

Title

From “a fair game” to “a form of covert research”: Research Ethics Committee Members differing notions of consent and potential risk to participants within social media research

Abstract

Social media (SM) research presents new challenges for research ethics committees (RECs) who must balance familiar ethical principles with new notions of public availability. This paper qualitatively examines how UK REC members view this balance in terms of risk and consent. While it found significant variance overall, there were discernible experience-based trends. REC members with less experience of reviewing SM held inflexible notions of consent and risk that could be categorised as either relying on traditional notions of requiring direct consent, or viewing publically available data as ‘fair game’. More experienced REC members took a more nuanced approach to data use and consent. We conclude that the more nuanced approach should be best practice during ethical review of SM research.

Introduction

Social media (SM) provides a new methodological research tool allowing researchers to draw on data from websites such as facebook, twitter and youtube, as well as password-protected and non-password-protected chatrooms and forums (Kaplan & Haenlein, 2010). The benefits of this research centres on the huge repository of easily accessible, seemingly public data, available at low cost. However, different views on the publically available nature of SM platforms and the perceived privacy of SM users presents an ethically grey area for Research Ethics Committees (RECs) charged with ensuring that research using SM data is conducted responsibly and ethically¹. Ordinarily, researchers' responsibilities to their research participants are outlined in a range of disciplinary codes of conduct (BERA: Jones, 2011; BPS, 2010) and in the off line context these responsibilities have clear boundaries that are familiar to researchers and RECs alike (Hedgecoe, 2016). However, in the context of SM research, new challenges that obscure the "fundamental rights of human dignity, autonomy, protection, safety, maximization of benefits and minimization of harms" (Markham & Buchanan, 2012; p.4) are presented that render familiar ethical principles considerably less so. As such, when ostensibly publically available data is collected without SM users' awareness and/or permission (as can be the case with research involving powerful or marginalised groups (Hammersley, 2009)), [privacy violations based on improper access that have been identified as a concern by certain authors \(Zimmer, 2010\)](#), [privacy violations based on improper access \(Zimmer, 2010\)](#) can take place that would have been unlikely to occur through research lacking a SM component.

The issue of consent is particularly problematic in the case of SM research. First, consent may not always be possible or practical. Second, whilst consent is an ethical and legal requirement for any human participant research (HPR) (WMA, 2001), a number of scholars

reject that SM data fulfils the definition of HPR, with this remaining an on-going live issue in the literature (Henderson et al, 2013; Keim-Malpass et al, 2013; Hutton & Henderson, 2015; Metcalf and Crawford, 2016). Third, the static, compulsory consent that takes place at the beginning of the research process has been questioned by some SM researchers, who prefer the use of changeable constructs of consent based upon the perceived risk to SM users as research participants (Hutton and Henderson, 2015). And finally, some scholars argue that terms and conditions consent - where SM users agree to a platform's data use policy prior to creating an account – are problematic in acting as a proxy for more rigorous attempts to seek permission from SM users for research (Morrison, McMillan & Chalmers, 2014). This latter issue has come to the fore in a number of high profile cases, including a joint Facebook-Cornell University study in 2014 where empirical research on behaviour and mood in response to variations in emotionally charged content viewed by users was published without the Facebook users' knowledge (Adams et al, 2014). As suggested by Zimmer (2010), when data is “considered freely accessible for collection and research, regardless of what the subject might have intended or desired” (p. 322), [-privacy violations through the improper access identified by Zimmer \(2010\) as a potential concern.](#) ~~privacy violations based on improper access can easily occur.~~ This issue is only likely to grow alongside the rising profile of big data research and the lack of sufficient regulation regarding ethical best-practice for SM research (Vayena et al, 2015).

Such challenges have not been entirely resolved by the guidelines that currently exist to steer best practice (for example, see (Ess, 2002; Markham & Buchanan, 2012)), and the uncertainty surrounding ethical decision-making for SM research is further reinforced by the bottom-up, researcher-led approach that has been advocated by the Association of Internet Research (AoIR: Ess, 2002) and others (Markham & Buchanan, 2012).

The complexity involved with the interplay between privacy, user awareness and consent as well as the lack of a standardised approach to SM research means there is little to guide RECs in their decision-making. Indeed, preliminary evidence from the US (Buchanan & Hvizdak, 2009), and in the form of a survey from a UK university (Carter et al, 2016) suggests that RECs (and their Institutional Review board counterparts) are struggling to adapt to these unique challenges during their review of such research (Henderson et al, 2013; Boyd & Crawford, 2011).

This paper builds on this research, and provides a deeper exploration of how REC members from a number of different UK institutions balance the benefits of SM research with the potential harm to SM users during their review of such research. In particular, this paper examines how the notions of risk and consent in SM research are considered by REC members. Below we provide a brief review of the differing notions of consent that need to be considered for SM research, thereby highlighting the potential challenges facing REC members assessing research ethics applications involving SM data. We then present our study, which involved interviews with 19 REC members. Our findings explore the level of SM experience of our REC member interviewees and how this influences their conceptualisation of consent and risk of SM ethics applications. Our discussion will contextualise our findings within the current developments of SM research and provide recommendations of how to implement a broader range of protections for SM users as participants within research. Though we note that as an in-depth, small-scale pilot study, it is not possible to generalise these findings across all contexts; the broad conclusions that we draw are designed to anchor further explorations of the ethics of SM research in data rich preliminary work.

