# Ethical dilemmas using social media in qualitative social research: A case study of online participant observation

Dr Kath Hennell

Liverpool Hope University (UK)

Professor Maria Piacentini

Lancaster University (UK)

Dr Mark Limmer

Lancaster University (UK)

## Abstract

Social media platforms that enable users to create and share online content with others are increasingly used in social research. This paper explores the complex ethical issues associated with using social media for data collection, drawing on a study of the alcohol consumption practices of young people. It aims to contribute to debates about the practical and ethical challenges facing researchers using social media as a data collection tool, and to demonstrate how a reflexive approach to the research and the context in which the research takes place is critically important for supporting and enabling an ethical approach. The article concludes by recommending that researchers who face ethical dilemmas associated with the use of social media maintain an on-going dialogue with their relevant ethics committees and other researchers to identify potential solutions and to share their findings.

## Introduction

The widespread use of social media in everyday life has seen a corresponding increase in the use of social media content as a data source for research studies (Townsend and Wallace, 2016). New platforms, and their associated data forms, have emerged rapidly with new and particular affordances, and this presents new challenges for ethical practice, guidance and scrutiny processes, as they evolve to adjust to this changing research landscape (Williams et al., 2017). This paper explores four key interrelated areas of ethical concern (risk of harm, private vs public data, informed consent, and anonymity) in the virtual research world (Townsend and Wallace, 2016), taking a public health study that uses social media data, in the form of online participant observation, as its context.

Legislative frameworks govern any ethical journey and this particular journey sits within the legislative framework of the UK, European law and the European court of justice. At the point when the data in this study were collected research of this type was subject to the UK Data Protection Act 1998 and the European court of justice ruling on the ‘Right to be Forgotten’. However, since then new regulation (the European General Data Protection) has updated these provisions and clear affirmative action is now required in the giving of consent for the use of personal data. A more detailed discussion of the constantly evolving landscape of data management legislation and guidance lies outside the focus of this study.

Halford (2017) highlights some tensions that exist between formal ethics procedures and social media research including that social media studies draw on already created data (in contrast with empirical research studies that generate new data), and hence data cannot be protected in the familiar and traditional ways. Further, the interconnected nature of, and capacity for, social media data to evolve and recirculate in virtual spaces places the data outside the control and protection of the researcher. In recognition of these tensions ethical approaches have been developed to address the specific requirements of social media research. A fruitful avenue is offered by adopting a ‘situational ethics’ approach (British Sociological Association (BSA), 2002; British Psychological Society (BPS), 2013), which enables the researcher to respond to ‘ethically important moments’ that arise during the course of the research, rather than relying on the *a priori* prediction of ethical concerns (Miller, 2012). By situating the research within its context, ethical decisions can be made on the basis of emergent issues as they arise, whilst using social research principles as a guide (BSA, 2002; Social Research Assocation (SRA), 2003).

Here, we build on the situational ethics perspective and contribute to the growing body of literature by examining and reflecting on ethical dilemmas, challenges and concerns related to the use of social media within research. Specifically, we map out four areas of ethical concern when conducting research in the social media space, and then reflect on our experiences of these ethical challenges as they emerged in our empirical study. Through our engagement with the key ethical concerns, we respond to calls for research that contributes to the academic conversation about ethical standards in social media research (Townsend and Wallace, 2016; Halford, 2017; Johnson et al., 2013). By highlighting the ethical challenges emerging from the first author’s PhD research (a qualitative study of young people’s alcohol practices, drawing on focus group data and participants’ social media content), we demonstrate the importance of being able to respond reflexively according to the ethical situation that we are confronted with. This approach is epistemologically in line with the dynamic and changing nature of social media as a research context (Association of Internet Research, 2012).

Qualitative social media research is not homogenous and there are many different approaches to using social media in research including experimental manipulation, capturing and analysing large amounts of publicly available social media content and participant observations (Woodfield and Iphofen, 2017). In addition, studies can be a snapshot in time, longitudinal or over a specific time period. One of the most commonly used ways in which social science researchers have used social media data is where publicly available data is collected using data mining tools to study communications and interactions (Williams et al., 2017). The methodological and technical challenges of using this type of big data within research have been well summed up by Williams et al. (2016) as the 6Vs of Volume, Velocity, Variety, Veracity, Virtue and Value. Attention to ethical dilemmas within other types of social media research is less well documented, however, and this paper provides a much-needed broader dimension to these ethical debates on social media research. Our purpose here is to reflect on a range of ethical issues that may be faced by qualitative scholars who analyse social media content or interactions as part of their research. We provide a literature review of key themes and then discuss a range of issues that were negotiated as part of our own research. Finally, we conclude by offering some ethical guidance and suggestions for qualitative researchers undertaking social media research.

## Literature review

This section summarises some of the key debates within the existing wider literature on using social media data in research and draws on Townsend and Wallace (2016)’s four key interlinking areas of ethical concern within the virtual research world: risk of harm, private vs public data, informed consent and anonymity.

