

Published in final edited form as:

Qual Res. 2010 June 1; 10(3): 357–376. doi:10.1177/1468794110362874.

Benefits of “Observer Effects”: Lessons from the Field

Torin Monahan, Ph.D.¹ and Jill A. Fisher, Ph.D.²

¹Dept. of Human & Organizational Development, Vanderbilt University, Peabody #90, 230 Appleton Place, Nashville, TN 37203-5721, phone: 1 (615) 322-8732, torin.monahan@vanderbilt.edu ²Center for Biomedical Ethics and Society, Vanderbilt University, 2525 West End Ave., Suite 400, Nashville, TN 37203, phone: 615-936-2609, fax: 615-936-3800, jill.fisher@vanderbilt.edu

Abstract

This paper responds to the criticism that “observer effects” in ethnographic research necessarily bias and therefore invalidate research findings. Instead of aspiring to distance and detachment, some of the greatest strengths of ethnographic research lie in cultivating close ties with others and collaboratively shaping discourses and practices in the field. Informants’ performances – however staged for or influenced by the observer – often reveal profound truths about social and/or cultural phenomena. To make this case, first we mobilize methodological insights from the field of science studies to illustrate the contingency and partiality of all knowledge and to challenge the notion that ethnography is less objective than other research methods. Second, we draw upon our ethnographic projects to illustrate the rich data that can be obtained from “staged performances” by informants. Finally, by detailing a few examples of questionable behavior on the part of informants, we challenge the fallacy that the presence of ethnographers will cause informants to self-censor.

Keywords

ethnography; methods; observer effects; Hawthorne effect; reactivity; investigator bias; science studies; science and technology studies; staged performance

A frequent criticism of ethnographic research is that “observer effects” will somehow bias and therefore invalidate research findings (LeCompte and Goetz, 1982; Spano, 2005). Put simply, critics assert that the presence of a researcher will influence the behavior of those being studied, making it impossible for ethnographers to ever really document social phenomena in any accurate, let alone objective, way (Wilson, 1977). Implicit in this negative evaluation of ethnographic methods is the assumption that other methods, particularly quantitative methods, are more objective or less prone to bias (Agar, 1980; Forsythe, 1999). This paper is an initial response to that criticism.

Observer effects – also sometimes referred to as “researcher effects,” “reactivity,” or the “Hawthorne effect” – are often understood to be so pervasive that ethnographers must make de facto explanations about how they will attempt to minimize them (McDonald, 2005; Shipman, 1997). By doing so, however, ethnographers effectively legitimize the concern. For instance, a key part of grant proposals is a description of the methods that ethnographers will mobilize to prevent their presence from becoming an intervention or changing the

behaviors and activities of those whom they are studying (Agar, 1980).ⁱ Of course, the implication is that individuals will behave better (e.g., more ethically, more conscientiously, more efficiently) when being observed. In part this concern is a response to ethnographers' relative methodological indifference – unlike researchers using statistical methods – to measuring the extent of any bias introduced or calculating the reliability and validity of their data (Atkinson and Hammersley, 2007).ⁱⁱ Importantly, observer effects are framed as inevitably bad because they indicate a “contamination” of the supposedly pure social environment being studied (Hunt, 1985). Some methodologists advise qualitative researchers to hone an awareness of possible observer effects, document them, and incorporate them as caveats into reports on fieldwork (Patton, 2002). Others encourage ethnographers to seek out explicitly evidence of observer effects to better understand – and then mitigate – “researcher-induced distortions” (e.g., LeCompte and Goetz, 1982; Spano, 2006). The possibility that the ethnographer can both have an effect and *by doing so* tap into valuable and accurate data is seldom explored in contemporary literature on methods (e.g., Speer and Hutchby, 2003).

This paper proposes that the capacity of ethnography to shape the discourses and practices being observed can be considered a strength of the method. While outsiders may see the data as “biased,” ethnographers should be prepared to argue that informants' performances – however staged for or influenced by the observer – often reveal profound truths about social and/or cultural phenomena. To make this case, first we mobilize methodological insights from science studies to illustrate the contingency and partiality of all knowledge, from physics to ethnography, and to challenge the notion that ethnography is inherently less objective than other research methods. This section underscores the points that “observer effects” (broadly conceived) are unavoidable in knowledge production regardless of the scientific field and that it is the different value that is placed on scientific disciplines that creates hierarchies – as well as criticisms – of particular methods. Second, we draw upon our ethnographic projects to illustrate the rich data that can be obtained from “staged performances” by informants.ⁱⁱⁱ In addition, by detailing a few examples of questionable behavior on the part of informants, we challenge the fallacy that the presence of ethnographers will cause informants to self-censor. We are neither suggesting that ethnographers should strive to produce observer effects nor are we arguing that anything goes in qualitative research. Instead, our own experiences in the field convince us that observer effects are not in and of themselves liabilities to ethnographic inquiry. Instead, observer effects can and do generate important data and critical insights.

Methodological Insights from Science Studies

Scholars working in the sociology of knowledge, science and technology studies (STS), and related fields know that knowledge is not simply made up of pre-existing, exogenous Truths about the natural or social worlds, but is instead produced through the interrelation of competing interests and values, organizational configurations, and material properties and constraints (Latour, 1986; Hess, 1995). In short, knowledge is socially constructed. As feminist scientist Ruth Hubbard (1988) explains, “Facts aren't just out there. Every fact has a factor, a maker... making facts is a social enterprise” (5). A crucial part of the scientific

ⁱBoth authors have received comments to this effect from grant reviewers on different projects submitted to the U.S. National Science Foundation (NSF) and National Institutes of Health (NIH). In these instances, reviewers who self-identified as quantitative researchers expressed concern that ethnographic observation cannot help but act as a corrupting form of intervention.

ⁱⁱThis is not to say that there are no methods for measuring validity in qualitative research, rather these techniques are not as common as those used in quantitative research. For a discussion on formal approaches to validity in qualitative research, see Cho & Trent (2006).

