Researching sensitive issues online: implications of a hybrid insider/outsider position in a retrospective ethnographic study

Professor Carrie Paechter
Educational Studies
Goldsmiths, University of London
New Cross
London SE14 6NW UK
c.paechter@gold.ac.uk

Abstract

This paper is concerned with methodological issues arising from a retrospective partially insider study of a divorce support website. I argue that, while we need to conduct detailed retrospective studies into the development of online communities, such studies bring methodological challenges, due to their retrospective nature, the potential size of the data set, and the problems of dealing with past manifestations of sites that continue to function. After an introduction to online and offline ethnography and insider/outsider researcher positioning, I discuss my hybrid insider/outsider status with respect to the research site. I focus in particular on researcher positioning, field entry and delimitation, public/private boundaries, ethical issues, and questions of time with respect to carrying out retrospective online studies while maintaining an ongoing real-time engagement with a research site.

Introduction

In this paper I reflect on some of the methodological issues involved in the retrospective study of the first year of a UK-based divorce wiki and support site, set up four years ago. I shall focus in particular on questions of insider/outsider status and how these affect entry to and delimination of the field, public/private boundaries, research ethics, and time compression in online archive research. I will discuss how these issues are particularly highlighted by the sensitive nature of the research focus, and indicate ways in which problems can be resolved.

These reflections arise out of the initial, pseudo-observational phase of an ongoing retrospective study of the first nine months of an online divorce wiki and support community. The site, Wikivorce, (www.wikivorce.com)[1] was set up in 2007 as an online resource for people going through divorce. My interest is in how the site became established, and in the factors that allowed it to operate successfully as an online learning community during its first nine months of operation. Given the increasing use of the internet as a source of advice and support, we need to have a clear idea about how online learning communities form and develop, and how they can be maintained as close, mutually responsible groups even as their success leads to rapid growth. This requires detailed analysis of such communities in the early stages of their development, and the study discussed here is intended to do this. In this paper I discuss the methodological and ethical issues involved in the earlier phases of this study, in which I was working almost entirely with the online records of activity during 2007, while interacting with the community as a user in real time.

The study was carried out retrospectively for several reasons. First, and crucially, I wanted to explore the development of an online learning community. It was only once the site was established that it became clear to me (as a user) that this was what it was. A retrospective study in a case of this kind also means that the researcher is not able to affect developments his or herself; there is clearly a danger of this if one sets out to study a new site in the hope that a learning community will emerge. Second, I wanted to be able to limit the data set, so that I could work with all that was available (rather than just particular topic areas) while still being able to examine aspects of it in detail. By the end of 2007 there were approximately 8 500 forum posts, which, while a daunting amount, could just about be handled using ethnographically-based methods if analysed through progressive iterative focusing techniques. Third, even retrospective research in this area is extremely sensitive, as I discuss below. Many Wikivorce members are deeply involved in traumatic, life-changing events that can feel beyond their control. A retrospective study permits participants to have gained a degree of emotional distance from the events being researched (Melrose 2002). Many discussions on the site also detail ongoing court cases, some of which involve children. I therefore judged that my research would be more acceptable to members if it focused on the past, rather than on current activity.

Ethnography, offline and online

Kozinets (2010)defines ethnography as:

> An anthropological approach to the research of culture based upon
> participant-observational techniques: ethnography's goals are a
> detailed and nuanced understanding of a cultural phenomenon, and a

representation that conveys the lived experience of culture members as

well as the meaning system and other social structures underpinning

the culture or community (190)

This superficially clear definition hides some contested issues. The first concerns

whether the key defining feature of ethnography is a process or a product.

Hammersley and Atkinson's (2007) highly influential account treats it as a research

process. They argue that while ethnography does not have a standard, well-defined

meaning, it usually includes five central features: people's actions are studied in

everyday contexts; data is collected from a range of sources, mainly informal

conversation and participant observation; data collection is relatively unstructured,

with interpretive categories emerging from the analysis; it is generally small-scale and

focused on a few cases; and data analysis involves interpretation of meanings,

function and consequences of human actions in context. This, they suggest, is a

development of people's ordinary modes of making sense of the social world.

Others, however, emphasise ethnography as a product. Forsey (2010), for

example, argues that ethnography 'is a research *product* based on methods aimed at

understanding and explaining the cultural context of lived experience' (75, italics in

original). Although both perspectives acknowledge 'thick description' as the central

outcome of ethnographic investigation, Forsey argues that the methods used have to

be adapted to the research context and that ideal typical accounts of ethnographic

practice as a primarily observational process can inhibit researchers in the field. In

particular, he challenges the idea that participant observation lies at the centre of

ethnographic practice. He suggests that what ethnographers actually do would more

accurately be described as 'engaged listening'.