### Risk of harm

There is a general consensus that social science research should do no harm (Webster et al., 2013) and while we recognise that the potential for social research methods to cause any serious physical harm to participants is low (Dingwall, 2008), our understanding of harm is as a wide-ranging concept, including serious physical harm, but also psychological and social harms (European Commission, 2010). Harms can include, but are not exclusive to: users being embarrassed by the research; the research leading to reputational damage affecting a user’s social or family life or their career; online or offline harassment; or being subject to menacing behaviour or differences in how people treat them. Dawson (2014) describes a study where a researcher published the usernames of members of an online community, which led to unwelcome, additional attention and scrutiny of one of the participants, who felt they had to leave the community after being involved for many years. Williams et al. (2016) advise that the publication of content may also open up the possibility of harm if the content is sufficiently inflammatory, citing the example of racist content. These wider harms are possible in the context of social media research particularly if user anonymity is breached or if content is shared with new and unintended audiences at the risk of embarrassment, reputational damage and/or physical harm (Townsend and Wallace, 2016).

### Private vs public data?

The public nature of much social media content (Beninger, 2017) leads to a blurring of the boundaries between the public and private domains, thus creating challenges around the level of ethical regulation required for the collection, analysis and dissemination of data gathered from different on-line contexts (Sloan and Quan-Haase, 2017; Robards, 2013).

To aid decision-making in this arena, much of the debate has centred around the public vs private nature of this type of data, addressing issues around defining privacy in this very public domain, particularly when access to content is unrestricted (Beninger, 2017; Woodfield and Iphofen, 2017). Accordingly, a general distinction has been made in the literature between social media data that is publicly accessible (and potentially therefore ethical for research use), and social media accounts that make use of restricted permissions with the highest possible privacy protections (thus potentially unethical to use for research purposes without consent) (Woodfield and Iphofen, 2017). Based on these distinctions Facebook and Instagram content can be regarded as private, as users can control who views their data through privacy settings (Kosinskietal., 2015). This is despite some Facebook data (namely basic Facebook profile information) being publicly available and indexable by search engines. Conversely, the majority of Tweets are in the public domain and accessible by anyone, even without a Twitter account. Although Twitter users can choose to protect their Tweets and restrict them to approved followers, in practice this is rare (Lafferty and Manca, 2015; Woodfield et al., 2013). However, privacy and expectations of privacy are not straightforward.

Zimmer (2010) explores some of the key issues relating to the public vs private status of data, providing an account of how Facebook data from a group of students gathered for a specific project (the 3T project), was released and used for other purposes. Zimmer argues that the students in that study did not consent to or expect their publicly available social media data to be used in this way, and thus the researchers effectively breached their contextual privacy. That is, the assumed level of privacy is based on the norms and expectations within any given social media platform (Brown et al*.*, 2016). For example, sharing with friends on Snapchat and posting on a chatroom forum are likely to be governed by different information norms. There were also additional concerns expressed about the potential for identification of individuals through the combination of the two data sources, emphasising the blurring of the public/private line (Zimmer, 2010).

The status of data as public or private is not definitive or static, since the affordances of social media enable text and non-text content to be recirculated, commented on, shared and changed. The effect is that in some cases, data that was originally been considered private, have come to be made public – a form of data described by Lange (2007) as publicly private/privately public. Facebook’s permitted data use policy informs users that information and content shared by users but created by another person is not under the control of the person who originally posted the content (Facebook, 2016a). This policy is effectively making private content, shared by others, public. In addition, the ‘tagging’ of friends (the identification of someone else in a post, status update or image), makes them visible to the wider social network of the user posting the content, connecting users whom would not otherwise have been connected (Altshuleret al*.*, 2013).

In this contested area of private vs public data some scholars have further distinguished between what data are ethical to use and what are not based on the type of data rather than the restrictions available. For example, user generated social media images (e.g. pictures or videos) are generally considered to be more personal and more sensitive than textual data (Humphreys, 2016). Although these images maybe publicly available there are additional issues to be considered which render them private, such as: individuals being identifiable; additional individuals may be included; or the research may be of a sensitive topic, e.g. mental health or sexual violence (Humphreys, 2016 ; Williams et al., 2017).

Given the complexity of the public/private nature of social media postings, debate continues around the use of social media data for research purposes. For Beninger (2017), this debate raises the issue of data ownership. She distinguishes between legal and moral ownership of data and argues that regardless of legal ownership, moral ownership of social media content is with the user. In her study on social media users’ views on the ethics of online research, some users were uneasy about their data being used for research and concerned that their data may be misrepresented and taken out of context. For these reasons, Beninger’s participants did not regard their data as public (Beninger, 2017). Studies on Twitter users’ views of their data use, conclude that the context of the original tweet is important for understanding user expectations of privacy (Fiesler and Proferes, 2018; Williamset al*.*, 2017).

Consideration around the private/ public aspects of data can assist both the researcher and institutional ethics committees to establish the level of ethical regulation required. However, where studies gather data via virtual social networks this becomes problematic. *A priori* definitions of data as public or private may be compromised because the data content may change through links to other interactions, evolving and recirculating in the virtual space and visual images may include non-research participants. Hence, a ‘situational ethics’ approach maybe most appropriate in this setting (Kosinskiet al., 2015).