ⁱⁱⁱEach project discussed mobilized qualitative methods that included ethnographic approaches to research. Detailed information about the specific methods for each of these projects can be found in the papers or books cited in the discussions of each example below.

enterprise is the methods used in the creation of facts. These methods help to establish the rules that researchers follow within a field and enable distinctions between researchers in different disciplines. Not only do methods provide a blueprint for how to do science, but they also shape the types of questions researchers can pose and the types of facts that are legible within each discipline (Kaplan and Rogers, 1994).

All knowledge and methods for generating facts are not seen as equal. Professional research communities, government agencies, and private industry attach different values to the contribution of each type of science or knowledge, and this influences societal understandings of valuable and less valuable research. The rank ordering of disciplines from the “hardest” natural science (i.e., physics) to the “softest” social science (i.e., anthropology) is indicative of the hierarchy imposed on fields and methods, with quantitative approaches ruling over qualitative ones (Porter, 1995). Hierarchies are not merely symbolic; in addition to differences in prestige bestowed on each field, practitioners are rewarded salaries and resources concomitant with each discipline’s ranking.^{iv} Moreover, knowledge production in the humanities, especially literature and the arts, is often not even acknowledged as such by those in both the natural and social sciences.

In spite of popular conceptions to the contrary, laboratory and quantitative methods are not more objective or less biased than field and qualitative methods. Mathematics and physics are no less socially constructed than anthropology and sociology (Restivo, 1985; Restivo and Bauchspies, 2006). All knowledge is contingent on the interests of the scientists creating it, the tools and procedures they use to measure the phenomena under investigation, and the analytic frameworks they use to interpret their results (Knorr-Cetina, 1999; Latour, 1986). On one hand, this means that knowledge is always circumscribed by the ability to represent the natural or social worlds. On the other, it means that the beliefs and expectations of researchers or scientific communities have powerful effects on the ability to measure and interpret the world.

An interesting example of the latter can be found in the history of science. Before 19th century science established the basis for current understanding of the mechanics of reproduction, there were many models to explain the differentiation of the sexes (Tuana, 1988; Beldecos et al., 1988). One of the models established by Aristotle was that males have more “heat” than do females, and it is the presence of heat that accounts for anatomical differences (Tuana, 1988). The question of why one sex would have more heat than the other was “solved” in the 2nd century when Galen traced the differences to the anatomical structure of veins and arteries attributing heat to the flow of “pure” blood through the arteries on the right side of the body and the lack of heat to “impure” blood in veins on the left side of the body. Seed that developed on the right side (in the right testicle and right ovary for instance) would by consequence lead to male offspring, and seed that developed on the left side would become female offspring. This was an inventive anatomical explanation to support a widespread belief about sex differences. What is remarkable about Galen’s claim is that it continued to be perpetuated for hundreds of years (until the 17th c.) in spite of anatomical evidence to the contrary. As Nancy Tuana (1988) explains:

anatomists persistently held to the view that the female seed was defective because of the impurity of the blood that fed it. Although careful attention to the actual

^{iv}John Jackson Jr. (2008) reflected on the hierarchies created to rank the social sciences:

There is a general pecking order in the social sciences. We all know that. It moves from economics down through the likes of political science and psychology, finally landing in the realm of sociology and anthropology. The closer one gets to serious mathematics as constitutive [*sic*] of the center of the discipline’s exploits, the higher one’s salary, the less diverse one’s colleagues (in terms of categories such as race or gender), and the more powerful one’s academic department. There are exceptions to this formulation, but it holds true quite a bit of the time, no?

structure of the veins and arteries of the testicles and ovaries would refute this view, anatomists continued to overlook this error... It is perhaps not surprising that even an anatomist as careful as Vesalius would perpetuate such an error. The scientific theory he had inherited demanded this “fact.” The belief that female seed arose from the “serous, salty, and acrid” blood of the left testes was the only viable explanation of the perceived differences between women and men. (49)

In other words, explanatory theories shape the perceptions of even the most careful scientists (Denzin, 1989). Observation is inflected with the values and beliefs of the observer in the laboratory and in the field whether the object of inquiry is a molecule, physiology, or human behavior.

This perspective is not unknown in the sciences. As Donna Haraway (1988) points out, “The only people who end up actually *believing* and, goddess forbid, acting on the ideological doctrines of disembodied scientific objectivity enshrined in elementary textbooks and technoscience booster literature are nonscientists” (576, original emphasis). The field of physics bears out Haraway’s point. Specifically, Heisenberg’s uncertainty principle acknowledges the impossibility of simultaneously observing, measuring, and representing nature (Lukacs, 2001). The uncertainty principle states that in an effort to gain precision about individual parts of a system, researchers lose precision in their knowledge of the whole system. For example, researchers can measure the momentum of a particle or its position, but both variables cannot be measured at once. The observer effect is the recognition that researchers are interacting with the system, usually through the instruments of measurement, and changing the phenomena being studied. In spite of these caveats, however, the field of physics is not generally thought of as biased. Nor should it be. The point here is that observer effects in physics share obvious similarities to those in ethnography, but ethnographers struggle more against *outsider claims* that the research is less valid, reliable, generalizable, and so forth than do other scientists.^v

If all pursuits of knowledge invariably have observer effects of some sort, how then can unbiased, objective science ever be achieved? In its purest sense, it cannot. However, this is not the same thing as saying that science is doomed to be biased and subjective. Instead it means that there are limits to scientific representations of Truth. What *all* natural, behavioral, and social sciences do exceedingly well is represent (multiple) truths. Feminist philosophers of science have for decades been critiquing hegemonic conceptions of Truth and advocating for more nuanced theoretical approaches to knowledge production (e.g., Haraway, 1991, 1997; Harding, 1998, 1991; Longino, 1990). A primary component of this orientation to science is that all perspectives are partial.^{vi} Because knowledge is created through combinations of disciplinary, methodological, and theoretical approaches, there are limitations to claims that can be made about the natural or social phenomena under investigation. As Sandra Harding (1991) proposes in her model of “strong objectivity,” knowledge production needs to be inclusive of more and different types of people with differing values and interests to create a more robust set of truth claims.

