work that was done in the research.

Credentialing counting can also be used to construct evidence of the analytical honesty of qualitative researchers. Those researchers who are sensitive to the need to demonstrate analytical honesty may provide one or both of two types of evidence: (a) the representativeness of findings and (b) the objectivity of coding processes.

Some experts have encouraged qualitative researchers to develop findings that are representative of their phenomenon of interest (Lee, 1999), in the sense that the findings are an accurate reflection of the data. For example, Dutton et al. (2001) set a specific quantitative standard for representativeness in their study of the role of issue selling in organizational change; they only included a category in the findings if it was mentioned in at least 10% of their data points. Other researchers have set more vague standards. For example, Heracleous and Barrett (2001) identified their key themes by seeing if they were present in “several” of the texts that they analyzed. As Pratt (2008) noted, there is no clear set of rules on how “cutoffs” for representativeness should be established or managed.

Another way that counting can be helpful to qualitative researchers is to use calculations of intercoder reliability as evidence of the objectivity and accuracy of coding processes. Intercoder reliability reflects the degree to which different people would assess and categorize the data in a similar fashion. Miles and Huberman (1994) suggested that intercoder reliability could be assessed using the following formula: $\text{reliability} = \text{number of agreements} / (\text{total number of agreements} + \text{disagreements})$, and this method is used by many researchers.

Counting: When to Avoid or Closet

Thus, qualitative researchers may choose to count their data to develop autonomous findings, to corroborate or supplement other findings, or to develop evidence that supports the credibility of their analysis. However, sometimes none of these purposes will be sufficiently useful to qualitative researchers to induce them to count their data. Furthermore, on other occasions researchers may decide after the fact that the counts they have produced are not useful. Next, we discuss when these situations are likely to arise. We specify when researchers should consider (a) avoiding counting entirely or (b) closeting their counting.

Avoid Counting

There are at least two occasions when it may not be worthwhile for qualitative researchers to engage in any counting: (a) When it is important for researchers to gain access to the perspectives of insiders and (b) When researchers wish to pursue unexpected findings during an inductive data collection process.

Losing the perspectives of insiders. Qualitative research often involves researchers trying to understand the perspectives of their research subjects, because those individuals who are insiders to a process frequently see things in a different way than outsiders do. By gaining access to the perspectives of insiders, researchers can also gain access to new and surprising ways of seeing the world. However, the process of counting may focus researchers away from the perspectives of their subjects, because counting requires researchers to assign objective characteristics (i.e., numbers) to the issues of interest in a study. When raw qualitative data, which typically include more direct representations of subjects' thoughts or actions, are converted into a different kind of information (i.e., numbers), that conversion may actually distance a researcher from subjects' perspectives (Alvesson, 2003), in the same way that translating words from one language to another can convey literal meanings but can also miss important nuances or connotations. The result may be a set of numbers that contains less information about insiders' views than the raw data did.

Limiting researchers' ability to generate insights from serendipitous findings. In qualitative research processes, insights into theory often arise serendipitously. Barley (2006) argues that qualitative research is often interesting and generates important insights because it is inductive: “qualitative researchers often discover something because they usually approach topics with little clue as to what they'll find” (p. 19). Other researchers may begin their research with the intention of building on existing theory only to find that standard accounts don't fit their observations. They then face a choice over whether to shift tacks and explore more novel explanations. A choice to explore surprising issues can often lead to useful discoveries (Locke, Golden-Biddle, & Feldman, 2008), but it also necessarily means that researchers would be taking an approach that was different from what they had done before. However, counting typically requires researchers to be consistent in their approaches to data gathering and to data analysis. Researchers who want to be responsive to unexpected discoveries may then decide to sacrifice the consistency in data collection that facilitates counting.

To illustrate, consider Denison and Mishra's (1995) research on organizational culture and effectiveness. They described their research methodology as follows:

A specific set of questions guided the interviews, but the interviewer resisted the approach of asking the same questions each time and summarizing the interviews by counting and aggregating responses. Instead, each successive interview was used to expand understanding of the organization. For example, after the first set of five to ten interviews, clear areas of overlap and redundancy (or sometimes conflict and inconsistency) began to emerge. At this point, the results were

summarized and served as a basis for the development of a new set of questions intended to develop an understanding of the organization that was both broader and deeper. (p. 208)

The implication was that counting was not relevant to, and may even have detracted from, the inductive approach they took to their research.