### Informed consent

A further key ethical principle in social science research relates to informed consent. This relates to the provision of clear explanations of all aspects of study participation, enabling potential participants to fully understand the participant role, to place trust in the researcher and subsequently secure agreement for participation in the proposed study (Wileset al*.*, 2005). However, as Salmons (2017) points out, the traditional consent protocols based on conversations between researcher and participant (and which are usually signed before the onset of the study) are more difficult in practice in the digital context. Consent in this space is complicated by the blurring of boundaries and associated ownership of data (Beninger, 2017), linked to the earlier discussion around how social media content evolves and re-circulates (Halford, 2017). Further, building trust between the researcher and the participant - important for gaining consent - is also challenging in the virtual world (Salmons, 2017) and practically impossible in the case of ‘big data’ research (Halford, 2017). However, studies of users’ views on the use of their social media content suggest that some users want to be asked for consent (Fiesler and Proferes, 2018), and they expect anonymity when research is published (Williams et al., 2017). Salmons (2017) suggests that, similar to researchers in offline settings, researchers operating in online spaces need to build trust by conveying messages to potential participants that exhibit credibility and trustworthiness. In the context of data collection through virtual social networks this can translate into the creation of an online identity that demonstrates and displays both credibility and trustworthiness (Salmons, 2017). For example, if a potential participant undertakes an online search of the researcher and fails to discover any evidence of their existence or expertise this is likely to negatively affect their credibility. Credibility can be established through the creation of an online identity that links to the research, a relevant research institution and perhaps, an online blog, webinar or other studies undertaken by the researcher (Salmons, 2017). In this way potential participants who search for the researcher will find a wealth of information that conveys credibility and authenticity.

Social media researchers using online participant observation methods share some similar issues relating to informed consent as those undertaking other more traditional types of longitudinal studies. The need to regard ‘informed consent’ as an ongoing process which is continually revisited and contextualised throughout the study (Salmons, 2017) is a key aspect of research in the online space. Similarly, the BSA (2002) recommends that where the research is conducted over a period of time consent should be regarded as a process, which includes renegotiation. This is particularly important where research is over a prolonged period and participants may forget they are being studied as can easily be the case with online research (BSA 2002). For this reason it is important in online observational studies that consent is conintually revisited and renegoatiated.

### Anonymity

Ensuring anonymity is a particular challenge in relation to publication of research conducted in the online world (Dawson, 2014; Zimmer, 2010). This is challenging, since we know from studies of users that they are most comfortable with their content being used for research purposes when that content is anonymised (Fiesler and Proferes, 2018). Anonymisation procedures are frequently problematic when using social media data in any study as there is the potential for the publication of social media data to result in it being traced, revealing the identity of the user and exposing them to the risk of harm (Townsend and Wallace, 2016; Williams et al., 2016). This is particularly concerning when data is sensitive or users are from vulnerable groups.

However, one of the issues associated with maintaining anonymity online is the traceability of content, comments, statuses and quotes (Woodfieldet al*.*, 2013; Beninger, 2017). Search engines have become so effective that it is possible to trace much content back to its origin using a simple search function (Johnson et al., 2013; Roberts, 2015). This traceability highlights the potential for identifying both the online profile of the user and their offline identity (Zimmer, 2010). To prevent the identification of research participants to third parties, Robards and Lincoln (2017) made the decision not to quote directly from the research participants’ profile pages in their study. A further option is to paraphrase parts of quotes or posts that are obviously searchable, such as names, places and sometimes text (BPS, 2013). This type of content editing has been challenged as this has the potential to change the meaning of the original content, and affect the authenticity of the reporting (Johnson et al., 2013). The traceability debate is further complicated where both text and image-based data are collected, since images are less conducive to such alterations. Additionally, social media organisations are using images and facial recognition software to identify individuals (Facebook, 2016b), another factor mitigating against changes to images. Traceability is enhanced where metadata (e.g. location and author data) is stored in some images (Johnson et al., 2013), and where some social media platforms require content to be reproduced in its original format and attibuted to the username (Sloan and Quan-Haase, 2017; Townsend and Wallace, 2016).

We now move on to the empirical study, and explore the emerging ethical concerns.

## The research study and ethical approval

### Study overview

The ethical dilemmas and concerns discussed relate to an in-depth public health PhD qualitative study drawing on both social media data and group interviews, collected from mixed gender friendship groups over fourteen months in the North of England. Three friendship groups of young people (total 23 people, aged between 16-21 years) were recruited to the study through offline methods using a snowballing approach. These initial contacts subsequently functioned as gatekeepers to wider networks of young people, enabling friendship groups to be recruited to the study. The research was discussed with potential participants in face-to-face meetings, with study information provided, and consent was established. Consent was subsequently re-confirmed via social media after 24 hours. During the 14-month period of the study each group participated in three or four interviews, with each discussion lasting up to 75 minutes. Social media content from Twitter, Facebook and Instagram including visual images, hashtags, microblogs, on-line conversations and memes relating to drinking practices, competing practices and interconnecting practices was collected through online participant observation via the researcher linking to the participant’s own social networks.