Just as measuring devices used by physicists provide useful ways of getting at particular truths about the world, so too does ethnography serve as an ideal vehicle for exploring truths about meaning-making practices, social relations, and power (Hess, 2001; Juris, 2008; Wall,

^vIt is common for quantitative researchers to criticize the “anecdotal” nature of qualitative research, yet quite uncommon for qualitative research to deconstruct statistical methods. This likely stems from the perception that qualitative methods – “just talking to people” – is easier and requires less training than do quantitative methods (Forsythe, 1999: 131).

^{vi}Building upon Haraway and others in science studies, Casper Jensen and Peter Lauritsen (2005) make the point that “this partiality is not in itself a problem; indeed, our argument is that it is a *condition* of research. More problematic is the forgetfulness of this partiality” (68, original emphasis).

2009). Or more relevant to debates over quantitative and qualitative methods, the questions that can be answered with statistics are different from those to be answered with ethnography, but the limitations of each, not just of ethnography, need to be acknowledged. The entrenchment of hierarchies in evaluating methods is in part a result of professional norms and political and economic motivations that both influence which types of questions are considered most important to society and constrain possible answers to those questions (Woodhouse et al., 2002). Unfortunately, so-called disinterested research that does not disrupt the status quo often supports and reproduces modern social problems, from weapons proliferation to environmental degradation (Monahan, 2008; Restivo, 1988). Building upon this problematization of the possibility of non-biased research, the following sections draw upon empirical examples to respond more directly to the criticism of ethnographic methods as compromised by observer effects.

Benefits of Informants' "Staged Performances"

Notwithstanding the inherent biases with all methods, a persistent critique of ethnography is that researchers can never obtain access to accurate representations of social phenomena (Wilson, 1977). The very act of mediation, of messy interactions with informants, of partial exposure to group practices somehow contaminates pure data that is always out there, just out of reach. A typical response of ethnographers, especially those operating in the anthropological tradition, is that with sufficient time, informants will become inured to the presence of the researcher, let down their guard, and behave "normally" (Geertz, 1973; Stoddart, 1986). According to this defense, which we find mostly persuasive, it is simply too difficult for informants to maintain a façade for researchers for months or years at a stretch. Indeed, with time researchers become integrated into the communities they study so that no façade is necessary because the barriers between researchers and informants become less important than the social practices in which individuals and groups are engaged. This is not the same thing as saying that cultural differences are erased or power differentials neutralized. Instead, the ethnographer comes to realize that his or her project is – and always was – subordinate to the relations, functions, and logics of the community being studied, and that communities find places for researchers and assign meaning to their activities (Emerson, Fretz, and Shaw, 1995). Put simply, informants have agency and will exercise it to make sense of and influence researchers and research results. The responses of communities to researchers are important data in and of themselves, revealing a great deal about the communities being studied.

From a positivist perspective, the very fact of informant responses to or engagement with ethnographers taints the data (Denzin and Lincoln, 2000; Burawoy, 1991). Thus, some researchers are haunted by the rich ethnographic material they have obtained and are plagued with existential misgivings that informants were simply staging "false" performances for the researcher's benefit. For instance, Charles Bosk (2001), reflecting on his research in U.S. medical settings, confesses:

I have been blessed and cursed with "theatrical" natives with more than enough savvy to figure out my interests. Those disciplinings that modeled the surgical conscience... and those tortured, quasi-religious meditations on the ethics of genetics and its applications... – how genuine were they really? Might they have not taken place if I had been absent and subjects felt no obligation to show just how seriously they took those obligations in which I was most interested? I did not much entertain these doubts at the time of the fieldwork. The data were often too good to question in this way. But now I wonder if I need to reconsider how an attentive audience of one with a tape recorder or stenographer's notebook, taking down every word, might cause a self-conscious subject to play his or her social role "over the top." (204)

Bosk's implication is that if performances are staged or scripted in response to the presence of an ethnographer, then the data are suspect and should probably be discarded. This is a troubling conclusion for us for a few reasons. First, it presumes that "natives" are fixed in their social settings and that their beliefs are held constant, unless – and until – they are compelled to perform for the visiting researcher.^{vii} This underplays and therefore undervalues the *constant* reconstruction of cultural meanings and group identities that occurs through engagement with a steady stream of outsiders and insiders, policies and practices, technologies and symbols, and so on.

Meaning is not out there to be found by the researcher; it is continuously made and remade through social practice and the give-and-take of social interaction, including interaction with the researcher. As Robert Emerson, Rachel Fretz, and Linda Shaw (1995) explain:

The task of the ethnographer is not to determine "the truth" but to reveal the multiple truths apparent in others' lives... Relationships between the field researcher and people in the setting do not so much disrupt or alter ongoing patterns of social interaction as reveal the terms and bases on which people form social ties in the first place... Through participation, the field researcher sees first-hand and up close how people grapple with uncertainty and confusion, how meanings emerge through talk and collective action, how understandings and interpretations change over time. In all these ways, the fieldworker's closeness to others' daily lives and activities heightens sensitivity to social life as process. (3-4)

Therefore, some of the greatest strengths of ethnographic research lie in cultivating close ties with others and dispelling the illusion that robust data are best achieved through distance. As Casper Jensen and Peter Lauritsen (2005) elucidate, "The traditional social scientist wants to keep his strength by staying distanced... these strategies are misplaced. Arguably, the problem of the social scientist is not that his connections are too many and too strong, but that they are too few and fragile" (72).

"Staged performances" are important because they are deeply revealing of how individuals perceive themselves and would like to be perceived. Developing close relationships with informants, while not necessary for witnessing and interpreting performances, can certainly assist researchers in grasping the many intended – and sometimes contradictory – messages behind such plays. Rather than invalidate or cautiously tolerate data derived from staged performances, we embrace such data – not as a representation of any singular Truth, necessarily, but as rich symbolic texts that lend themselves to multiple interpretations and provide critical insights into the cultures being studied. It is important to remember that observations are data to be interpreted, not the "results" themselves of the study, and as such, data need to be analyzed by the ethnographer in light of the context in which they were generated.