The redescription of the ethnographic process as centred around listening rather than observing is salient for the practice of ethnography online. This is particularly the case when researching a community in which members communicate with each other entirely through synchronous and asynchronous text input. Such studies, because of their online location and reliance on text, cannot really be described as observational at all in the usual sense. This does not mean that the data is not extremely rich (Thomsen et al. 1998). That it is in many cases analogous to written speech, however, means that the characterisation of ethnographic practice as focused on listening captures the process much more accurately.

Research into online communities is now well established, and takes various forms. In this paper I want to focus on some of the implications of a partially insider positioning when conducting retrospective qualitative case study research online. Although it is possible, with some caveats, to conduct non-participative purely 'observational' studies of online communities (Langer and Beckman 2005), a considerable proportion of the work on internet communities is participatory. Such studies involve the researcher in signing up to the relevant community and interacting with it, either overtly as a researcher or covertly as a member. Like 'face-to-face' ethnography, they take place in real time. There are considerable debates about the ethical position of covert online research (Beaulieu 2004; Berry 2004; Ess 2009; Hine 2008; Langer & Beckman 2005; Mann & Stewart 2000; Markham 2005), which I shall not go into here except as they affect my own study, discussed below. However, in the overt and in some covert cases the researcher interacts as a member of the community, studying the group's actions and reactions as they take place. This

happens through participation in forum discussions and chat rooms, and interviews both online and offline using a range of methods and interfaces.

A salient issue here is what it means to participate in an online community, for both researchers and community members themselves. In face-to-face interaction, everyone present is visible: someone may remain silent but their presence will still be noticed by participants. Online life, by contrast, contains an explicit peripherally participant role, that of the lurker. A lurker is someone who follows online interaction without overtly participating in it. This can (in member-only communities) involve logging in but remaining silent, in which case others may register some level of presence (such as being listed as online). Alternatively, it can involve simply reading message posts as a guest. Such activity is not uncommon both for new users, who may lurk while they familiarise themselves with a site and decide whether to register (Hine 2008; Nonnecke et al. 2006) and for longstanding members, who may read discussion boards without logging in, or only log in when they want to post messages themselves. Nonnecke et al (2006), for example, note that many lurkers still think of themselves as community members, and that, in most communities, more people lurk than post.

This greater ambiguity about participation in online communities, compared to their face-to-face counterparts, again calls into question the nature of participant observation. If lurking is an acceptable and recognised means of participation in an online community, does this mean that it will be sufficient to allow a researcher to produce an appropriately 'thick' description of that community, or is it necessary to have a more active form of participation? Beaulieu (2004) argues that while lurking allows a researcher to follow events while working out a strategy for active

participation, avoiding interaction at all may lead one to miss things, and that participation also allows the checking of interpretations. She points out that initial lurking can be a risky strategy: people who have lurked for a while then come out as researchers can be rejected by the community. Others argue that active participation is essential to access the lived experience of community members (Hine 2005; Thomsen et al. 1998).

Insider and outsider researcher positions

Whether ethnographic research is considered to centre around participant observation or engaged listening, the researcher remains at the centre of the research process (Dwyer & Buckle 2009; Ellis 2011; Eppley 2006; Hine 2008; Hodkinson 2005; Humphrey 2007; Labaree 2002; Perryman 2011; Taylor 2011; Watts 2006). Social researchers are part of the world they study (Hammersley & Atkinson 2007), and their orientations both within and outside the field have to be reflexively accounted for in ethnographic writing. This requirement has led to considerable discussion about the relative importance of the researcher's position as insider or outsider to the field in question. As this is a key issue in the research under discussion in this paper, I shall introduce some of the issues at this stage, but will continue to address them throughout, as they apply to various aspects of the research process.

Traditionally, the ethnographic researcher is an outsider who journeys to another culture and attempts to capture its essential features through participation (Ellis 2011; Forsey 2010; Labaree 2002). In the last thirty years, however, there has been a move towards studying groups of which the researcher is, or has been, a member. This trend has led to discussion about the role of the researcher in such

contexts, much of which has focused around the position of the researcher as insider, outsider, or a mixture of the two (Eppley 2006).

Even when someone starts off as an insider, the research process itself challenges this status (Acker 2000; Hodkinson 2005; Humphrey 2007; Taylor 2011; Watts 2006). Labaree (2002) suggests that, while the mainly outsider researcher has to 'go native' in order to understand the local culture, insiders have, by corollary, to 'go observationalist', distancing themselves introspectively from phenomena. Insider positioning also necessitates the observation of oneself and one's relation to the research process; in this way, research makes outsiders of us all. While this is in some cases referred to as 'autoethnography' and seen as a subtly different approach (Ellis 2011), others have argued that accounts of the researcher's position in the field should underpin all ethnographically-based studies (Ball 1990).