Interpretation of the data involved the integration of all source data using NVivo software. Thematic analysis was used to identify and analyse patterns of meaning (Braun and Clarke, 2006) and analysis of visual images drew on the work of Banks (2008), who proposes the notion of an internal narrative focusing on content and an external narrative referencing the context of the image. The empirical findings of the study will be available in forthcoming papers.

This particular ethical journey started in a similar way to those of other researchers, with an anticipatory and pre-study review of ethical considerations. Following this, the study then became the subject of ethical scrutiny by the formally constituted institutional Faculty Ethics Committee (FEC). The FEC undoubtedly had considerable expertise and knowledge from which the study benefitted. However, new ethical issues arose within the study which related to the use of social media; these were either unforeseen during the pre-study review or were identified during the course of the study through wider reading. The contemporary nature of using social media as a data collection tool in research, combined with the continually changing nature of social media affordances, meant that the ethical challenges encountered were outside the practical experience of the FEC.

The following sections of this paper provide a reflection of the ethical problems related to the use of social media and from joining the participant’s social networks that arose during the study, addressing the four areas of ethical concern discussed in the literature review and exploring how these issues were addressed.

### Addressing the risk of harm

Particular safeguarding issues arose relating to the observations of the virtual everyday contexts of the drinking practices of participants. For example, it was deemed likely during the anticipatory review that data would show participants who were younger than the legal drinking age consuming alcohol, and there was a possibility that other illegal or illicit practices may also be exposed. While these concerns also feature within offline studies, the use of social media as a data collection platform has an added dimension of revealing activities which may show the participants and/or others at risk of harm in real time (Moreno et al*.*, 2013). This is particularly relevant with the advent of Facebook live, Twitter live and Instagram Live where users can stream live, unedited and unregulated video footage (Hern, 2017). The use of these online spaces may have resulted in the researchers observing these events as they are happening. As such the researcher, similar to other professionals working with young people, has a duty of care and responsibility to respond accordingly (BSA, 2002). The ethical challenge linked to the nature of an appropriate response in an online space has been recognised previously by other researchers (Johnson et al., 2013; Barnes et al., 2015). Barnes et al. (2015) collected Facebook status updates and made ethical decisions on a case-by-case basis, using the knowledge and experience gained by their lead researcher as a school teacher to support their decision making. Similar to Barnes et al. (2015) we chose to use this type of ethics-in-practice approach, assessing situations as they arose on an individual basis, the likelihood of harm happening, and gauging the seriousness and severity of any illicit activity. Participants were informed that there were limits to confidentiality; should the lead researcher consider there to be a likelihood of serious harm to themselves or others, then they would be obliged to report it. In making the decision whether significant harm was likely, we took into account who else was aware of the situation (such as venue staff, parents and the police), the support available to the young person at the time from friends and family, and whether the young person or someone else was in imminent danger. No incidents were deemed serious enough to require an intervention to secure personal safety.

A further ethical dilemma related to protecting the integrity of the data, while also protecting the online identities of gatekeepers and limiting the possibilities for harm to the lead researcher. Approval from the FEC was dependent on the creation of an online research project identity, considered to be an effective way of protecting the integrity of the data and reducing the likelihood of harm to both participants and the researcher. However, creating an online research project identity created both practical issues and additional ethical dilemmas. Practically and ethically, creating a unique research account for both Twitter and Instagram was easily accomplished, as both social media support the use of pseudonyms and the only requirement is a valid email address not currently linked to an alternative account. A credible online research project identity was therefore created on Twitter and Instagram using a university email address, and both accounts were closed on completion of the study.

Two options were investigated for setting up a project-based Facebook account. The first option was setting up a Facebook page similar to that of a business or company. However, this was not found to be feasible as there was no facility for accessing participants’ status updates or newsfeed from the page. The second option was setting up a personal profile using a fake account. This had the outcome of meeting the FEC requirements, but raised the ethical dilemma of contravening Facebook rules as using a fake account contravenes Facebook policy (Facebook, 2016c). The use of a fake identity also raised the ethical dilemma of potentially misleading the research participants as to the identity of the researcher (Carteret al., 2016), and thus not exhibiting the credibility and trustworthiness that Salmons (2017) advocates for creating a research relationship.

Hence, although we agree that covert research methods should not necessarily be dismissed as being intrinsically unethical (Calvey, 2017), in this study building a relationship and rapport with the participants was important to the success of the project, and the creation of a false identity (one potentially uncovered by something as simple as a brief look at the University website pages) may well have reduced the credibility of the researcher. Furthermore, Elovici et al. (2014) argue that fake accounts create unnecessary costs for the provider, as they police this policy, and even have the potential to lead to a prosecution. It was this policing that ended the deception as an automatic alarm was triggered when connecting with participants and Facebook froze the fake research profile account. In this instance, we explained to the participants why a false research identity had initially been set up which did not negatively impact on the study.