Take for example the following description of a performance staged for one of us (TM) by transportation engineers in the U.S. This study was designed to document the surveillance capabilities of intelligent transportation systems (ITS), which include the integration of a complex array of technological systems, ranging from video cameras on streets and highways, to license-plate recognition systems, to sensors embedded in roads to detect the speed, flow, and density of automobile traffic (Monahan, 2007). The primary objective of such systems is to increase the throughput of vehicles on streets and highways, using smart technologies to rationalize traffic flow rather than building new lanes or roadways. While it

^{vii}This view of informants and this approach to research is reminiscent of Orientalism. Edward Said (1979) criticized Orientalist scholars for their tendency to create the "Other" by holding cultures as separate and incommensurable and romanticizing the presumed constancy of the cultures being studied.

is not the everyday interest of ITS engineers, these systems also enable a host of surveillance possibilities, from monitoring individuals and their movements, to collecting aggregate data on commuters, to assisting the police with criminal investigations. The performance staged by engineers was intended to alleviate any concerns that the researcher might have about the surveillance functions of the systems the engineers oversaw.

Near the end of an observational session and interview with three ITS engineers in their control room, one of the engineers prompted another to demonstrate for me something they had programmed into their video surveillance system. The engineer who was addressed said “Oh yeah,” seemingly glad that his colleague reminded him, and then pulled up live video footage from an intersection. After punching a few buttons, he asked me to watch what happened when he turned the camera and zoomed-in to look into the window of a nearby apartment building. The other engineer explained:

So the camera... it's on the southeast corner and that's an apartment complex that's kind of northeast there, and as he zooms in you'll notice that whole corner, the image just goes away. It just blurs it out.

When he got close to being able to see people in the apartment, a gray splotch emerged suddenly on the screen, obscuring the view into the apartment. They demonstrated for me what they could see both with and then without the software-based privacy protection. While providing claims that the camera had the software embedded in it, they still had the capability to alter the configuration remotely, turning it on or off. The reason they gave for programming the privacy patch was that they did not want to accidentally see people in their places of residence, which occasionally happened when the system lost power and the cameras moved into a default position – pointing right at people's windows.

This was clearly a performance for my benefit rather than something that the engineers do on their own when no outsiders are present. It was intended to communicate to me that they really care about privacy and are not interested in scrutinizing people, just analyzing and optimizing traffic flow. This shows, in part, that they know that people are concerned about potential threats to privacy, which is not something I even mentioned, and that they know that their systems look like surveillance systems that could compromise individual privacy. Thus, the performance also revealed that the engineers conceive of surveillance as being about scrutinizing individuals, especially in private places, instead of seeing surveillance as being more about the scrutiny and control of people more generally, perhaps even in groups. Finally, because they have the capability of activating and deactivating the privacy protections at will, and they were not trepidatious about letting me know that, I learned that they believe that they should be trusted and that they feel capable of shouldering that responsibility. Without the performance, it would have been much more difficult to obtain these data. Still, one cannot say that these findings are corrupt because they were given freely with a purpose in mind, namely of influencing the researcher and his interpretations. As with all data, qualitative or quantitative, these data require interpretation, which is the task of the researcher, the research community, and any other readers.

A second example of a staged performance was witnessed by one of us (JF) as part of research on the clinical trials industry. The purpose of the project was to explore the implications of the pharmaceutical industry's outsourcing of drug studies to private practices and for-profit research facilities in the U.S. (Fisher, 2009). Part of the focus of the study was to better understand the process of informed consent, including how consent forms were explained and administered and how human subjects made decisions about participating in pharmaceutical clinical trials (Fisher, 2006). The staged performance occurred when a physician invited me to observe the informed consent visit of an elderly woman and her son

who were considering enrolling the woman in a clinical trial to test the safety of an investigational treatment for Alzheimer's Disease.

After the woman and her son had left the physician's office (having signed the consent form), the physician turned to me and asked me to pretend that I was a clerk from the U.S. Food and Drug Administration (FDA) and to evaluate the informed consent process. That this was the first thing he said to me following the session reveals the extent to which he was aware of my presence and likely performing, at least in part, for me. While one might question how the informed consent visit might have gone had I not been present, any observer effects that occurred do not, however, compromise the data. Because I was not representing the FDA, it did not matter how the informed consent visit went in a technical sense (which was more or less flawless); I was not there to evaluate the physician. Instead, my purpose was to understand issues pertaining to informed consent. The physician's staged performance did not diminish the importance of the insights I gleaned from the observation.

Specifically, the single most interesting finding from this interaction was the conflict between why the woman and her son had come to the clinic that day – in hopes of finding a miracle cure for the woman – and the purpose of the study – to determine if the treatment was safe enough to continue its clinical development. Scholars have discussed “therapeutic misconceptions” that accompany participation in medical research (Appelbaum and Lidz, 2008), especially safety studies which are also known as Phase I trials. This occurs when potential human subjects believe that the clinical trial will have a therapeutic benefit for their condition when in fact the study is not designed to provide a benefit (as was the case with the Alzheimer's trial) or when that the efficacy of the treatment is the very purpose of the study. There are often concerns that researchers or clinicians are not doing their part to minimize the occurrence of therapeutic misconceptions because it might be in their interest not to correct subjects' misunderstandings about the purpose of the studies. In the case of the informed consent visit I observed, the physician was very careful about explaining and reiterating that the study would not improve the woman's condition. Nonetheless, the son continued to search for some benefit to his mother if she were to participate. He settled in the end on the diagnostic benefits to the woman as a way to justify her enrollment in the study in spite of the lack of therapeutic benefit. In other words, what this interaction revealed is that even when observer effects could be said to be happening, much can still be learned from the observation.

What is important is that ethnographers interpret their data in light of the possibility that their informants are engaging in staged performances. In both of these examples, critical insights into the research questions were given. In the first case, ITS engineers demonstrated the ability of the technology to be used as a surveillance tool, their awareness of privacy concerns that are raised by this potential application of traffic cameras, and their conviction that they should be trusted. In the second case, the physician being on his best behavior could not mitigate the risk of potential subjects' to misunderstand the purpose of clinical trials, which reveals how complex and delicate the informed consent process can be. Additionally, observer effects involving the putting forth of good behavior have the potential to expose the pervasiveness and even naturalization of less than desirable behavior or interactions, which is the area we turn to next.