There are considerable advantages to a mainly insider position, but it is important also to be aware of risks and disadvantages. Starting research as an insider can make the initial processes much easier. Insider researchers have argued that they are better placed to identify appropriate research questions, and that their prior knowledge of the field can make them less liable to being misled by participants (Hodkinson 2005). Most insider researchers also point to the comparative ease of access for someone already known. Such ready access, it is argued, is not just to the research site itself, but also to the 'real' perspectives of participants; insider research is considered by its practitioners to bring greater intimacy and openness to research interviews (Acker 2000; Breen 2007; Dwyer & Buckle 2009; Hodkinson 2005; Humphrey 2007; Keval 2009; Labaree 2002; Taylor 2011; Watts 2006). Generally, those researching from a mainly insider position argue that this status makes for

richer, thicker descriptions that are more likely to reflect the 'actual' community culture, though it is also pointed out that so much may be taken for granted within the community that things are left unsaid as 'obvious' which would be fully spelled out to an outsider (Acker 2000; Breen 2007; Dwyer & Buckle 2009).

On the other hand, insider positions can also bring drawbacks. Communities can have particular expectations when they are researched by one of their members, particularly around how the community is represented (Labaree 2002; Watts 2006). Taylor (2011) notes that there is an expectation that an insider will be sympathetic to the group in their analysis, and that the community may continue to see them as internally accountable (Humphrey 2007). Others have found that their insider position causes them to be 'read' in a certain way which may limit, rather than extend, their access. Insider research may also compromise, or be affected by, wider social relationships (Halstead 1996; Keval 2009; Perryman 2011; Taylor 2011; Watts 2006). For example, Humphrey (2007) found that those participants who were personal friends pushed her to disclose confidential information about others.

My own position in the Wikivorce study is a combination of insider and outsider. This is partly due to the context of the research and partly results from choices made during the negotiation of access. I am a longstanding member of the site, posting under a pseudonym, so in that sense I am an insider. I was an active member during the period I am studying and remain one today. On the other hand, my membership during 2007 was the same as anyone else's; apart from the constant analysis that is usual for any social researcher going about their daily life, I approached the site simply as a user. At that time, therefore, I was an insider, and a researcher, but not an insider researcher. As a researcher, my position in relation to

the data from 2007 is hybrid: I read the posts without intervening, but at the same time I have considerable insider knowledge. This is both a remembered knowledge of what it felt like to be on the site at that time (and in some cases a memory of reading a particular post for the first time) and knowledge of subsequent events, including how some of those members present themselves today. Furthermore, I have to be aware that I come to the research with my own pre-constructed assumptions about the community and what is happening in it (Labaree 2002). There is also the question of my dual role on the site. For reasons I explain below, in 2011 I have both a researcher role (as Carrie) and a participant role (under a pseudonym) and the connection between the two is known only to members of the management group. Given that Carrie does not post very frequently, and her blogs are about research, not about her feelings, she might well not be considered by most participants to be a full community member, although I (as Carrie) still feel that I have this status. As my pseudonymous persona, on the other hand, I am definitely an insider, and recognised as such, with (in 2011) over 1500 postings over four years, and an ongoing relationship with other longstanding members of the site.

My multiple position as a currently active Wikivorce member, while also being a participant in and now researcher of Wikivorce 2007, brings both advantages and disadvantages. Current membership, both as Carrie and under a pseudonym, allows me to remain aware of what it is like to be part of the community in real time, reading conversational 'turns' as they appear rather than as part of a whole thread, opening up the site to see what has happened to a particular individual or discussion, catching up with what one has missed in the evening after work. As Carrie, I also interact directly with Wikivorce members through an occasional research blog in the

site's blogging area. In addition to serving as a reminder that the research is taking place and an indication of the sorts of things it may come up with, this allows me both to feed back initial findings to members, some of whose 2007 interactions are themselves part of the data set, and to test out emerging concepts against member perceptions. This has proved a fertile way of communicating with participants and stakeholders, who have responded with perceptive comments.

While maintaining this link with the current site, it is, however, necessary to keep in mind that the 'feel' of the site in 2007 was in many ways quite different from the way it is now. Quite apart from the greater complexity of what is now offered, the speed of post turnover has increased enormously, and I am only fully aware of this through my memory of what it was like before. Clearly, my history as an insider is what gives me access to this change, and is a benefit of having that status. My insider knowledge as a long-term member has also allowed me to untangle some of the changing identities on the site (Taylor 2011). In some cases I have a memory of someone's change of pseudonym, and of the reasons behind it; in others I am able to trace probable changes by recognising writing styles, and then confirm them through people's responses to their posts (for example when someone writes 'I agree with X' when 'X' is now calling him or herself Y).

Nevertheless, it can be difficult to separate my knowledge of the site over four years from what I directly encounter in 2007. In particular I am aware of what has happened to people since, and I have to take care not to let this affect my analysis. There has also been the problem, common to those who research settings with which they are already involved, of making the familiar strange. I had hoped to be able to do this by using an archived version of Wikivorce from the Wayback Machine

11

(http://www.archive.org/web/web.php), but this has not been possible in practice, due to problems with the archived site. My approach to the issue has been twofold. Most importantly, in working through the 2007 threads I started with those which I rarely visit as a user, so spent two weeks reading only financial postings. Second, by taking the forum thread as a unit of analysis, and thereby reading 'conversations' whole rather than as posts come in, I have deliberately made a distinction between my usual interaction with the site and that required for research. The implications of this in terms of time compression are discussed below.

