
Article

Sampling Hurdles: “Borderline Illegitimate” to Legitimate Data

Anupama Garg, MSc, DIA, PhD
Institute for Health Research
Lancaster University
Lancaster, United Kingdom

© 2008 Garg. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

In this paper the author discusses how sampling access and recruitment problems encountered in an in-depth interview study heightened her sensitivity to “borderline illegitimate” data. The term *illegitimate data* usually refers to the data collected during a covert study, whereas “legitimate” data are collected during an overt study. Hence, data collected during any nonconsented period(s) of an overt study lie on the borderline of illegitimacy and legitimacy, and constitute what the author calls borderline illegitimate data. Such data need legitimization before use. The borderline illegitimate data were collected during the pre- and postinterview stages of her study as they explained how medical and ethnic cultures and sensitivity to racism as a topic combined to create sample recruitment difficulties of the study. The author later legitimized them by sharing them with the participants, guaranteeing anonymity, and asking their permission to use them.

Keywords: sampling, ethics, consent, borderline illegitimate data

Author’s note: I am indebted to the Economic and Social Research Council, grant ref. PTA-026-27-1627, for funding my postdoctoral fellowship during which this article has been written. I would also like to thank Prof. Chris Hatton, Prof. Carol Thomas, Dr. Anne Grinyer, and the reviewers for their comments, which have helped to refine this article.

Introduction

To begin with, I will provide a brief introduction to my study, focusing on the ethical principles and practices I followed and the sampling access and recruitment problems I encountered. This will lead to a discussion of what are legitimate, illegitimate (Faculty of Social Sciences Committee on Ethics, 2002) and “borderline illegitimate” data; and to why and how I turned the latter type of data into legitimate data to be used in my thesis and subsequent publications.

My study, its ethical principles and practices, and the sampling problems

My study *Racism at Work: Indian Origin Doctors' and Their Wives' Well-Being* (2007) was theoretically underpinned by the Afrocentric feminist epistemology of Collins (1990) and Lazarus's cognitive mediational stress (CMS) theory (Lazarus, 1984; Lazarus & Folkman, 1984). Based on CMS theory the perception of racism was seen as the experience of racism, and it was conceptualized as a type of stress (Garg, 2007). Both of these theories consider individuals who are targets of stress as the only ones who can provide data on the nature of their stress, coping, and the stress's effects on them (Collins, 1990; Lazarus, 2000). In other words, they are the knowers. Furthermore, according to Lazarus, qualitative in-depth interview studies are more appropriate than quantitative studies for investigating stress and well-being as they are more accurate in revealing what is happening in individuals' lives and its personal significance to them.

It is also important to clarify why a feminist theory was used in a study (mainly) of and for men. This is based on my belief in the humanist vision of feminism which is committed to “eradicating the ideology of domination that permeates Western culture at various levels” (hooks, 1982, p. 195). Black British feminists of Indian and Pakistani origin also argue that challenging racism is intrinsic to feminism and highlight a history of Black women and men working together against racism (Amos & Parmar, 1997; Bhavnani & Coulson, 1997). A similar view was taken in this study, where I consider myself to be assisting in the struggle of male Indian-origin doctors against racism by providing information to the Indian Doctors' Associations about their members' experiences of racism at work. In this way, the study fulfills the primary aim of Black (and other) feminists to empower dominated group individuals so that they can resist their domination (Campbell & Bunting, 1991; Waugh, 1998). It is also important to note that although the study is focused on men, it is pro women because it highlights wives'/women's roles in times of adversity and shows the diversity and complexity of their lives which is one of the requirements of feminist research (Lernam, 1992). Some might still ask why Indian-origin women doctors' experiences of racism were not sampled. This is because ethnic minority women experience not only racism but also sexism at work (Healy, Bradley, & Mukherjee, 2003; Smith, Wadsworth, Shaw, Bhui, & Dhillon, 2005). Therefore, any effects on Indian-origin women doctors' well-being are likely to be the result of the cumulative effects of both racism and sexism at work (Kasl, 1995; Krieger et al., 2006), whereas I wanted to explore only the effects of racism at work.

Based on this theoretical guidance, data were first collected through separate face-to-face depth interviews with Indian-origin male hospital doctors (employees of the National Health Service [NHS]) and their wives on doctors' experiences of racism at work, the well-being of both doctors and their wives, and how the racism-related stresses and strains traveled between the spouses. These were followed by second clarification interviews with each participant over the telephone.