The remaining option was to use a personal Facebook account linked to the lead researcher. However, this led to further ethical challenges; research data could be compromised and the identity of the researcher’s friends, family and ‘gate keepers’ for the research project could be exposed, as they would become visible to the research participants (a potential outcome of users becoming connected via shared links with the researcher as a third party) (Eloviciet al*.*, 2014). This potentially fuels the blurring of the private/public boundaries in the research setting (Johnson et al., 2013), and was complicated by the fact that the gatekeepers were bound by professional codes of conduct that governed their use of social media. Gatekeepers had requested not to be linked to the young people they had introduced to the project. Although the ethical principle of protecting the private lives of gatekeepers within research is not necessarily different from other types of qualitative research, new and different safeguards to protect their online information were needed. We proposed to solve this ethical challenge by ‘unfriending’ all existing Facebook friends (for the lead researcher) and only accepting research participants as friends. However, this uncovered a further ethical complication, with the emerging possibility that research participants would have access to researchers historical data thus again compromising the privacy of gatekeepers and the researcher’s friends and family (Facebook, 2016a). The chosen option, on balance, was for the researcher to close their own personal Facebook profile, open a new profile using the researcher’s real identity but including details of the fake research identity as a ‘nickname’ and only accepting participants as ‘friends’.

### Addressing the public vs private data dichotomy

In this study, the participants had given informed consent to their tweets, Facebook content and Instagram posts all being used as data. However, the issues raised by Halford (2017) relating to the evolving nature and recirculation of social media data did raise ethical concerns within this study, particularly evident with Facebook content.

The sharing nature of Facebook meant the project had access to data that was continually evolving as a result of additional content (e.g. comments, likes, tagging and shares) posted by the participants’ wider social networks (Barnes et al., 2015; Robards, 2013). This data potentially gave a deeper insight into the participants’ own postings but also posed an ethical challenge around its use.

According to the literature, this data could be considered ‘private’ as it was posted on a social media platform that has restricted permissions. However, as the researcher could see the content, these users had not made use of the highest privacy restrictions and had not restricted posts to be viewed only by ‘friends’. So the question we faced when examining user posts was “is this a ‘public’ or ‘private’ posting”? Facebook’s own terms and conditions state that each user owns their own content and has control over how that content is shared. Although content posted by participants’ friends and then shared by the participants is ‘owned’ by the user sharing it according to Facebook terms of use. However, the question still remains how ethical is it to use this data? In addition, the evolution and recirculation of the content was deemed important to the study, since it provided information on the performativity of alcohol consumption practices online.

We decided to only use content from non-participants as data when that content directly related to the research participants’ online performance of drinking or competing practices, and when the identity of that non-participant could be protected. This position is similar to the ethical conclusions reached by Kosinski et al. (2015), who propose that the use of data generated or reference to content from non-participants is acceptable, if analysis is aimed exclusively at study participants (Kosinskiet al., 2015: p553).

### Addressing informed consent

Two ethical challenges relating to consent and the use of social media were highlighted in the study: (1) negotiation of informed consent, and (2) informed consent and data ownership.

 ***Negotiation of informed consent***

Similar to other researchers who wish to gain institutional ethical approval, the researcher followed internal institutional guidance for a formal, paper-based consent procedure, which included providing accessible information and consent materials for potential participants. This process was agreed by the FEC who also noted that young people in a group situation may consent to participate in a research study, even if they do not wish to participate, as research is affected by its particular social context. Consequently, consent was given at an initial face-to-face meeting with participants, and then followed up by individual confirmation of consent within twenty-four hours. All the young people recruited to the study were given the researchers’ contact details, and encouraged to make contact with any concerns they had, or if they wanted to withdraw from the study for any reason. Since participants were recruited to the study through face-to-face contacts, gaining informed consent took the form of more traditional consent protocols with the researcher able to build rapport and trust during this time. Corroboration of consent was accomplished with the use of social media within the agreed timescale, with no participants declining.

While the FEC did not require adherence to the BSA recommendation that research over a prolonged period of time consent should be regarded as a process, which includes renegotiation, the researcher adopted this approach. At the beginning of each group interview with the research participants, we revisited the study aims and our expectations around participant contribution and involvement. The participants were then asked to confirm their continuing involvement in the research. Revisiting their consent in this way was of particular importance as the virtual lives of the young people were observed on an on-going basis via social media platforms. The importance of renegotiating consent became apparent when the researcher contacted one of the participants to arrange a meeting, raising the issue of the participant having ‘fallen out’ with another study participant, and to enquire how they wished to proceed with the study. The participant was confused, and had to be reminded that she had posted about the disagreement on Facebook. The participant had either forgotten the information had been posted or alternatively had not remembered that the researcher was part of her online social network. Either way it demonstrates how easy it can be for participants to forget they are participating in an online research study.