The innocent ethnographer in cyberspace: working through the insider/outsider interface

In this section I shall discuss some of the implications of my hybrid status for the research process as a whole. I shall focus on gaining entry, operating in the field, ethical issues, and questions of time in online qualitative research.

Entry

In negotiating entry into the field as a researcher I made full use of my insider position as a respected member of Wikivorce, both because it enabled me to know whom to contact, and because my posting history established my bona fides as someone committed to the site as a whole. I emailed the key gatekeeper, the owner and founder of the site, explaining who I was, identifying myself as a longstanding member, and asking his permission to carry out the research. This was granted, and we then agreed a process through which I would formally enter the site as a research field. I had originally envisaged that once the research commenced I would cease

posting as a regular member. However, it was suggested by the management team that instead I should (against the usual Wikivorce policy) have two identities, one continuing my pseudonymous presence on the site and one as a researcher, because 'It seems to me that carrie still has a lot to contribute as [pseudonym]' (email from management team member, June 2010). This indicated an anticipated clear distinction, borne out in practice, between the two personae, as well as a sense that something would be lost if I subsequently only posted as Carrie. After some discussion it was also agreed that the two identities would not be connected in any way, in order to protect my anonymity as a user. The ethical implications of this are discussed below.

This dual role within Wikivorce also, up to a point, deals with a problem identified by Howard (2002), which concerns what it means to 'enter' an online field. He argues that:

> for some researchers claiming to do ethnography online, going into the
> field is little more than a state of mind because there is so little
> convergence between their lives and the subjects' lives: there is no
> physical entry into or exit from the community. (559)

Having a dual identity meant that it is obvious both to me and to the community when I am participating in Wikivorce as a user and when as a researcher, as well as giving me a clear point (signing on in my researcher role) at which entry to the field takes place each day. It allowed me to share in the day-to-day life of Wikivorce (albeit in 2010 and 2011) and retain a sense of the actual (rather than the archived) functioning of the community, as well as to share my thoughts and findings with members as the research proceeds. I shall discuss the implications of this further below.

The fieldwork began formally with a permanent announcement about the research, posted by the managing team, on the front page of the site. This both establishes my bona fides as a longstanding member trusted by the owner and alerts members to the research taking place. This announcement is linked to an article in Wikivorce's zine, in which I am once again introduced as a trusted longstanding member and in which I describe the research, with a link to my university homepage so that people can see who I am and the work I do. The response to this was very muted, in contrast to the experiences of others seeking to research online communities (Hine 2008; Kozinets 2010).

What is the field?

A key issue in online qualitative research is what constitutes the field (Hine 2005), and how the researcher enters and operates within it. Markham (2005: 801) argues that in both online and face-to-face research, 'drawing boundaries around the research context, or "identifying the field" involves a series of decisions that both presuppose and reveal the researcher's underlying ontological and epistemological assumptions', as well as affecting the researcher's choices and practices (Markham 2004). In my own case I took the 'field' to be everything that I could still access that was available in Wikivorce between March and December 2007. The caveat 'that I could still access' points to a number of limitations. First, the Wiki-based support material undergoes sporadic revision, which means that it is difficult for me to see, even through site archives, what was available in 2007. Second, I have no access to the chatrooms of the research period. Third, it is possible to edit entries. Although I have no certain way of telling, it is my belief (as a regular user) that most entries have

not subsequently been edited, except occasionally to hide identifying features such as names of children. They are, however, occasionally removed entirely, by the management, usually for legal reasons, so leave no public trail at all. Furthermore, one very prolific poster during this first year deleted almost all of her posts early in 2008 during a dispute with the owner (since resolved): while I can see that she made a comment, in most cases it is impossible to tell what it is. Some people have removed their profiles or blog pages from the site, or had them removed by the moderators. Finally, it is not uncommon for people to change their pseudonyms, usually to reflect a change in their feelings about themselves (for example, from 'devastated' to 'reborn'): when this takes place it affects all their forum postings, past and present. Although it is generally possible for the alert reader to tell from the forum lists that this has happened, it means that what I am seeing in the 2007 sections of the forum is not exactly what I would have seen at the time.

Even given these limitations, the data set was enormous: my field of research comprises approximately 8 500 forum postings, arranged in conversational threads, from March to December 2007, plus 475 blog entries (and associated comments) from the same period. 'Conversations' could comprise anything from two posts (or, occasionally, a single one unanswered) to over sixty. This is clearly a considerable amount of data to handle in any rigorous qualitative way. While some previous researchers studying smaller online communities have been able to analyse every post in detail over an extended period (Fayard & DeSanctis 2005; Langer & Beckman 2005), this approach was clearly not possible given such a wealth of data.