Refuting the Fallacy of Informant Self-Censoring

Often when U.S. funding agencies receive grant proposals for projects that will use ethnographic methods, some reviewers will complain that informants will act differently in front of researchers, that they will censor themselves and be on their best behavior, therefore attenuating the validity of the data obtained (Agar, 1980).viii Such comments, which are

intended to influence study designs and funding outcomes, embody both explicit and implicit truth claims. The explicit claims are that observer effects will occur and that every effort should be taken to minimize them; the implicit claims are that this kind of bias is unique to ethnographic research, that it is bad because it prevents access to “pure” data, and that other methods – particularly quantitative ones – are more scientific because they are not subject to these particular risks of data corruption.

Whereas we discussed the biases inherent in all research activities in the first section of this paper, the goal of this section is to question the veracity of claims about self-censorship and good behavior by research subjects. Drawing upon a few examples from our research projects, we illustrate that informants tend to show little compunction about behaving unethically or making discriminatory statements in the presence of researchers. It may be the case, still, that they are trying to be on their best behavior, but their best behavior sometimes leaves a lot to be desired and, more importantly, provides insight into their understanding of their roles in relation to others. One could argue that questionable behavior in front of researchers is a *more valid* research finding than self-reporting or other measures because it is what cannot be contained by, or what subjects choose not to hide through, self-censorship.

For example, in a collaborative ethnographic research project of ours, we witnessed a medical procedure where a physician’s failure to inform his patient – even in the presence of outside researchers observing the consent conference – may have endangered the patient. For this project, we have been studying the use of new identification and tracking technologies in hospitals (Fisher and Monahan, 2008). One of the more exotic technologies being used is a radio-frequency identification (RFID) human implant, which is inserted into the triceps region of patients’ arms. Once the chip is implanted, it can be scanned at close range by hand-held readers at hospitals to obtain a unique identifying number, which can then be used to access a patient’s health records. The ostensible value of such a system is that if a patient with an RFID implant arrives at a hospital and is unable to communicate with hospital staff, then the number communicated by the implant will give staff access to a patient’s name, medical history, allergies, and so on, so that rapid and appropriate treatment can be provided (Monahan and Wall, 2007). During a visit to one of our fieldsites, we observed and interviewed an elderly patient being implanted with one of these chips.

We were already present in the physician’s office when the patient arrived. We requested and received permission to interview the patient and physician and observe the implant procedure, and we obtained informed consent from both the physician and patient. At the same time, the physician pushed a series of “release” forms at the patient but did not explain them or tell the patient how the medical device worked or what its limitations were. The patient, who was an energetic and voluble man in his 90s, signed the release forms without reading them. When we asked why he was electing to be implanted, he told us a frightening story about collapsing in his home four months earlier and being unable to dial 911 or call for help. Someone finally did call paramedics, who came and took him to the hospital, where he eventually recovered. Medical staff at the hospital then convinced him to participate in a study on RFID implants. As the patient’s narrative unfolded, all in the presence of the physician doing the procedure, it became clear that the patient believed that the implant would direct medical assistance to him in the future, should he ever collapse again and be unable to call for help. This is not how the technology works, however. It does not read vital signs; it does not have a locator device, such as a global positioning system unit; and it does not transmit information at all without the very close presence (less than 3 feet) of an RFID reader. Thus, because the patient thought that the technology would automatically send for help in the future, he might be less likely to struggle to reach a phone

^{viii}See endnote #1 for additional information about reviews of grant proposals.

and dial 911. In short, he could be placed in greater danger with the implant than without it. The physician implanted the chip anyhow.

What is remarkable about this scene, especially for purposes of this discussion of self-censorship and best behavior on the part of those being studied, is that the physician heard from the patient everything that we did, but he chose not to correct the patient. For whatever reason, the physician was not particularly concerned about the fact that the patient had a profound misunderstanding about the capabilities of the technology or that the patient might be placed in greater physical danger later as a result. This is consistent with other evidence about “normal” misbehavior that happens in the context of medical care or research (DeVries, Anderson, and Martinson, 2006; Ziebland et al., 2007). It may be the case that self-censorship was occurring and that this was better-than-usual behavior for the physician, but one cannot say that observer effects prevented interesting and important data from being divulged.^{ix}

In another example from the clinical trials industry study described in the previous section, physicians and research staff were unguarded about their perceptions of human subjects based on gender, race, and class differences. Participation in pharmaceutical clinical trials is stratified with the poor and racial minorities enrolling in the riskiest drug studies and middle-class and whites enrolling in studies that have the most potential for personal benefit (Fisher, 2009). Given these inequalities, research staff rationalize the distribution of risk and burden by mapping it onto stereotyped characteristics of different groups. For example, poor women and women of color were often portrayed problematically as leading disorganized lives that are prone to crisis, being unreliable because of lack of transportation, and being irresponsible for having “too many” children. I (JF) heard statements from research staff such as the following: “Lots of times poor women have children attached to them, and it makes it a lot harder [to complete studies]” and

I would love to be able to go to our county hospital which is *overrun* by women who have lots of unplanned pregnancies and say, “Line up, come on over here, and let us do this for you [sterilize you].”... Maybe sometimes it’s unfair [to say, but] you have to have women who can understand *their* commitment toward the clinical trial that they’re participating in. I mean it’s not all about what *we* can do for *you*.

Likewise, differences in the participation among racial groups in clinical trials were described in equally problematic ways. Compared to white participants, research staff argue that Asians are non-altruistic, blacks are suspicious, and Hispanics are very compliant (Fisher, 2009).