Sampling was designed after consultation with the leadership of the British Association of Physicians of Indian Origin (BAPIO) and consisted of advertising and snowballing. Advertising included a call for participants that was published in the electronic newsletter of BAPIO with an

introduction to me as the researcher; the study's objectives, methods, and inclusion and exclusion criteria; and information about how to contact me by mail, telephone, and e-mail (Garg, 2007). Snowballing involves existing participants' introducing new participants to the study until an appropriate sample size is reached or thematic saturation is achieved (Bowling, 1997) and is recommended where no sampling frame exists and prospective participants are likely to fear victimization (Denscombe, 1998). Therefore, snowballing was seen as a fallback strategy in this study as some difficulties with sampling were expected because racism is a politically sensitive topic and because of the medical culture of silence in which the prospective participants were embedded (Cooke, Halford, & Leonard, 2003; Hoosen & Callaghan, 2004). The medical culture of silence represents a culture wherein junior doctors do not complain about negative treatments, such as harassment, from their seniors, for three main reasons: (a) junior doctors think that their complaints will not be taken seriously and that complaining might even make their positions worse (Hoosen & Callaghan, 2004), (b) junior doctors need references from their seniors to get the next job and promotion (Health Policy and Economic Research Unit [HPERU], 2006), and (c) medical training, particularly surgical training, is largely an apprenticeship, the success of which greatly relies on good relations between juniors and seniors (Hargreaves, 1996).

Ethical principles and practices

The study's overall design, particularly the call for participants, participant information statement and consent forms, interview topic guide, and data transcription and storage (including an anonymization procedure involving the use of pseudonyms for participants and organizations, and copyright agreement), was submitted for ethical review and was approved by an NHS Multi-Centre Research Ethics Committee (MREC) in April 2005 (Garg, 2007). This approval was required as participants included health professionals who were employees of the NHS (Department of Health [DH], 2001). An open-ended consent process was built into the study (Economic and Social Research Council [ESRC], 2005; Gregory, 2003). In practice, this involved asking the participants whether they are happy to go ahead with their interviews and giving them a choice to withdraw at any stage of the research process without giving me any reasons for their withdrawal. In short, the study was completely overt, following accepted ethical principles and practices of social science research with health professionals, and was approved by an NHS MREC.

Sampling problems encountered

Only one couple volunteered to take part through advertising. Eventually a sample of eight couples was achieved through snowballing stemming from this couple in addition to one further couple who took part in the pilot phase of the study. In this way, recruitment was slow and low, even with snowballing. The recruitment was low as each couple referred one to four couples but only one or sometimes none of the couples participated in the study. The recruitment was slow as the average time from making a first contact to interviewing the couples was 2 months, the longest being 11 months for one couple.

With this context in mind, let me clarify the notion of both legitimate and illegitimate data, what I mean by *borderline illegitimate data*, and why I collected them.

Legitimate, illegitimate, and borderline illegitimate data

According to the ethical codes of research practice, researchers should collect only the type of data specified in the consent form and only during the specified periods for which participants' consent has been obtained (Faculty of Social Sciences Committee on Ethics, 2002). The data thus obtained and their use in research is considered legitimate. On the other hand, data collected in covert research, where participants are unaware of the purpose of research, how and what type of data will be collected, their rights to withdraw consent, and so on, are called illegitimate data (ESRC, 2005; Grinyer, 2001). In other words, illegitimate data collection involves a degree of deception and thus "violates both the principle of informed consent and the participant's right to privacy" (Faculty of Social Sciences Committee on Ethics, 2002, section: "Illegitimate" data; paragraph 2).

As discussed, as I had not only received written informed consent from the participants but was also following the open-ended consent process, my study was completely overt, with clearly identified legitimate data collection periods. These were face-to-face and telephone interviews.

In this environment of openness, when asked to refer new participants to the study, the participants sensed my difficulties with sampling, discussed their rationale for them, and expressed their concerns for my future. These discussions were held mostly during the postinterview period but were sometimes held during the preinterview period. The latter was the case when new participants had been made aware of sampling difficulties of the study by the previous participants that had referred them to me. These discussions were obviously outside the specified data collection periods in my study. If such data are collected in an overt study, I consider them to be borderline illegitimate data. In other words, data collected during any nonconsented periods of an overt study lie on the borderline of illegitimacy and legitimacy and constitute what I call borderline illegitimate data. I consider the collection and use of such borderline illegitimate data to be ethically defensible if the data are crucial for the study and the researcher guarantees participants' anonymity and obtains their permission to use them (Faculty of Social Sciences Committee on Ethics, 2002).