Closeted Counting

If qualitative researchers choose to count their data, we maintain that there are times when they should keep those counts in the closet, that is, conceal them. There are four circumstances when counting should be closeted. The first of these is similar to those identified by Sutton in his article: counts should be closeted when audiences may be biased against numbers. Second, we argue that counting, or at least the extensive descriptions of it, should be closeted when those descriptions waste space in a manuscript. Third, counts should be closeted when supplementary or corroborative counting produces weak quantitative data. Finally, we explain why we disagree with Sutton's point that qualitative research should be closeted, as he put it, "When weak qualitative data leads to good insights" (1997, p. 99). We suggest that in that situation, it is not the qualitative data that should be closeted but the results of any counting that has taken place.

When audiences may be biased against numbers. At a recent Academy of Management meeting, we heard from a colleague who had an experience with qualitative researchers who were opposed to counting. She was developing an article that contained some new ideas about how qualitative researchers could go about counting their qualitative data. She recounted a previous presentation when she described the methodology and then put up a table that included numbers to demonstrate the outcome of the method. To her surprise, her audience of qualitative researchers booed the numbers. Her story confirmed our own impression that some qualitative researchers have a bias against counting.

An associated challenge for qualitative authors is that, given the multiple audience problem, it is difficult to know in advance when such bias will manifest itself. In Sutton's article, he wrote that three journals in his core discipline of psychology, *Journal of Applied Psychology*, *Organizational Behavior and Human Decision Processes*, and *Journal of Personality and Social Psychology*, would be likely to evaluate research more positively if it did not contain qualitative data. We are not as confident as Sutton was in identifying journals that are more or less likely to include counting, but one way in which researchers may be able to reduce uncertainty is by considering who at a journal is likely to be the chosen editor for qualitative research. Many journals will have one or two members of an editorial board who are

experienced qualitative researchers. Prospective authors could take a look at that editor's qualitative work. If he or she counts their data, they may be more open to counting than an editor who does not count their data. If that editor never counts their data, this may be a signal that authors should closet their own counting. In a recent editorial on publishing qualitative research at the *Academy of Management Journal*, Pratt (2009) similarly recommended "modeling" someone who successfully published qualitative work; our advice narrows the choice of this model to members of the editorial board at the targeted journal.² Although this would not solve the multiple audience problem, at least it would increase the odds that a crucial member of that audience would react favorably to the article's research methodology.

When descriptions of counting waste space in a manuscript. As we noted earlier, the purpose of credentialing counting is to demonstrate why one should have confidence in the findings of a qualitative analysis. In many articles, credentialing counting includes extensive counts of data sources (e.g., Corley & Gioia, 2004; Reay et al., 2006; Suddaby & Greenwood, 2005). According to Pratt (2008), the majority of recent award-winning qualitative articles include tables and figures as part of their methods description (and some include other forms of counting as well). It appears that credentialing counting may be emerging as part of a developing standard for qualitative research.

One might argue that this is a good thing, because it provides a partial solution to the lack of clear guidelines for the presentation of qualitative research. However, it could also be argued that sometimes authors devote too much manuscript space to descriptions of credentialing counting. Since qualitative researchers often struggle to comply with journal length requirements (Pratt, 2008), one might question whether the pages spent presenting the results of credentialing counting could be put to better use by either shortening a manuscript or by discussing more interesting aspects of the research.

One argument against devoting much research time or manuscript space to presentations of credentialing counting is that such presentations often do not communicate much about the actual quality of the method or associated insights. Some of the inevitable, and perhaps wonderful, messiness of qualitative research is that there may not be predetermined cutoffs of the simple *amount* of data that is sufficient to generate interesting theory. Sutton (1997) argued that the quality of the insights produced from qualitative research is frequently independent of the amount of data that produced them. Barley (1990) noted that when the volume of field notes is routinely cited in ethnographic studies as "evidence of sustained observation and as a not-so-subtle indicator that real work was done" (p. 228); this is not sufficient to indicate research quality or guarantee useful insights. Similarly, Easterby-Smith et al. (2008) argued that researchers should emphasize "proximity to the life worlds of those studied" (p. 423) rather than

quantity of data collected in their methods descriptions. In sum, while credentialing counting on a modest scale may be useful, overly extensive counting of this type should generally be closeted or avoided as it is minimally useful and is likely to displace other material.³