In thinking about how to approach the analysis of such an enormous amount of available information, I considered how I would behave in the early days in a

15

research field if I were doing an analogous face-to-face study. Of course such analogies only work up to a point. In particular, in face-to-face case study research it is not possible to have access to such a high proportion of interpersonal interactions. Given that my initial approach when entering a new field would be simply to observe as much as possible for a while, I decided to regard the forum postings initially as observational data. I read all the forum entries for the whole of the period under study, taking notes, which I treated as field notes. I did this by working through all the threads in a particular topic-based section, starting with the first thread in the section and continuing thread by thread. This allowed me to select particular threads that could be analysed in more detail, using traditional data coding techniques. I was also able, through this process, to identify a group of potential virtual (and in some cases, actual) key informants. These individuals were then followed up (using links from their profiles), tracking and analysing both their blog entries and their 2007 forum posts. This dual approach to the site content allows the community to be investigated both through the lens of the forum 'conversation' and from the point of view of the contributing individual.

Ethical issues

The ethical issues involved in a study of this kind are considerable. First, the research area itself is sensitive, and the data extremely so. Second, the ethics of online research, particularly in social networking sites, are problematic and contested. This is partly related to more general insider/outsider debates and partly because of the nature of online communities and communication.

16

Data within Wikivorce is sensitive in several ways. First, it involves family law cases, which are subject to considerable confidentially safeguards. Second, because many people use the forum as a place to ask for advice about possible divorce settlements, they post financial information in considerable detail. This means that, as a researcher, I would have easy access to such information in relation to particular individuals if those people reveal their real identities to me as part of the research. Madge and O'Connor (2002) note that one of the problems with web-based studies is that it is impossible to verify the identity of participants: in my research it is probably better that I cannot, despite potential difficulties with respondent veracity.

Third, the data is sensitive because it comes from people at a particularly traumatic time of their lives. This, coupled with the disinhibiting effects of anonymous online communication (Barak et al. 2008), means that many posts are very emotionally raw. Even with the passage of time, care has to be taken to use this material sensitively and empathically. This issue is complicated because of the way people treat online social media as at least partially private spaces, despite their public accessibility(Barnes 2006). Berry (2004), for example, argues that online communities are bound by respect and trust among members. This view was echoed in a post from November 2007 in which a member, having realised with surprise and alarm that anyone could read her posts, suggested that they should be accessible to members only. In making this suggestion she seemed to be implying that members could be trusted with personal information whereas outsiders could not. As the site at this time had over 6000 members, and anyone with an email address could join with a minimum of trouble, this was a somewhat naïve view, but one that other researchers suggest is widely shared elsewhere (Zimmer 2010).

This brings us to the question of informed consent and what it means in the context of publicly accessible material. There are two central problems here: one concerns the status of publicly available forum posts and how one handles such material; the other is to do with informed consent and who can give consent to internet-based research. The status of forum postings remains a matter of debate in internet research. Some researchers claim that as public documents they do not require any form of anonymity or consent for quotation, while others suggest that they should be treated as literary works, with attendant copyright issues (Kozinets 2010). Langer and Beckman (2005) argue that analysing message board contributions is analogous to doing the same with readers' letters to newspapers, and that it is consequently not necessary to inform the authors or to use alternative pseudonyms.

Markham argues that the issues arising from internet research require 'astute, reflexive methodological attention' (2005: 815). She suggests that:

> Instead of asking "how we can protect human subjects through various
> types of research design?" we will frame better questions and find
> richer answers by shifting our focus toward the participant. Putting the
> human subject squarely in the center of the research both shifts the
> ethical considerations and allows for socially responsible research.
> (815)

My partially insider status was instrumental in providing an approach to this issue. I considered how I, as a Wikivorce member, might feel about someone using my own 2007 postings as part of their research. My conclusion was that it was reasonable to expect the site owner, as a trusted and involved participant, to give consent to the research as a whole on behalf of the community, but that I would want to give

individual consent, probably using my online pseudonym, for any interviews, and that it should not be made any easier than it is now for my postings to be 'joined together' by a subsequent reader to give a more comprehensive picture of me. Shoemaker (2009) points out that what upsets people about data mining is that, while it draws on individual pieces of publicly available data, it brings them together in ways that as a whole undermine privacy. There is an analogous issue in conducting the Wikivorce study: in any individual posts people do not give away a great deal about themselves, but by reading someone's entire posting history I can obtain, and communicate, a much more detailed picture. This in itself might be worrying for me as a participant, but combined with individual interviews it could have seriously distressing results. Consequently, I felt that, as a member, I would not want it to be possible for anyone reading the research to be able to connect my public-domain posts with anything that I might say in a private interview. This issue is in turn exacerbated by the tendency, in both online and face-to-face research, for participants to attempt to 'crack' the pseudonym code (Kozinets 2010). Given that it takes only seconds to enter direct quotations into a search engine and come up with the original posting, this has the potential to leave members open to exposure to others in the community or beyond, including their ex-spouses (Couldry 2008). While there is nothing in quoted forum posts that is not already in the public domain, the juxtaposition, in a research analysis, of posts from the same individual, can leave someone exposed in ways that may upset or otherwise harm them.