This indicates that the presence of an ethnographer does not automatically trigger informants to become politically correct. Yet, based on the body language of most of my informants who talked about perceived differences among potential human subjects and the care with which they seemed to be choosing their words, I got the strong impression that they were trying to be on their best, most PC behavior. What does this mean for observer effects and how one can interpret these data? On one hand, the views of my informants can be seen as simply biased – and sometimes classist or racist – statements they made that only reveal something about the person making these claims. On the other hand, however, these perceptions of potential subjects become normalized and perpetuate enrollment practices

^{ix}The question of what is the ethical obligation of ethnographers in cases like these is clearly an important one. Perhaps it is the ethnographers’ ethical duty to intervene in order to protect their subject. In this particular case, we did not correct the patient during the consent conference. One of us (TM) did, however, attempt to talk with the patient afterwards, but the patient was ushered by the physician into another room after the procedure, and we were unable to speak with him later. This was an unfortunate and morally problematic outcome to the observation. Should we have the opportunity to re-live the experience, we would certainly have taken the opportunity we had with the patient to correct his misunderstanding.

and by doing so have real effects. For example, research staff shared with me that they saw it as their duty to persuade people not to enroll in clinical trials when the staff did not think they were “capable” of being committed, dependable study subjects. As one staff member explained that she frequently used this technique on low-income women interested in participating in pharmaceutical studies:

Usually you try to give them enough information that they [decide not to participate] themselves, you know? Maybe I’ll say to you, “Well, I don’t know these hours are not going to coincide with your work or with your lifestyle, you know?” Or “You’ve got such bad veins that I really don’t recommend you get in this study, you know?” You look for something to kind of sway the thing [decision] to go in the other direction [of not participating]... They are not going to make it [through the end of the study] because they don’t have the tools to do it. And you don’t want to tell them [that] you don’t feel that they’re mentally capable.

Thus, perceptions of research staff determine who may be allowed to enroll in clinical trials. In spite of any tendency for informants to self-censor, research staff revealed what they believe is acceptable to say about different gender, racial, and class groups to an outsider and how these beliefs shape their practices.^x

Negotiating observer effects such as the ones presented here requires interpretation on the part of ethnographers. In the example of the RFID implant, the question should not be how badly would the physician act without the ethnographers present (the answer to which might be that he would have acted the same). Instead, it presents an interesting moment in the relationship of the physician to the patient *because* the ethnographers were present. Had we not questioned the patient about why he was agreeing to get the implant, the physician may never have heard the extent to which the patient was confused about the implant’s function. Nonetheless, the interaction illustrated a revealing lack of responsibility on the part of the physician to inform the patient and protect his safety. Likewise, in the clinical trials example, the task was for the ethnographer to map the larger trends in clinical trial participation onto the attitudes of research staff. What would be an ineffective – and certainly biased – approach would be to seek a definitive answer from informants about the stratified differences in drug study participation. Instead, the goal of ethnography is to interpret and make meaning out of the relationships among groups being studied. In both examples, if what we witnessed was informants being on their best behavior, we can interpret our observations accordingly to make claims about what informants’ concerns are for how ethnographers or others might perceive them and the ways in which they may or may not be trying to make themselves look better.

Conclusion

There are politics involved whenever claims are made about the strengths or weaknesses of particular research methods. This can be seen when qualitative researchers question the practical significance of “statistically significant” findings by quantitative researchers, or when they say that quantitative research is simply unable to access or understand the local meanings that subjects attach to social phenomena. Similarly, when researchers in the quantitative tradition refer to qualitative research as “anecdotal,” “without validity,” or subject to a host of biases, such as “observer effects,” these may be legitimate concerns about the accuracy and generalizability of data, but they are also territorial efforts to discredit those working in qualitative methodological traditions (Wilson, 1977; Forsythe, 1999). Our position is that all research methods possess inherent biases, but that the

^xSimilarly, in her ethnographic study of maternity care, Bowler (1993) found that midwives in the U.K. perpetuated and mobilized racial stereotypes to shape the care they provided to women.

hegemonic value system for social science research is that of scientism – a system that aspires toward value-free, objective findings while denying that this is a biased, partial, and exclusionary form of knowledge production. Qualitative research is often falsely assumed to be less objective and more prone to bias than are quantitative approaches.

Ironically, when ethnographers take seriously the injunction to minimize observer effects, they may be restricting their access to rich data in the field. If the keys to successful ethnographic research are gaining access to fieldsites, developing rapport with informants, and obtaining insider status (Fox, 2004), then the scientific move toward distance and separation pulls ethnographers out of their element, thereby privileging modes of discrete data extraction over engaged, collaborative participant observation. Efforts by ethnographers to minimize observer effects not only lend validity to the critique that bias and data-corruption are occurring, but they also impose artificial constraints upon fieldwork and subsequently diminish the scientific endeavor.

Of course, we recognize that ethnographers are compelled to constrain their own practices in these ways to obtain funding, publish, and otherwise succeed in academia. Given the widespread nature of the biases against ethnographic methods, it would be unethical for us to advise researchers – especially junior researchers – to eschew these scientific rituals that vouchsafe acceptance of ethnography while effectively subordinating it to quantitative modes of inquiry. This is further evidence, however, that there is still a long way to go in cultivating recognition across social science fields of the unique strengths, value, and validity of ethnographic methods. It is insufficient to see ethnography as an optional precursor to truly robust quantitative research that will be informed by ethnographic findings, which is a conciliatory move in several fields. This narrative further subordinates one method to another.^{xi} Ethnography is not simply a scout that is sent into a strange and possibly hostile field to spy on natives and report back to authorities so that “real” research can follow. Even as a metaphor alone, this is a colonial and extractive approach to research that does not resonate well with ideals of engagement, participation, and collaboration that are present in the best – but unfortunately not all – ethnographic work.

Within this context, which is the context of the politics of knowledge production as well as research methods, this paper has explored some salubrious outcomes of observer effects, especially when they take the form of staged performances, and questioned the extent to which self-censorship occurs in the field. Staged performances should be warmly accepted as gifts from informants; they are valuable treasures of meaning, abundantly wrapped in multiple layers of interest, assumption, and concern; they are alluring conceits overflowing with interpretive possibility. As anthropologists know, a gift is never merely a gift (Mauss, 1990). The giving of a gift creates and solidifies relationships and obligations, enmeshing people in shared social fabric through ritual. One does not discard the gift or scowl upon it (or its givers), no matter how ugly or useless the gift might seem. The obligation for ethnographers given the gift of a staged performance is to interpret it deeply. Ethnographers should reflect upon what staged performances communicate about what informants desire, how they would like to be seen, what they hold up as ideals, what they think might be important for an outsider to know, how they perceive researchers, and more. All of these messages are present in staged performances in a much more highly concentrated form than in routine interactions and mundane daily practices. This is a gift that must be accepted and passed on to others.

^{xi}Lynne S. Giddings (2006) makes a compelling argument that the move toward “mixed methods research” also effectively reduces robust and engaged qualitative inquiry to pragmatic instrumentality, usually for purposes of securing funding for research.