When counting produces weak numbers. In order for counting to produce useful findings, the numbers produced must be valid: they need to be both specific and meaningful (Gherardi & Turner, 1987). Numbers should be specific in the sense that the differences between units are clear, and meaningful in the sense that a difference is an indicator of something. Gherardi and Turner (1987) cite Suttles's (1978) study of Chicago slums as providing an example of valid numbers. Suttles assessed the degree of conflict in a particular area by counting the number of gang fights per month. His count was composed of specific units (the number corresponded to the number of fights) and also meaningful (more fights were an indicator of more conflict between social groups).

The study conducted by Ashforth, Kreiner, Clark, and Fugate (2007) illustrates how counts in an exploratory study can produce misleading numbers. The authors used an exploratory qualitative approach to document the diverse ways in which managers attempt to normalize dirty work. As part of their data analysis, they developed counts of the "number of times interviewees mentioned a particular phenomenon" (p. 155). They also acknowledged that these counts were flawed due to their exploratory approach. The potential flaws related to two ways in which the data collection process varied across respondents in this study. First, themes that were ultimately counted emerged spontaneously in interviews rather than in response to specific questions and as a result were not considered or raised by all interviewees. Second, respondents answered different sets of questions since interviewers "were free to explore interesting themes in more detail and were not required to ask every question in the protocol" (p. 154). Both issues meant that counts developed "likely under represent the phenomenon" (p. 155). In other words, the counts of the number of times the themes emerged were not meaningful, at least to the extent that they did not represent the actual prevalence or influence of the themes.

While Ashforth et al. (2007) chose not to include numbers in the presentation of their data due to concerns about their validity, in other cases authors have chosen to include relatively weak numbers in a presentation of corroborative or supplementary counting. Consider the aforementioned study by Maguire et al. (2004), where they counted the legitimizing characteristics of actors to corroborate their conclusions about who the key actors were. Their chosen actors had 10 and 8 characteristics, respectively, and the next two actors both had six characteristics. Although these units were specific (a score of 1 was assigned for each characteristic), it was not clear to what degree an increase in units corresponded to an increase in legitimacy. It could be said that those numbers

on their own do not present compelling evidence regarding the differences in legitimacy between the actors.

Our contention is that if authors conduct supplementary or corroborative counting, and generate only weak numbers, they should closet that counting. Quantitative data, whether it is in a qualitative paper or in a quantitative one, should be held to a high standard. If this standard is not met, the paper should stand (or fall) on the merits of its qualitative data alone. In our opinion, if the weak counting was removed from the Maguire et al. (2004) article, there remains more than sufficient quality in the data, research methods, and insights to justify its publication in the *Academy of Management Journal*.

We suspect that this kind of weak counting occurs in two kinds of cases. The first is when quantitatively trained researchers attempt qualitative research for the first time. Such researchers are trained to work with quantitative data, so they generate such data and work to produce findings. On other occasions, qualitative researchers do not embrace counting (at least for a particular project), but their editorial audience sees merit in it. For both cases, we would repeat that if counting is conducted to corroborate or supplement qualitative findings, it should still be held to a high standard. (Of course, if researchers plan to conduct autonomous counting and produce weak numbers, their research is probably fatally flawed). We would further say that authors, reviewers, and editors must be aware of the plurality of approaches possible to qualitative research (Easterby-Smith et al., 2008), and recognize that counting is not a requirement of great insight.

When weak qualitative data leads to good insights. Sutton argued that due to the prevalence of norms for systematic qualitative research, "... insights are more convincing to most of us when they are grounded in lots of qualitative data" (1997, p. 100). He suggested that when "weak" qualitative data—by which he means small amounts of qualitative data—lead to good insights, researchers should consider closeting their qualitative data. He acknowledged that some might argue that this practice is ethically questionable, but counters that it is the lesser of two evils, because hiding the origins of the insights "... averts unfair bias that will result in the rejection of sound ideas" (p. 101).