None of the young people in the study withdrew their consent to participate in the study. Rather those who no longer wished to continue simply chose not to turn up at the pre-arranged times, or did not commit to arranging a meeting time for the interviews. The latter circumstance was the case with a complete friendship group of seven young people who did not commit to a further meeting; this was despite meeting with the researcher, signing consent forms, re-consenting twenty-four hours later and accepting the researchers ‘friend requests’ on social media. Their silence was interpreted as an indication of withdrawal of their consent to participate in the study, and that meaningful informed consent had not been given (Bourke and Loveridge, 2014). However, young people who had given informed consent and then participated in at least one friendship group discussion (before absenting themselves from later discussion groups) were considered to have consented to participate in the study. This assumption was confirmed through the social media conversations when reasons for not attending interviews included leaving the college where the discussion groups were held, and no longer being in friendships with other participants. The combination of online data collection with more traditional data collection methods proved a useful combination in the gaining (or otherwise) of consent and its renegotiation over time.

 ***Informed consent and data ownership***

The issue of informed consent and data ownership was also relevant to this study as participant’s social media content interacted, evolved and rec-circulated within their social networks. The researcher was inundated with an abundance of data not only generated by the research participant, but as a result of their interactions. The previous discussions in the literature review section lead us to ask whether it can be deemed public or private, who owns the data and also to assess the potential for embarrassment or risk of damage to an individual’s reputation. In this case, the data could be considered to be public - as the user had not restricted access to it - although when shared on Facebook the affordances for higher privacy settings are available thus making it private. Also although legal ownership of the data is no longer with the original user we agreed moral ownership is, and we therefore decided it was only appropriate to use non-participants’ data if it related to participant’s own practices, could be anonymised and did not damage any user’s reputation or be likely to cause embarrassment.

### Addressing anonymity

As was noted previously, a commitment to anonymity and privacy for all participants is a key principle within social science research guidance (BSA, 2002; SRA, 2003). Although we knew the offline identities of the participants, the question of how to protect their anonymity was raised during the course of the study.

In an attempt to prevent a breach of anonymity whilst maintaining authenticity and meaning in this study, the following actions were undertaken. First, all names, places, faces and other obvious searchable terms in any content used in the reporting of this study were blanked or pixelated, and images were checked for metadata (BPS, 2013). This action was intended to preserve anonymity whilst not changing meaning or affecting authenticity. Second, the study participants were informed about the risk of being traced from any quotes or data used and participants were asked to consent to the use of their social media content in the reporting of the research (Dawson, 2014). Random searches of the edited quotes that were planned to be used were conducted using two different search engines and the users were not subsequently identified (Barnes et al., 2015). In addition, a decision had to be made about the use of tweets in published material as the reproduction of participant’s tweets alongside their user names would effectively expose their offline identities. For this reason, the decision was made to not reproduce tweets in published material. However, the changing nature of online technologies raises the possibility that published data may be traceable in the future. In addition, in our study research participants who did not necessarily know each other became visible to each other and therefore their anonymity could not completely be preserved. For these reasons, both Dawson (2014) and Beaulieu & Estalella (2012) suggest promises of complete anonymity to research participants cannot (and should not) be made, since this is impossible to deliver on

## Discussion and implications for future social media research

This paper has reflected on the ethical challenges associated with one study that used social media data namely a longitudinal online participant observational study. In doing so we responded to calls for researchers to participate in conversations about the ethical dilemmas encountered when using social media data in research (Kosinski et al., 2015; Johnson et al., 2013). Furthermore, we contribute to the development of social media research and the evolution of ethical principles for both researchers and ethics committees on the use of social media data in social research.

The use of new technologies in this participant observational study has introduced some distinctive problems, such as determining what social media data is ethical to use and what is not, and some old problems in a new form, including preserving anonymity, safeguarding from harm in a world of live streaming and (re)negotiating ongoing consent. Thus, some thought needs to be given to how research in online spaces evolves and develops, specifically to enable researchers to make ethical decisions during the course of their research, and to help ethics committees as they scrutinise the ethical decision-making of researchers. Ethical guidelines as they relate to more traditional research cannot always simply be transferred to researching social media without careful examination and sometimes adaptation (Woodfieldet al., 2013), for example protecting the anonymity of research participants.

Recognising that social media and online technologies are constantly evolving and changing is important for any new study, and researchers and ethics committee members alike need to be vigilant to these changes and respond to new ethical challenges as they arise (Beaulieu and Estalella, 2012). This experience has however highlighted that researchers and ethics committees cannot always anticipate all the ethical challenges that may arise prior to beginning an online research project. The continually changing nature and affordances of social media mean that ethics committee and researchers often lack the requisite knowledge or competences needed to make contextually informed ethical judgements during anticipatory review (Johnson et al., 2013; Woodfield et al., 2013). The discussions around the four interrelated areas of ethical concern addressed lead to some practical suggestions for future researchers and institutional ethics committee to consider when designing, assessing and undertaking studies that collect data within the online environment.