Self-censorship on the part of informants is another way that observer effects are purported to manifest. Undoubtedly, self-censorship or behavior modification do occur when informants feel scrutinized. Ethnographers have argued that over time self-censorship fades away, especially if the researcher becomes taken for granted by informants and/or integrated into the community being studied. There are also ways that ethnographers can perform “validity checks” on data they suspect to be dubious, namely by comparing discourse to practice and looking for tensions, or by triangulating articulations from multiple informants and looking for inconsistencies. In employing such pragmatic validity checks, though, one should not lose sight of the fact that all data are open to interpretation and reinterpretation; it is too simplistic to think of some data as true and other data as false. Instead, there are many webs of signification that the ethnographer must navigate to construct a coherent story of cultural meanings, logics, and structures (Geertz, 1973; Clarke, 2005).

Another way of approaching the problem of self-censorship is by carefully noting instances where it does not appear to be happening. We related a few examples from our research projects where people acted against their professional code of ethics (i.e., physicians not obtaining informed consent from patients for voluntary medical procedures) or made racist, sexist, or classist statements seemingly without concern that their words would become data. In such instances, which are not at all rare, disturbing practices and articulations on the part of informants may actually be *more valid* for occurring in front of researchers. The reasoning behind this conclusion is that self-censorship may be occurring but that these unethical practices or prejudicial claims escaped that filter. In other words, informants might behave much worse when researchers are not present, but it is unlikely that they behave much better.

It is our hope that this serious discussion of observer effects will inform debates about the merits of ethnographic research more broadly. We are under no illusion that all ethnographic projects are equally robust or free from corrupting interests, but ethnographic methods should not be subordinated to other methods simply because they do not employ the same techniques, aspire to the same outcomes, or adhere to the same criteria. Additionally, staged performances and self-censorship, or lack thereof, are valuable data in and of themselves, not something to be dismissed. It is the ethnographer’s close proximity to and interaction with informants, rather than distance and separation, that affords the transformation of observer effects from distasteful bias to serendipitous boon.

Acknowledgments

This material is based upon work supported by the U.S. National Science Foundation under grant numbers 0423672, 0642797, and 0907993 and by the U.S. National Institute of Mental Health through a Kirschstein National Research Service Award, number 5F31MH070222.

References

- Agar, Michael. Getting Better Quality Stuff: Methodological Competition in an Interdisciplinary Niche. *Urban Life* 1980;9(1):34–50.
- Appelbaum, Paul S.; Charles, W Lidz. The Therapeutic Misconception. In: Emanuel, EJ.; Crouch, RA.; Grady, C.; Lie, R.; Miller, F.; Wendler, D., editors. *The Oxford Textbook of Clinical Research Ethics*. New York: Oxford University Press; 2008.
- Atkinson, Paul; Martyn, Hammersley. *Ethnography: Principles in Practice*. 3. New York: Taylor & Francis; 2007.
- Beldecos, Athena; Sarah, Balley; Scott, Gilbert; Karen, Hicks; Lori, Kenschaft; Nancy, Niemczyk; Rebecca, Rosenberg; Stephanie, Schaertel; Andrew, Wedel. The Importance of Feminist Critique for Contemporary Cell Biology. *Hypatia* 1988;3(1):61–76.

- Bosk, Charles L. Irony, Ethnography, and Informed Consent. In: Hoffmaster, B., editor. *Bioethics in Social Context*. Philadelphia: Temple University Press; 2001. p. 199-220.
- Bowler, Isobel. "They're not the same as us": Midwives' stereotypes of South Asian descent maternity patients. *Sociology of Health & Illness* 1993;15(2):157-178.
- Burawoy, Michael. Introduction. In: Burawoy, M., editor. *Ethnography Unbound: Power and Resistance in the Modern Metropolis*. Berkeley: University of California Press; 1991. p. 1-7.
- Cho, Jeasik; Allen, Trent. Validity in Qualitative Research Revisited. *Qualitative Research* 2006;6(3): 319-340.
- Clarke, Adele. *Situational Analysis: Grounded Theory after the Postmodern Turn*. Thousand Oaks, Calif: Sage Publications; 2005.
- Denzin, Norman K. *Interpretive Interactionism*. Newbury Park, CA: Sage; 1989.
- Denzin, Norman K.; Lincoln, Yvonna S. The Discipline and Practice of Qualitative Research. In: Denzin, NK.; Lincoln, YS., editors. *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage; 2000. p. 1-28.
- DeVries, Raymond; Anderson, Melissa S.; Martinson, Brian C. Normal Misbehavior: Scientists talk about the ethics of research. *Journal of Empirical Research on Human Research Ethics* 2006;1(1): 43-50. [PubMed: 16810336]
- Emerson, Robert M.; Fretz, Rachel I.; Shaw, Linda L. *Writing Ethnographic Fieldnotes*. Chicago: University of Chicago Press; 1995.
- Fisher, Jill A. Procedural Misconceptions and Informed Consent: Insights from Empirical Research on the Clinical Trials Industry. *Kennedy Institute of Ethics Journal* 2006;16(3):251-268. [PubMed: 17091560]
- Fisher, Jill A. *Medical Research for Hire: The Political Economy of Pharmaceutical Clinical Trials*. New Brunswick, N.J: Rutgers University Press; 2009.
- Fisher, Jill A.; Torin, Monahan. Tracking the Social Dimensions of RFID Systems in Hospitals. *International Journal of Medical Informatics* 2008;77(3):176-183. [PubMed: 17544841]
- Forsythe, Diana E. "It's Just a Matter of Common Sense": Ethnography as Invisible Work. *Computer Supported Cooperative Work* 1999;8:127-145.
- Fox, Renée C. Observations and Reflections of a Perpetual Fieldworker. *The ANNALS of the American Academy of Political and Social Science* 2004;595(1):309-326.
- Geertz, Clifford. *The Interpretation of Cultures: Selected Essays*. New York: Basic Books; 1973.
- Giddings, Lynne S. Mixed-methods Research: Positivism Dressed in Drag? *Journal of Research in Nursing* 2006;11(3):195-203.
- Haraway, Donna. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies* 1988;14(3):575-599.
- Haraway, Donna J. *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge; 1991.
- Haraway, Donna J. *Modest_Witness@Second_Millennium.FemaleMan_Meets_OncoMouse: Feminism and Technoscience*. New York: Routledge; 1997.
- Harding, Sandra. *Whose Science? Whose Knowledge?: Thinking from Women's Lives*. Ithaca, NY: Cornell University Press; 1991.
- Harding, Sandra. *Is Science Multi-Cultural?: Postcolonialisms, Feminisms, and Epistemologies*. Bloomington: Indiana University Press; 1998.
- Hess, David J. *Science and Technology in a Multicultural World: The Cultural Politics of Facts and Artifacts*. New York: Columbia University Press; 1995.
- Hess, David J. Ethnography and the Development of Science and Technology Studies. In: Atkinson, P.; Coffey, A.; Delamont, S.; Lofland, J.; Lofland, L., editors. *Sage Handbook of Ethnography*. Thousand Oaks, CA: SAGE Publications; 2001. p. 234-245.
- Hubbard, Ruth. Science, Facts, and Feminism. *Hypatia* 1988;3(1):5-17.
- Hunt, Morton M. *Profiles of Social Research: The Scientific Study of Human Interactions*. New York: Russell Sage Foundation; 1985.
- Jackson, John L, Jr. Anthropology: The Softest Social Science?. *The Chronicle of Higher Education* (online), July 29. 2008. Available from <http://chronicle.com/review/brainstorm/index.php?id=680>