We disagree with Sutton's arguments about closeting qualitative data under these circumstances. First of all, we question whether qualitative data that generates good insights should be called "weak" data. Even if there is not much of it, one could argue that data that help a researcher develop interesting, provocative new ideas should not be called weak. Second, in many cases the insights stimulated by small amounts of data would probably not have been realized without some background work by researchers. Qualitative scholars have emphasized the importance of theoretical sensitivity (Strauss & Corbin, 1990), which is a personal quality of researchers that allows them to be aware of meanings in the data. Researchers can develop theoretical sensitivity by

being aware of relevant literature and by having direct experience with a phenomenon of interest. One could argue that researchers are much more likely to generate strong insights from weak (or strong) qualitative data if they have done the work necessary to be sensitive to those insights. Thus, in many cases the so-called “weak” data that stimulated strong insights may often be only small part of a broader set of personal experiences and relevant literature, all of which contributed to those insights.

We suggest that when small amounts of qualitative data have generated strong insights, researchers should present that data, along with a description of how they acquired the theoretical sensitivity necessary to be open to those insights. We further suggest that they closet any counting they have conducted. **As Pratt (2008) noted, one of the problems with presenting numbers is that it can evoke a positivist orientation among editors or reviewers who may react to the inclusion of numbers by (a) demanding even more numbers or (b) enacting and applying quantitative standards of evaluation to the research. This can result not only in more pressure to count, but also in the enactment of the quantitatively oriented idea that small amounts of data constitute weak data. Put another way, by presenting numbers, qualitative researchers risk undermining the legitimacy of any insights they derived from small amounts of qualitative data.**

Discussion and Conclusion

We began this article by noting that scholars do not agree on how to assess the quality of qualitative research. We hope that the ideas we have put forward can help us reach a better understanding, and therefore some consensus, over when qualitative researchers should consider counting their data, when they should avoid counting, and when they should closet their counting. At the very least, we hope that we have initiated a discussion that will move us toward that desired end.

To conclude this article, we wish to discuss three issues. First, we consider whether closeted counting is already happening in the field. Second, we discuss the implications of our ideas for triangulation. Finally, we reflect on the importance of being purposeful when making decisions about counting.

Is Closeted Counting Already Happening?

Although we cannot be sure without talking to the authors involved, we have identified many examples of what appears to us to be closeted counting in qualitative research. This in turn suggests that authors have already recognized its benefits. For example, Heracleous and Barrett (2001) chose their key themes by seeing whether they were present in “several” of the texts that they analyzed. If the authors deliberately avoided the inclusion of a number in their manuscript, this is an example of closeted counting. In some cases, closeted

counting can be somewhat more open: Maitlis (2005, p. 29) for example, reported that she

determined the overall level of stakeholder sensegiving in two ways. The first was by counting the number of stakeholders involved in sensegiving for that process. The second was through a more impressionistic “gestalt analysis” of the frequency and intensity with which stakeholders engaged in the above sensegiving activities, which provided a richer, more holistic assessment than number of stakeholders alone would have done.

In this case, counts were done and acknowledged but not reported. This form of closeted counting poses risks—reviewers alerted to the existence of counts may reasonably request to see them. More complete closeting, in which authors hide entirely that counting has taken place, avoids this problem.

Whither (or Wither) Triangulation?

We do not take issue with the inherent logic of triangulation, but we suggest that it does not make sense to corroborate strong, interesting, well-substantiated qualitative findings with weak counting. One way of interpreting this argument would be to view it as a call for stronger triangulation. This is not our intent. It is difficult enough to conduct excellent qualitative research without having to add an entire quantitative study. We are suggesting that qualitative researchers and those who read and evaluate it should ask themselves the following question before considering triangulation: are the findings convincing and interesting as they stand now? If the answer is “yes,” then why bother conducting additional corroborative or supplementary counting, or presenting page upon page of credentialing counting?

Counting Should be Approached Purposefully

Finally, we think it is worth noting that counting in and of itself does not require any great abilities on the part of a researcher. According to recent research, numerous animals have rudimentary counting abilities, including lions, apes, dolphins, cats, robins, and even fish (Agrillo, Dadda, Serena, & Bisazza, 2009). Thus, there is no reason to be any more impressed by qualitative research just because it includes counting (at least when that research has been conducted by human beings).