This led to the following approach: Where I quote from forum posts, I do so verbatim (including idiosyncratic stylistic markers, spelling and grammar errors, and attendant emoticons). I think this is important from the point of view of the research,

as it gives an accurate picture of how the forum 'feels'. Where appropriate, within these quotes, I replace names with others chosen more or less at random (replacing a consistent pseudonym with an inconsistent one). I do not attribute even pseudonymous names to the authors of the posts quoted, though anyone wanting to find out who said what could do so very easily. This means that, if I quote the same individual more than once, it is not obvious except to someone who deliberately checks the origin of each quotation, so that it is this person, rather than myself, that is performing the 'joining up' operation. Furthermore, when I move to the interview phase of the study, I will use entirely different pseudonyms for those interviewed, and will not connect interview data in any way with individual postings while writing up. Although this will disadvantage the analysis somewhat, I believe that it is the only way in which I can preserve the confidentiality of interview data.

Related to this is the issue of informed consent in online contexts. Markham (2005) points out that it is difficult to obtain informed consent from an actual person if they desire anonymity. This issue is complicated by the attachment of people to their online pseudonyms and their concern for their online persona's reputation (Dibbell 1993; Kozinets 2010). In some ways it could be considered sufficient to obtain informed consent (for example, for an online interview) from the online persona, although this runs up against the impossibility of checking whether online individuals are self-responsible adults or members of vulnerable groups. These questions are not going to be easily resolved, and may come down to a requirement for the researcher to take the responsibility of acting with integrity and care (Markham 2005) rather than follow legalistic ethical guidelines.

My dual role as Wikivorce member and researcher with different personas, adds a further layer of ethical complexity. As Carrie, I am openly interacting with the site when logged in with my real name. This signals to members (if they remember that I am a researcher) that anything they say to me becomes information that I know for the purposes of my research. The site works in such a way that in order to send anyone a private message you have to first go to their user profile, where my role is clearly explained, so I think this aspect of things is relatively unproblematic. However, in my pseudonymous persona, I am a longstanding member with a personal history, which includes interactions going back nearly four years with a good number of members. Some of those members know considerable detail about my personal circumstances and, more important here, I know a fair amount about theirs. This is not public information: it is information that I was given as a fellow member (not as a researcher) and thus in a different position of trust. Dealing with this knowledge (which I can't just forget (Taylor 2011)) requires me once again to put the participants and my responsibility to them, at the centre of my research. This responsibility includes not mixing my two personae, including never being open about the connection between them, so that people do not become afraid that things said to me in my pseudonymous role will be incorporated into the research. While this is sad for me (because, for example, it means that I cannot attend occasional face-to-face events in the persona that others are most familiar with) it is an inevitable consequence of the research process.

A final way in which I have attempted to '[put] the human subject squarely in the center of the research' (Markham 2005: 815) has been through including the current Wikivorce community in my thinking about some of these issues. While I was

preparing this paper I also wrote a blog entry about the issue of the private/public boundary of the community, and how people treat their forum and blog posts, with a view both to alerting people to the issue and to eliciting their views about it. This led to a lively discussion that suggested both that people were generally aware of the public nature of the site, and that they partially got around this by using the chat and personal message functions. Most respondents, indeed, saw a clear distinction between the public and private parts of the site, and posted accordingly:

> So yes wiki is public but as time goes by members often get to know
>
> one another and confide, ask questions, give and receive support, so i
>
> suppose there are two sides to wiki the one you see on the forum and
>
> the other in chat and private messages.

This suggests that, contrary to appearances, members protect themselves and their identities both from their former spouses and from other interested or curious parties. This reflects Couldry's (2008) view that people adjust their personal stories on social networking sites, holding back information that they would prefer not to make public, thus protecting an older public/private boundary.

Questions of time

In this final section I want to consider some of the time-related issues that have arisen out of simultaneously carrying out retrospective archival research while maintaining a longstanding real-time presence in the same online community. These all concern the relationship between my archival reading of Wikivorce posts and the real-time experience of being a member of the community. In some ways this is about my ability to move between roles and personae but in others it is about what it means to

do online research as a community member. In also underlines the need for even retrospective, archive-based researchers to experience real-time interaction in a community.