The borderline illegitimate data were crucial for two main reasons in this study. First, they explained how a combination of medical and ethnic cultures and sensitivities to racism as a topic underlay the sample recruitment difficulties of the study. Second, after 11 months of limited success in recruiting the participants, I would be able to use these data as a rationale for stopping the fieldwork as any further extension was unlikely to yield a large number of participants.

Tempting borderline illegitimate data and their usefulness

The following excerpts are part of the borderline illegitimate data. Their use was ultimately permitted by the participants, and participants' names are pseudonyms chosen by them.

Reasons for the failure of advertising

The following extract explains Dr. K's fears about taking part in the study.

Dr. K: Call me paranoid but these people are always watching.

Researcher: I would like you to elaborate on it. Is there an experiential basis for it or how did you reach this conclusion?

Dr. K: It is my experience from work. . . . [I believe] NHS managers are spies. They

record everything. I am sure they have things about me; I hope there is nothing against me.

This was followed by Dr. K's expressing worries about my future job prospects. He indicated that I could be viewed as antiestablishment and remembered as someone who had researched racism and is thus unemployable. This concern for me and Dr. K's earlier comment reveal his belief that if his participation in the study became known to his employers, he would lose his job. However, this mistrust in the NHS management was not unique to Dr. K; other participants had similar views. Lee (1993) has also highlighted that if a population is concerned about victimization in case their identity is revealed either during data collection or through deductive disclosure from the data included in publications, any involvement in the study is seen as a threat. Hence, if a community feels threatened and is secretive, it is likely to yield few, if any, participants through advertising (Wagstaffe & Moyser, 1987).

Another insight was provided by Dr. Raj, who said,

Most of its [BAPIO's] members are senior Staff Grades and Associate Specialists. These people are much older and have therefore perhaps gone beyond the point of frustration. They have lost an element of fight or they have reconciled to their level of achievement.

In this situation, taking part in the study would be considered a rather futile and unsettling exercise. Peter and Siegrist (1997) also found this withdrawal or "giving up" behavior among middle managers in a German car firm, especially among those who had experienced a lack of upward mobility despite trying hard in the early stages of their careers. Associate specialist and staff grade doctors, collectively referred to as middle-grade doctors, are equivalent to the middle managers of industry, who usually have little or no chance of career progression (Lemos&Crane, 2000). Therefore, many BAPIO doctors' age- and grade-related apathy could be another plausible reason for their poor response to advertising.

Reasons for the low recruitment rate

I have chosen the following quote from Dr. R's interview as he eloquently described the family-related reasons for the study's low recruitment rate: "[Your] difficulty with recruitment relates to the fact that most couples are not prepared to subject themselves to a great deal of intrusion in their private lives and open their heart to a stranger."

The most important aspects of Dr. R's explanation are "intrusion in private lives" and "stranger." The former raises two issues regarding family research. (a) The family is a private institution (Sapsford, 1993; Wharton, 2005). (b) Any study of the family, particularly one exploring the nature of interspouse relations, could be viewed as the worst kind of intrusiveness (Lee, 1993). Therefore, it is difficult to recruit couples in any family research. The "stranger" aspect is particularly relevant for South Asians as they are reluctant to talk about family matters with strangers or non-family members (Goodwin & Cramer, 2000).

Furthermore, Dr. R raised the question of the psychological costs of taking part in the study and said, "See your research is a living post-mortem. It is too painful to remember the past so not many couples can participate."

To build on Dr. R's explanation, I will argue that perhaps the psychological costs of recounting the painful past outweighed any gains for some prospective participants, leading them to not participate in the study. This could be understood within a framework of distributive justice,

particularly the exchange principle, which stipulates that an interaction can take place only if each party to the interaction estimates the cost and rewards of taking part to be roughly equal (Miell & Crogham, 1996; Turner, 2007). In other words, a prospective participant might decide to become a participant if his or her projected costs, such as time and both physical and psychological efforts, are roughly equal to the satisfaction gained from taking part in the study, for example the possible catharsis-related relief (Byrne-Davis et al., 2006) and the altruistic satisfaction of helping others in a similar position. Many participants mentioned the altruistic satisfaction as their main reason for participating in the study. Despite this, I would argue that perhaps for many prospective participants the anticipated psychological costs outweighed the benefits of taking part in the study, resulting in its low recruitment rate.