First, we suggest that ethics committees promote an ongoing dialogic and reflexive process with researchers when considering any research studies using social media data. This could be supported by developing the expertise of researchers, PhD students and committee members who may need to be better equipped to think more deeply about the ethical implications of studies involving social media data. For example, what data maybe ethical to use and what data may not be ethical to use. Second, there is a need to record, share and publish knowledge, experience and insights into the challenges and solutions that are facing researchers in this context to enable future researchers to learn from previous challenges and to contribute to discussion and debate on how to tackle ethical issues using social media data. In addition to publication in peer reviewed journals this needs to be in a more responsive format. This could be done by developing case studies which could be shared within university communities, with social media users and to the wider academic community through online blogs, podcasts, training workshops and searchable databases. This would enable researchers to share their experiences in a speedier and more timely manner; critical in a rapidly changing research context.

Third, in a virtual world that is complex and evolving, guaranteeing anonymity (unless a commitment is made to not publishing content) would seem to be virtually impossible. Researchers should therefore acknowledge this, and be open and honest with research participants about the possibility of being identified whilst doing everything possible to maintain anonymity where appropriate and to reduce the risk of harm.

Finally, we advocate that researchers give detailed consideration to the handling of consent in fluid contexts, including how participants are reminded of consent issues over the course of the project.

**Conclusion**

A number of conclusions have been drawn, but an overarching concern is that research studies undertaken within an online context is multifaceted, is continually changing, and is often outside the traditional expertise of institutional ethics committee, requiring ethical processes that are flexible and adaptable. These should focus on ethical principles rather than ethical regulation, enabling a move to a more situational/reflexive ethics approach. As more of our lives are lived in and through on-line spaces, research in this context is becoming more relevant, important and popular. Therefore it is important that the research community engage in robust debate about the particular challenges of undertaking ethical research when using social media data. Finally, despite the ethical challenges facing researchers in the online context we would not wish to deter researchers seeking to investigate people’s experience through social media as this constitutes a contemporary social space within which people live many aspects of their lives, and therefore is an appropriate and data rich context for research.

**Acknowledgments**

We are grateful to the all the young people who participated in this research and graciously shared their social media content.

**Funding**

LiLac - Lancaster and Liverpool Universities Collaboration for Public Health Research.

**References**

Altshuler, Y., Elovici, Y., Cremers, A., Aharony, N. and Pentland, A. (2013) *Security and Privacy in Social Networks.* New York: Springer.

Association of Internet Research (2012) *Ethical decision-making and Internet Research*. Association of Internet Researchers. Available at: <http://www.aoir.org/reports/ethics2.pdf> (Accessed: 03/03/2018.

Banks, M. (2008) *Using visual data in qualitative research.* London: Sage.

Barnes, N., Penn-Edwards, S. and Sim, C. (2015) 'A dialogic about using Facebook status updates for education research: a PhD student's journey', *Educational Research and Evaluation,* 21(2), pp. 109-121.

Beaulieu, A. and Estalella, A. (2012) 'Rethinking research ethics for mediated settings', *Information, Communication & Society,* 15(1), pp. 23-42.

Beninger, K. (2017) 'Social Media Users’ Views on the Ethics of Social Media Research', in Sloan, L. and Quan-Haase, A. (eds.) *The SAGE Handbook of Social Media Research Methods*. London: Sage, pp. 57-73.

Bourke, R. and Loveridge, J. (2014) 'Exploring informed consent and dissent through children's participation in educational research', *International Journal of Research & Method in Education,* 37(2), pp. 151-165.

Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology,* 3(2), pp. 77-101.

British Psychological Society (2013) 'Ethics Guidelines for Internet-mediated Research', Available: BPS (Accessed 23/02/2017).

British Sociological Association (2002) 'Statement of ethical practice', Available: BSA (Accessed 23/02/2016).

Brown, B., Weilenmann, A., McMillan, D. and Lampinen, A. 'Five provocations for ethical HCI research'. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*: ACM, 852-863.

Calvey, D. (2017) *Covert Research: The Art, Politics and Ethics of Undercover Fieldwork.* London, Thousand Oaks, New Delhi, Singapore: Sage.

Carter, C., Koene, A., Perez, E., Statache, R., Adolphs, S., O'Malley, C., Rodden, T. and McAuley, D. (2016) 'Understanding academic attitudes towards the ethical challenges posed by social media research', *ACM SIGCAS Computers and Society,* 45(3), pp. 202-210.

Dawson, P. (2014) 'Our anonymous online research participants are not always anonymous: Is this a problem?', *British Journal of Educational Technology,* 45(3), pp. 428-437.

Dingwall, R. (2008) 'The ethical case against ethical regulation in humanities and social science research', *Twenty-First Century Society,* 3(1), pp. 1-12.

Elovici, Y., Fire, M., Herzberg, A. and Shulman, H. (2014) 'Ethical considerations when employing fake identities in online social networks for research', *Science and engineering ethics,* 20(4), pp. 1027-1043.

European Commission (2010) 'Guidance Note for researchers and evaluators of Social Sciences and Humanities Research (draft)', (Accessed 04/05/2017).

Facebook (2016a) *Data policy*. Available at: <https://www.facebook.com/about/privacy/> (Accessed: 07/03/2016.

Facebook (2016b) *How does Facebook suggest tags*. Available at: <https://www.facebook.com/help/122175507864081> (Accessed: 07/03/2016.