The most obvious difference between the experience of doing archival research and participating in a community in real time is the way that time becomes compressed in the archives. Over the course of two months, I read nine months' worth of forum posts. In addition, I read them differently from the way that a 'normal' member would. The asynchronous nature of the forum means that a conversation of, say, eight turns, might take place over two days, or an evening, or even, in some cases, several weeks, months or years, with the discussion being picked up through updates or being found by someone who had not originally been party to it. Except in cases where over a month has elapsed between conversational turns, this is completely opaque to the archive reader. Furthermore, the experience of reading entire archived threads is qualitatively different from that of being a real-time participant, because it lacks the experience of wondering what has happened in a particular thread, or to a specific person, and looking out for the next posting. The conversation is presented as a completed whole, in which the option to participate has been, if not removed, certainly rendered pointless. In cases where a member is asking for support over something that is happening to them that day, this emphasises the researcher's separation from the events and helps to make the familiar strange. At the same time, the experience of reading forum posts all day for days on end brings a different sort of intensity to the researcher's relationship with the community.

The difference in intensity of the experience of Wikivorce as a 2007 archive and as an ongoing website and forum is particularly salient given that the site was still

in its infancy during the research focus period. The far greater volume of traffic in 2011 means that the experience of community membership is now qualitatively different: for example, one can no longer easily log on at the end of the day and read and reply to all posts in the previous 24 hours. In analysing the data, I therefore have to use my insider experience to access a sense of what Wikivorce used to feel like: a continuing presence on the site is insufficient and even in some ways misleading.

Similarly, I have to be exceptionally careful not to allow my perceptions of the community as it is today to affect my analysis of what it was in 2007. Like most qualitative researchers, I sometimes start with an analytical 'hunch' that then has to be checked out against detailed analysis of the data. There have been occasions when such supposed leaps of understanding have not been borne out in my detailed analysis, and I have been forced to the conclusion that they stemmed from my current participation, not from what was going on in the research period. This forgetfulness about which time period one is in is likely to be a common problem for people working with internet archives while maintaining real-time interaction. It is particularly difficult in this case because the only way reliably to access the forum archive is through the current site. This suggests that, where possible, retrospective studies should be carried out using archived versions of sites, rather than their current forms.

Conclusion

Internet research, as an ethnographically-based practice, is well established and, as discussed above, debates in the field are already reasonably mature. However, the pace of change of the field as a whole means that these discussions need to remain

24

active, to reflect the size of potential datasets, the nature of participation and the characteristics of participants. In particular, much early research into internet-based groups focused on MUDs and MOOs, which had a much more contained and specialist range of participants than current online fora.

As the online world is increasingly important in relation to people's offline lives, it is important to study in detail the communities in which people participate. This explosion of participation, however, brings with it considerable methodological and ethical difficulty. Data sets can be enormous: I found the Wikivorce 2007 data set daunting, but the comparable data from 2011 would be impossible to work with in the same way. The nature of researcher participation, participant awareness, and informed consent all need careful consideration, and solutions to the attendant problems found within the context of the particular study. Online research practice also once again brings to the fore the question about the justifiability of covert research: for example, what are the ethical implications arising from the possibility of a researcher who is truly invisible, because the data is already in the public domain (Langer & Beckman 2005)?

Although most internet-based qualitative research to date has been carried out in real time, there remains the need for retrospective studies. This is particularly important if we want to understand how online communities form and develop: it is hard to tell, when a site is set up, whether a community will form around it, and, if so, whether its formation will be a fertile subject for study. Consequently, if we are to illuminate how successful online communities become established (and, indeed, why some do not), we need to study their histories. This brings questions of insider/outsider status, consent, ethical practice and methodology which I have

attempted to address but with which I am still struggling. These issues are especially

pertinent if the community coheres around a sensitive subject such as divorce or other

life changes, or around mental or physical illness.

Although some methodological and ethical approaches can be transferred

between face to face and online qualitative research, we need to rethink some of our

practices and their meaning as we take them into online arenas. In doing this, the most

important thing is to maintain the position of the participants at the centre of the

research, and to attempt to act throughout with integrity and good faith.

References

Acker, S. (2000) In/out/side: positioning the researcher in feminist research, *Resources for feminist research*, 28(1-2), 189-210.

Ball, S. J. (1990) Self-doubt and Soft Data: social and technical trajectories in ethnographic fieldwork, *Qualitative Studies in Education*, 3(2), 157-171.

Barak, A., Boniel-Nissin, M. & Suler, J. (2008) Fostering empowerment in online support groups, *Computers in Human Behavior*, 24, 1867-1883.

Barnes, S. B. (2006) A privacy paradox: social networking in the United States, *First Monday*, 11(9), no page numbers.

Beaulieu, A. (2004) Mediating ethnography: objectivity and the making of ethnographies of the internet, *Social Epistemology*, 18(2), 139-163.

Berry, D. M. (2004) Internet research: privacy, ethics and alienation: an open source approach, *Internet Research*, 14(4), 323-332.

Breen, L. J. (2007) The researcher 'in the middle': negotiating the insider/outsider dichotomy, *Australian Community Psychologist*, 19(1), 163-174.

Couldry, N. (2008) Mediatization or mediation? Alternative understandings of the emergent space of digital storytelling, *New Media and Society*, 10(3), 373-391.

Dibbell, J. (1993) A rape in cyberspace, *Village Voice* (New York.