My sensitivity to the above insightful comments was high as they not only explained the sample recruitment difficulties but also showed the futility of extending the fieldwork in my study. Therefore, I noted them in my postinterview summary sheets. These were designed to be a simple housekeeping device, where I noted information such as implications for future data collection, possible improvements to the interviewing style, and who will transcribe the data, I or the research typist (Garg, 2007). However, despite being crucial, such borderline illegitimate data could be used only if I legitimized them.

Borderline illegitimate to legitimate data

At the beginning of the telephone interviews, I told the participants that I had noted their pre- and postinterview data about sampling difficulties and asked their permission to use them in my thesis and future publications. They were also given an assurance that these data, just like the rest of the data provided by them, would be anonymized by using pseudonyms. The relevant participants were also sent write-ups of my notes or the borderline illegitimate data, excluding any data analysis points (Robson, 1993), to check their accuracy. Participants had an option to modify these data if they felt they were inaccurate. All the participants agreed that their borderline illegitimate data could be used in my thesis and future publications, which, in turn, legitimized them. If the participants had refused, I would have not used their data as the use of data without permission is unethical.

Conclusion

The above illustrates, first, how participants could identify and explain sampling related difficulties in my study, and were thus the “knowers,” which is in line with the theoretical foundations of the study; and, second, how researchers must use participants’ explanations to generate a better understanding of the individuals, communities, and issues they are exploring. In addition, what are borderline illegitimate data? These are data provided by the participants and collected by the researcher during nonconsented parts of an overt study. The use of such data was ethically defensible in my study because (a) they were crucial in explaining the sampling difficulties encountered, (b) I shared these data with relevant participants and gave them the opportunity to amend them, and (c) I guaranteed participants’ anonymity through the use of pseudonyms and obtained their permission to use them. If the participants had refused the use of their borderline illegitimate data, despite their being crucial for my study, I would not have used them as it is unethical.

In this paper I have highlighted how medical and ethnic cultures and sensitivities of racism as a topic can combine to increase sampling difficulties in racism in medicine studies. These are perhaps the very reasons why the racism in medicine field is dominated by opinions with few

academic studies (Grant et al., 2004) and why racism at work and interspouse stress and strain transfer is a relatively unexplored area on an international basis (Garg, 2008). Considering this paucity of research and a need to build evidence in this field due to the increasingly multicultural nature of the British workforce (Kenny & Briner, 2007), I would argue that future studies should adopt quantitative methods, despite their pitfalls (Lazarus, 2000), as they are less likely to be affected by the aforementioned sensitivities related to culture and racism.

References

- Amos, V., & Parmar, P. (1997). Challenging imperial feminism. In H. S. Mirza (Ed.), *Black British feminism: A reader* (pp. 54-58). London: Routledge.
- Bhavnani, K., & Coulson, M. (1997). Transforming socialist feminism: The challenge of racism. In H. S. Mirza (Ed.), *Black British feminism: A reader* (pp. 59-62). London: Routledge.
- Bowling, A. (1997). *Research methods in health: Investigating health and health services*. Buckingham, UK: Open University Press.
- Byrne-Davis, L. M. T., Wetherell, P. D., Weinman, J., Byron, M., Donovan, J., Horne, R., et al. (2006). Emotional disclosure in rheumatoid arthritis: Participants' views on mechanisms. *Psychology and Health, 21*(5), 667-682.
- Campbell, J. C., & Bunting, S. (1991). Voices and paradigms: Perspectives on critical and feminist theory in nursing. *Advances in Nursing Science, 13*(3), 1-15.
- Collins, P. H. (1990). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. Boston: Unwin Hyman.
- Cooke, L., Halford, S., & Leonard, P. (2003). *Racism in the medical profession: The experience of UK graduates*. London: British Medical Association.
- Denscombe, M. (1998). *The good research guide for small-scale social research projects*. Buckingham, UK: Open University Press.
- Department of Health. (2001). *Research governance framework for health and social care*. London: Author.
- Economic and Social Research Council. (2005). *Research ethics framework (REF)*. Swindon, UK: Author.
- Faculty of Social Sciences Committee on Ethics. (2002). *Unpacking the moral maze: Ethical guidelines for social researchers*. Retrieved April 8, 2004, from <http://www.lancs.ac.uk/fss/resources/ethics/anonconf.htm>
- Garg, A. (2007). *Racism at work: Indian origin doctors' and their wives' well-being*. Unpublished doctoral dissertation, Lancaster University, Lancaster, UK.
- Garg, A. (2008). *Racism at work: Stress and strain transfer in Indian-origin medical marriages*. Unpublished manuscript.
- Goodwin, R., & Cramer, D. (2000). Marriage and social support in a British-Asian community. *Journal of Applied Social Psychology, 10*, 49-62.