- Jensen, Casper; Lauritsen, Peter. Qualitative Research as Partial Connection: Bypassing the Power-Knowledge Nexus. *Qualitative Research* 2005;5(1):59–77.
- Juris, Jeffrey S. *Networking Futures: The Movements Against Corporate Globalization*. Durham: Duke University Press; 2008.
- Kaplan, Gisela; Rogers, Lesley. Race and Gender Fallacies: The Paucity of Biological Determinist Explanations of Difference. In: Tobach, E.; Rosoff, B., editors. *Challenging Racism and Sexism: Alternatives to Genetic Explanations of Difference*. New York: The Feminist Press at The City University of New York; 1994.
- Knorr-Cetina, K. *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, Mass: Harvard University Press; 1999.
- Latour, Bruno; Woolgar, Steve. *Laboratory Life: The Construction of Scientific Facts*. Princeton, NJ: Princeton University Press; 1986.
- LeCompte, Margaret D.; Preissle Goetz, Judith. Problems of Reliability and Validity in Ethnographic Research. *Review of Educational Research* 1982;52(1):31–60.
- Longino, Helen E. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton: Princeton University Press; 1990.
- Lukacs, John. Heisenberg's Recognitions: The End of the Scientific World View. In: Lederman, M.; Bartsch, I., editors. *The Gender and Science Reader*. New York: Routledge; 2001. p. 225-230.
- Mauss, Marcel. *The Gift: Forms and Functions of Exchange in Archaic Societies*. London: Routledge; 1990.
- McDonald, Seonaidh. Studying Actions in Context: A Qualitative Shadowing Method for Organizational Research. *Qualitative Research* 2005;5(4):455–473.
- Monahan, Torin. "War Rooms" of the Street: Surveillance Practices in Transportation Control Centers. *The Communication Review* 2007;10(4):367–389.
- Monahan, Torin. Editorial: Surveillance and Inequality. *Surveillance & Society* 2008;5(3):217–226.
- Monahan, Torin; Wall, Tyler. Somatic Surveillance: Corporeal Control through Information Networks. *Surveillance & Society* 2007;4(3):154–173.
- Michael Quinn, Patton. *Qualitative Research and Evaluation Methods*. 3. Thousand Oaks, CA: Sage Publications; 2002.
- Porter, Theodore M. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press; 1995.
- Restivo, Sal. *The Social Relations of Physics, Mysticism, and Mathematics: Studies in Social Structure, Interests, and Ideas*. New York: Springer; 1985.
- Restivo, Sal. Modern Science as a Social Problem. *Social Problems* 1988;35(3):206–225.
- Restivo, Sal; Bauchspies, Wenda. The Will to Mathematics: Minds, Morals, and Numbers. *Foundations of Science* 2006;11(1-2):197–215.
- Said, Edward W. *Orientalism*. 1st Vintage books. New York: Vintage Books; 1979.
- Shipman, Marten D. *The Limitations of Social Research*. 4. London: Longman; 1997.
- Spano, Richard. Potential Sources of Observer Bias in Police Observational Data. *Social Science Research* 2005;34:591–617.
- Spano, Richard. Observer Behavior as a Potential Source of Reactivity: Describing and Quantifying Observer Effects in a Large-Scale Observational Study of Police. *Sociological Methods & Research* 2006;34(4):521–553.
- Speer, Susan A.; Hutchby, Ian. From Ethics to Analytics: Aspects of Participants' Orientations to the Presence and Relevance of Recording Devices. *Sociology* 2003;37(2):315–337.
- Stoddart, Kenneth. The Presentation of Everyday Life. *Journal of Contemporary Ethnography* 1986;15(1):103–121.
- Tuana, Nancy. The Weaker Seed The Sexist Bias of Reproductive Theory. *Hypatia* 1988;3(1):35–59.
- Wall, Tyler. *Doctoral Dissertation, Justice and Social Inquiry*. Arizona State University; Tempe: 2009. *The Fronts of War: Military Geographies, Local Logics, and the Rural Hoosier Heartland*.
- Wilson, Stephen. The Use of Ethnographic Techniques in Educational Research. *Review of Educational Research* 1977;47(2):245–265.

- Woodhouse, Edward; David, Hess; Steve, Breyman; Martin, Brian. Science Studies and Activism: Possibilities and Problems for Reconstructivist Agendas. *Social Studies of Science* 2002;32(2): 297–319.
- Ziebland S, Featherstone K, Snowdon C, Barker K, Frost H, Fairbank J. Does it matter if clinicians recruiting for a trial don't understand what the trial is really about? Qualitative study of surgeons' experiences of participation in a pragmatic multi-centre RCT. *Trials* 2007;8(1):4. [PubMed: 17257440]