Dwyer, S. C. & Buckle, J. L. (2009) The space between: on being an insider-outsider in qualitative research, *International Journal of Qualitative Methods*, 8(1), 54-63.

Ellis, C. (2011) Autoethnography: an overview, *Forum: Qualitative Social Research*, 12(1), Article 10.

Eppley, K. (2006) Review Essay: David Weaver-Zercher (ed) (2005) Writing the Amish: the worlds of John A. Hostetler, *Forum: Qualitative Social Research*, 7(3), Article 16.

Ess, C. (2009) Internet Research Ethics, in: A. Joinson, K. McKenna, T. Postness & U.-D. Reips (Eds) *Oxford Handbook of Internet Psychology* (Oxford, Oxford University Press), 487-502.

Fayard, A.-L. & DeSanctis, G. (2005) Evolution of an online forum for knowledge management professionals: a language game approach, *Journal of Computer-Mediated Communication*, 10(4).

Forsey, M. (2010) Ethnography and the myth of participant observation, *Studies in Qualitative Methodology*, 11, 65-79.

Halstead, F. R. (1996) The effects on staff morale of a change of leadership and period of industrial dispute in a secondary school., *School of Education* (Milton Keynes, The Open University).

Hammersley, M. & Atkinson, P. (2007) *Ethnography: principles in practice* (London, Routledge).

Hine, C. (2005) Internet research and the sociology of cyber-social-scientific knowledge, *The Information Society*, 21, 239-248.

Hine, C. (2008) Virtual ethnography: modes, varieties, affordances, in: N. G. Fielding, R. Lee & G. Blank (Eds) *The Sage Handbook of Online Research Methods* (London, Sage), 257-270.

Hodkinson, P. (2005) 'Insider research' in the study of youth cultures, *Journal of Youth Studies*, 8(2), 131-149.

Howard, P. N. (2002) Network ethnography and the hypermedia organization: new media, new organizations, new methos, *New Media Society*, 4(4), 550-574.

Humphrey, C. (2007) Insider-outside: activating the hyphen, *Action Research*, 5(1), 11-26.

Keval, H. C. (2009) Negotiating constructions of 'insider'/'outsider' status and explaining the significance of dis/connections, *Enquire*, 4, 51-72.

Kozinets, R. V. (2010) *Netnography: doing ethnographic research online* (London, Sage).

Labaree, R. V. (2002) The risk of 'going observationalist': negotiating the hidden dilemmas of being an insider participant observer, *Qualitative Research*, 2(1), 97-122.

Langer, R. & Beckman, S. C. (2005) Sensitive research topics: netnography revisited, *Qualitative Market Research*, 8(2), 189-203.

Madge, C. & O'Connor, H. (2002) Online with e-mums: exploring the internet as a medium for research, *Area*, 34(1), 92-102.

Mann, C. & Stewart, F. (2000) *Internet Communication and Qualitative Research: a handbook for researching online* (London, Sage).

Markham, A. (2004) Representation in online ethnographies: a matter of context sensitivity, in: S.-I. Chen, J. Hall & M. Johns (Eds) *Online Social Research: methods, issues and ethics* (New York, Peter Lang).

Markham, A. (2005) The methods, politics and ethics of representation in online ethnography, in: N. K. Denzin & Y. Lincoln (Eds) *The Sage Handbook of Qualitative Research 3rd Edition* (London, Sage), 793-820.

Melrose, M. (2002) Labour pains: some considerations on the difficulties of researching juvenile prostitution, *International Journal of Social Research Methodology*, 5(4), 333-351.

Nonnecke, B., Andrews, D. & Preece, J. (2006) Non-public and public online community participation: needs, attitudes and behavior, *Election Commerce Research*, 6, 7-20.

Perryman, J. (2011) The return of the native: the blurred boundaries of insider/outsider research in an English secondary school, *International Journal of Qualitative Studies in Education*, ifirst article.

Shoemaker, D. W. (2009) Self-exposure and the exposure of the self: informational privacy and the presentation of identity, *Ethics and Information Technology*, 12, 3-15.

Taylor, J. (2011) The intimate insider: negotiating the ethics of friendship when doing insider research, *Qualitative Research*, 11(1), 3-22.

Thomsen, S. R., Stroubhaar, J. D. & Bolyard, D. M. (1998) Ethnomethodology and the study of online communities: exploring the cyber streets, *Information Research*, 4(1), no page numbers.

Watts, J. (2006) 'The outsider within': dilemmas of qualitative feminist research within a culture of resistance, *Qualitative Research*, 6(3), 385-402.

Zimmer, M. (2010) "But the data is already public": on the ethics of research in Facebook, *Ethics and Information Technology*, 12, 313-325.

---

[1] I have made no attempt to anonymise the site. Given my overt presence on the site it is extraordinarily easy to find it using a search engine. Consequently I decided that it would be better overtly to publicise the site than to make a fruitless attempt to obscure it.