What is more important to the quality of a study is whether the researchers have considered when, how, or why counting is likely to help them move from their raw data to an interesting and convincing set of findings. To make that determination, researchers must approach counting purposefully. Researchers should consider the overall purpose of their study

and first evaluate whether they should avoid counting entirely. If counting has the potential to be useful to them, they should consider whether to conduct autonomous, supplementary, corroborative, or credentialing counting. Once that determination has been made and the counting has been done, researchers conducting supplementary, corroborative, or credentialing counting must decide whether they wish to present the results of that counting, or keep some or all of those results closeted.

In a similar vein, we suggest that when evaluating a manuscript, reviewers and editors should consider whether the authors have approached the decision about whether or not to count their data in a purposeful manner. If, for example, authors have approached their research from a positivist perspective and are interested in identifying patterns in an entire data set, then they would be justified in choosing to use autonomous counting. If authors have taken an interpretive approach to their work, have the stated goal of developing novel theory and do so successfully, and have decided to avoid counting, then that decision should be respected.

In conclusion, we hope that our ideas have provided some much-needed guidance for researchers. We have identified different purposes of counting; suggested when counting should be avoided entirely; and discussed when the results of counting should be closeted. While our arguments may not apply to every possible variation of counting, we hope that others will decide to take the ideas we have presented here and argue with them, modify them, and build on them, so that those who conduct and review qualitative research will have a better understanding of when scholars should and should not use counting as a means of analyzing qualitative data.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Funding

The Social Sciences and Humanities Research Council of Canada provided funding for this research.

Notes

1. Interpretivist researchers may combine qualitative analysis with counting where the intent is to explore distinctions between espoused and manifest schema. We are indebted to an anonymous reviewer for this insight.
2. We do not mean to imply that editors will be tricked into giving undeserved positive evaluations of papers that resemble theirs. Our experience has been that qualitative editors preach what they practice: they encourage authors to use the same kinds of methodological approaches that they use in their own research, probably because they believe those approaches lead to high-quality research.
3. Another option would be for credentialing information to be included in an appendix that editors and reviewers can use during

the review process but ultimately cut before publication to conserve space. In this case, the whole editorial team participates in the closeting of the counted data. We are thankful to an anonymous reviewer for this suggestion.

References

- Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2009). Use of number by fish. *PLoS ONE*, 4, e4786. doi:10.1371/journal.pone.0004786
- Alvesson, M. (2003). Beyond neopositivists, romantics, and localists: A reflexive approach to interviews in organizational research. *Academy of Management Review*, 28, 13-33.
- Amis, J. M., & Silk, M. L. (2008). The philosophy and politics of quality in qualitative organizational research. *Organizational Research Methods*, 11, 456-480.
- Ashforth, B. (2005). Becoming vanilla pudding: How we undermine our passion for research. *Journal of Management Inquiry*, 14, 400-403.
- Ashforth, B., Kreiner, G. E., Clark, M. A., & Fugate, M. (2007). Normalizing dirty work: Managerial tactics for countering occupational taint. *Academy of Management Journal*, 50, 149-174.
- Barley, S. R. (1986). Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments. *Administrative Science Quarterly*, 31, 78-108.
- Barley, S. R. (1990). Images of imaging: Notes on doing longitudinal field work. *Organization Science*, 1, 220-247.
- Barley, S. R. (2006). When I write my masterpiece: Thoughts on what makes a paper interesting. *Academy of Management Journal*, 49, 16-20.
- Bartunek, J. M., & Seo, M. (2002). Qualitative research can add new meanings to quantitative research. *Journal of Organizational Behavior*, 23, 237-242.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally College Publishing Company.
- Corley, K. G., & Gioia, D. A. (2004). Identity ambiguity and change in the wake of a corporate spin-off. *Administrative Science Quarterly*, 49, 173-208.
- Denison, D. R., & Mishra, A. K. (1995). Toward a theory of organizational culture and effectiveness. *Organization Science*, 6, 204-223.
- Dutton, J. E., Ashford, S. J., O'Neill, R. M., & Lawrence, K. A. (2001). Moves that matter: Issue selling and organizational change. *Academy of Management Journal*, 44, 716-736.
- Easterby-Smith, M., Golden-Biddle, K., & Locke, K. (2008). Working with pluralism: Determining quality in qualitative research. *Organizational Research Methods*, 11, 419-429.
- Fineman, S., & Mangham, I. (1983). Data, meanings and creativity: A preface. *Journal of Management Studies*, 20, 295-300.
- Gephart, R. P., Jr. (2004). From the editors: Qualitative research and the Academy of Management Journal. *Academy of Management Journal*, 47, 454-462.