- Grant, J., Jones, H., Kilminster, S., Macdonald, M., Maxted, M., Nathanson, B., et al. (2004). *Overseas doctors' expectations and experiences of training and practice in the UK* (Revised ed.). Milton Keynes, UK: Open University Press.
- Gregory, I. (2003). *Ethics in research*. London: Continuum.
- Grinyer, A. (2001). Ethical dilemmas in nonclinical health research from a UK perspective. *Nursing Ethics*, 2(2), 123-132.
- Hargreaves, D. H. (1996). A training culture in surgery. *BMJ*, 313, 1635-1639.
- Health Policy and Economic Research Unit. (2006). *Bullying and harassment of doctors in the workplace*. London: British Medical Association.
- Healy, G., Bradley, H., & Mukherjee, N. (2003). A double disadvantage: The workplace and union experience of minority ethnic women. *Equal Opportunities Review*, 121, 12-15.
- hooks, b. (1982). *Ain't I a woman: Black women and feminism*. London: Pluto.
- Hoosen, I. A., & Callaghan, R. (2004). A survey of workplace bullying of psychiatric trainees in the West Midlands. *Psychiatric Bulletin*, 28, 225-227.
- Kasl, S. V. (1995). Methodologies in stress and health: Past difficulties, present dilemmas, future directions. In S. V. Kasl & C. L. Cooper (Eds.), *Research methods in stress and health psychology* (pp. 306-318). New York: John Wiley.
- Kenny, E. J., & Briner, R. B. (2007). Ethnicity and behaviour in organisations: A review of British research. *Journal of Occupational and Organizational Psychology*, 80(3), 437-457.
- Krieger, N., Waterman, P. D., Hartman, C., Bates, L. M., Stoddard, A. M., Quinn, M. M., et al. (2006). Social hazards on the job: Workplace abuse, sexual harassment, and racial discrimination—A study of Black, Latino, and White Low-income women and the men workers in the United States. *International Journal of Health Services*, 36(1), 51-85.
- Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist*, 39(2), 124-129.
- Lazarus, R. S. (2000). Towards better research on stress and coping. *American Psychologist*, 55(6), 665-673.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.
- Lee, R. M. (1993). *Doing research on sensitive topics*. London: Sage.
- Lemos&Crane. (2000). *Annexe E, Tackling racial harassment in the NHS: Evaluating black and minority ethnic staff attitudes and experiences*. Retrieved July 26, 2007, from http://www.dh.gov.uk/prod_consum_dh/idcplg?IdcService=GET_FILE&dID=12759&Rendition=Web
- Lernam, H. (1992). The limits of phenomenology: A feminist critique of the humanistic personality theories. In L. S. Brown & M. Ballou (Eds.), *Personality and psychopathology: Feminist reappraisals* (pp. 8-19). New York: Guildford.
- Miell, D., & Crogham, R. (1996). Examining the wider context of social relationships. In D.

- Miell & R. Dallos (Eds.), *Social interaction and personal relationships* (pp. 268-318). Buckingham, UK: Open University Press.
- Peter, R., & Siegrist, J. (1997). Chronic work stress, sickness absence, and hypertension in middle managers: General or specific sociological explanations? *Social Science & Medicine*, 45(7), 1111-1120.
- Robson, C. (1993). *Real world research: A resource for social scientists and practitioner researchers*. Oxford, UK: Blackwell.
- Sapsford, R. (1993). Public and private. In M. Wetherell, R. Dallos & D. Miell (Eds.), *Interactions and identities* (pp. 129-147). Milton Keynes, UK: Open University Press.
- Smith, A., Wadsworth, E., Shaw, C., Bhui, K., & Dhillon, K. (2005). *Ethnicity, work characteristics, stress and health* (No. 308). Sudbury, UK: Health and Safety Executive.
- Turner, J. H. (2007). Justice and emotions. *Social Justice Research*, 20, 288-311.
- Wagstaffe, M., & Moyser, G. (1987). The threatened elite: Studying leaders in an urban community. In G. Moyser & M. Wagstaffe (Eds.), *Research methods for elite studies* (Vol. 14, pp. 183-201). London: Allen & Unwin.
- Waugh, P. (1998). Postmodernism and feminism. In S. Jackson & J. Jones (Eds.), *Contemporary feminist theories* (pp. 177-193). Edinburgh, UK: Edinburgh University Press.
- Wharton, A. S. (2005). *The sociology of gender: An Introduction to theory and research*. Oxford, UK: Blackwell.