Facebook (2016c) *What names are allowed on Facebook*. Available at: <https://www.facebook.com/help/112146705538576> (Accessed: 08/03/2016.

Fiesler, C. and Proferes, N. (2018) '“Participant” Perceptions of Twitter Research Ethics', *Social Media+ Society,* 4(1), pp. 1-14.

Halford, S. (2017) 'The Ethical Disruptions of Social Media Data: Tales from the Field', in Woodfield , K. (ed.) *The Ethics of Online Research*. UK: Emerald Publishing Limited, pp. 13-25.

Hern, A. (2017) 'Facebook live is changing the world, but not in the way we had hoped', *The Guardian*. Available at: <https://www.theguardian.com/technology/2017/jan/05/facebook-live-social-media-live-streaming> (Accessed: 11/08/2017).

Humphreys, A. (2016) *Social media: Enduring principles.* Oxford: Oxford University Press.

Johnson, N., Henderson, M. and Auld, G. (2013) 'Silences of ethical practice: Dilemmas for researchers using social media', *Educational research and evaluation,* 19(6), pp. 546-560.

Kosinski, M., Matz, S., Gosling, S., Popov, V. and Stillwell, D. (2015) 'Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines', *American Psychologist,* 70(6), pp. 543.

Lafferty, N. and Manca, A. (2015) 'Perspectives on social media in and as research: A synthetic review', *International Review of Psychiatry,* 27(2), pp. 85-96.

Lange, P. (2007) 'Publicly private and privately public: Social networking on YouTube', *Journal of computer-mediated communication,* 13(1), pp. 361-380.

Miller, T. (2012) 'Reconfiguring research relationships: regulation, new technologies and doing ethical research', *Ethics in qualitative research*, pp. 29-42.

Moreno, M., Kota, R., Schoohs, S. and Whitehill, J. (2013) 'The Facebook influence model: A concept mapping approach', *Cyberpsychology, Behavior, and Social Networking,* 16(7), pp. 504-511.

Robards, B. (2013) 'Friending Participants: Managing the Researcher–Participant Relationship on Social Network Sites', *Young,* 21(3), pp. 217-235.

Robards, B. and Lincoln, S. (2017) 'Uncovering longitudinal life narratives: scrolling back on Facebook', *Qualitative Research,* 17(6), pp. 715-730.

Roberts, L. (2015) 'Ethical issues in conducting qualitative research in online communities', *Qualitative Research in Psychology,* 12(3), pp. 314-325.

Salmons, J. (2017) 'Getting to Yes: Informed Consent in Qualitative Social Media Research', in Woodfield, K. (ed.) *The Ethics of Online Research*. UK: Emerald Publishing Limited, pp. 109-134.

Sloan, L. and Quan-Haase, A. (2017) 'A retrospective on state of the art social media research methods: Ethical decisions, big-small data rivalries and the spectre of the 6Vs', in Sloan, L. and Quan-Haase, A. (eds.) *The SAGE handbook of social media research methods*. London: SAGE Publications, pp. 663-672.

Social Research Association (2003) 'Ethical guidelines', (Accessed 23/02/2016).

Townsend, L. and Wallace, C. (2016) 'Social media research: A guide to ethics', *University of Aberdeen*, pp. 1-16.

Webster, S., Lewis, J. and Brown, A. (2013) 'Ethical considerations in qualitative research', in Ritchie, J., Lewis, J., McNaughton-Nicholls, C. and Ormston, R. (eds.) *Qualitative research practice: A guide for social science students and researchers*. London: Sage.

Wiles, R., Heath, S., Crow, G. and Charles, V. (2005) 'Informed consent in social research: A literature review', *NCRM Methods Review Papers NCRM,* 1.

Williams, M., Burnap, P. and Sloan, L. (2017) 'Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users’ Views, Online Context and Algorithmic Estimation', *Sociology,* 51(6), pp. 1149-1168.

Williams, M., Burnap, P., Sloan, L. and Jessop, C. (2016) *Using Social Media Data  in International Development Research, Monitoring & Evaluation*: NatCen Social Research & Social Data Science Lab.

Williams, M., Burnap, P., Sloan, L., Jessop, C. and Lepps, H. (2017) 'Users’ Views of Ethics in Social Media Research: Informed Consent, Anonymity, and Harm', in Woodfield, K. (ed.) *The Ethics of Online Research*: Emerald Publishing Limited, pp. 27-52.

Woodfield, K. and Iphofen, R. (2017) 'Introduction to Volume 2: The Ethics of Online Research', *The Ethics of Online Research*: Emerald Publishing Limited, pp. 1-12.

Woodfield, K., Morrell, G., Metzler, K., Blank, G., Salmons, J., Finnegan, J. and Lucraft, M. (2013) 'Blurring the Boundaries? New social media, new social research: Developing a network to explore the issues faced by researchers negotiating the new research landscape of online social media platforms', Available: NCRM (Accessed 09/03/2016).

Zimmer, M. (2010) 'But the data is already public: On the ethics of research in Facebook', *Ethics and information technology,* 12(4), pp. 313-325.