- Gherardi, S., & Turner, B. (1987). Real Men Don't Collect Soft Data. Quaderno 13, Dipartimento di Politica Sociale, Università di Trento.
- Golden-Biddle, K., & Locke, K. (1993). Appealing work: An investigation of how ethnographic tests convince. *Organization Science*, 4, 595-616.
- Heracleous, L., & Barrett, M. (2001). Organizational change as discourse: Communicative actions and deep structures in the context of information technology implementation. *Academy of Management Journal*, 44, 755-778.
- Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, 24, 602-611.
- Lee, T. W. (1999). *Using qualitative methods in organizational research*. Thousand Oaks, CA: SAGE.
- Locke, K., Golden-Biddle, K., & Feldman, M. S. (2008). Making doubt generative: Rethinking the role of doubt in the research process. *Organization Science*, 19, 907-981.
- Maguire, S., Hardy, C., & Lawrence, T. B. (2004). Institutional entrepreneurship in emerging fields: HIV/AIDS treatment advocacy in Canada. *Academy of Management Journal*, 47, 657-679.
- Maitlis, S. (2005). The social processes of organizational sensemaking. *Academy of Management Journal*, 48, 21-49.
- Marshall, C., & Rossman, G. B. (1995). *Designing qualitative research* (2nd ed.). Thousand Oaks, CA: SAGE.
- Miles, M. B. (1979). Qualitative data as an attractive nuisance: The problem of analysis. *Administrative Science Quarterly*, 24, 590-601.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: SAGE.
- Pfeffer, J. (1995). Mortality, reproducibility, and the persistence of styles of theory. *Organization Science*, 6, 681-686.
- Pratt, M. G. (2008). Fitting oval pegs into round holes: Tensions in evaluating and publishing qualitative research in top North American Journals. *Organizational Research Methods*, 11, 481-509.
- Pratt, M. G. (2009). For the lack of a boilerplate: Tips on writing up (and reviewing) qualitative research. *Academy of Management Journal*, 52, 856-862.
- Reay, T., Golden-Biddle, K., & Germann, K. (2006). Legitimizing a new role: Small wins and microprocesses of change. *Academy of Management Journal*, 49, 977-998.
- Silverman, D. (2000). *Doing qualitative research*. London: SAGE.
- Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: SAGE.
- Suddaby, R. (2006). From the editors: What grounded theory is not. *Academy of Management Journal*, 49, 633-642.
- Suddaby, R., & Greenwood, R. (2005). Rhetorical strategies of legitimacy. *Administrative Science Quarterly*, 50, 35-67.
- Suttles, G. D. (1978). *The social order of the slum*. Chicago: University of Chicago Press.
- Sutton, R. I. (1997). The virtues of closet qualitative research. *Organization Science*, 8, 97-106.
- Symon, G., Buehring, A., Johnson, P., & Cassell, C. (2008). Positioning qualitative research as resistance to the institutionalization of the academic labour process. *Organization Studies*, 29, 1315-1336.

Bios

David R. Hannah is an associate professor of management and organization studies at Simon Fraser University. He received his PhD from the University of Texas at Austin. In addition to qualitative research methods, his research interests include the protection of trade secrets and confidential information, media portrayals of leaders, and human-animal interactions in organizations. His work has been published in a number of academic and practitioner journals, including *Organization Science*, *Journal of Management Studies*, *Sloan Management Review*, *Journal of Business Ethics*, and *Journal of Management Education*. He was selected as a 2008 Ascendant Scholar by the Western Academy of Management.

Brenda A. Lautsch is an associate professor in the Segal Graduate School of Business at Simon Fraser University. She received her PhD in industrial relations and human resource management from the MIT Sloan School of Management. Her current research focuses on new work systems and inequality, knowledge management, contingent work, qualitative research methodology, and on the effects of flexible work arrangements, such as telecommuting, on work-life outcomes. Her articles have appeared in prestigious journals such as the *Industrial and Labour Relations Review*, *Industrial Relations*, and *Journal of Vocational Behavior*, and she is coauthor of *CEO of Me: Creating A Life that Works in the Flexible Job Age* (2008, Wharton School Publishing).