Literature review

There is a significant variation of opinion in terms of when consent is needed in SM research due to a lack of consensus about a whole range of issues. Some researchers argue that consent is needed as SM data represents human participant research (HPR); others reject this proposition, rather viewing SM data as 'published text' that precludes the permission of the SM user for data use. Beliefs about consent are also influenced by whether researchers view SM data as residing within the public domain (and therefore freely available for any researcher to use) or as private data where SM user expectations of any 'perceived privacy' about their data need to be considered during the research process (Markham & Buchanan, 2012; Moreno et al, 2013; Samuel, 2017). The type of research being conducted (quantitative versus qualitative datasets) and the discipline within which it is occurring can also play an important role in how researchers view the necessity to ask SM users to source their data (Paechter, 2013).

This lack of consensus results in RECs struggling to know how best to apply familiar ethical principles to SM research that blurs boundaries and provides a mix of different ideas about research ([Buchanan & Hvizdak, 2009](#); [Boyd & Crawford](#); [Henderson et al, 2013](#); [Carter et al, 2016](#)). Indeed, whilst the BPS (2007) has identified the 1) level of identifiability and 2) level of observation as two key dimensions affecting ethical consideration, within this framework, while consent is something that is required for identifiable and recruited participants, when subjects are not identifiable and/or unaware of being researched, consent becomes an uncertain requirement (BPS, 2007).

Without clear consensus about under what conditions consent for SM data is required, it is uncertain whether the use of SM data should be treated as being publically available for research purposes. It is similarly uncertain what the risks to participants may be when consent is not considered to be something that is required. In these situations, researchers typically rely on de-identifying their SM data as a standard requirement. However, recent research has found that even large-scale aggregated data can be easily re-identified upon release (Nils Homer et al, 2008; Gymrek et al, 2013; Narayanan & Felton, 2014), calling into question the extent to which de-identification can be considered a sufficient protection when dealing with SM data.

Furthermore, there is a lack of clarity in terms of the extent to which SM users are cognisant of the uses of their data, with only a few recent studies examining user perceptions of SM research. Such studies have found that users want more privacy regardless of the public nature of data (Williams and Burnap, 2017) and overall they forward a view of research utilising SM data that upholds the ethical principles of consent, anonymity and avoiding undue harm to participants (Beringer, 2014). Beringer (2014) reported that users positioned the need to improve the representativeness of findings and understand the privacy risks of platforms as being paramount to protection and trust of participants. Transparency in reporting results, while protecting the identity and reputation of participants, was central in maintaining their trust and continued support of research.

In addition, recent studies have documented a 'privacy paradox' (Barnes, 2006) where SM users' self-declared desire for more privacy online conflicts with their actual behaviour disclosing personal information on websites and mobile apps. It has been suggested that this

contradictory behaviour may be the result of a number of factors ranging from a lack of understanding and knowledge of risk and privacy-protective behaviours (Acquisti & Gross, 2006; Hargittai & Litt, 2013; Park 2013), to apathy and learned helplessness in the face of the perceived privacy that SM users haven't actually lost, but rather they never had in the first place (Hargitta & Marwick, 2016). In addition, the need to accept consent on SM platforms as a 'take it or leave it' offer (Custers 2016, 2013) is likely to further exacerbate such feelings of apathy and learnt helplessness. Some scholars therefore suggest that consent processes within SM are inadequate, and do not address the cognitive problems undermining privacy self-management in terms of SM users making decisions about privacy only selectively and on an often uninformed basis (Solove, 2013; Custers, 2013).

In response to the complexities of ethical decision-making during SM research, Vayena et al (2015) have suggested that *"the current regulatory framework emphasizes practices, such as obtaining informed consent and balancing the benefits of research against the risks of participation, that are out of place in non-clinical research"* (p.426). They, and others argue that while consent can happen in a number of different ways (ranging from lengthy consent forms to less onerous approaches such as opt-out or terms and conditions consent) it is unclear which is the best way for SM research relative to the practical and pragmatic difficulties presented in seeking consent from subjects that are virtually, and sometimes geographically distant from researchers. In addition, they also argue consent as a cornerstone of HPR is often inadequate due to consent forms that are "lengthy, complex and difficult to understand"; disclosures that "often do not inform subjects of all potential data uses and the harms that could result from misuse"; and limited opportunities for subjects to "withhold, revoke, or modify consent" (Vayena et al, 2015; p. 432). As a consequence, they say, obtaining consent for SM research can amount to inadequate approaches to subject

permission as per the norms of traditional approaches to participant consent and potential risk.

Others similarly argue that it is difficult to formulate simple general rules in SM research that rather needs to be an “inductive process” (Henderson et al, 2013; p.5; McKee and Porter, 2009). Indeed, it is suggested that there is “*no one-size-fits-all approach to consent that achieves optimal results*”, and that while low burden ‘secured consent’ where permission is given at a single point in time may be sufficient in some contexts, more sustained approaches where permission is continually probed may be required for other participants in different research contexts (Hutton & Henderson, 2015; p.186). In addition, while ethical norms also include SM users’ rights not to be researched, this is not always practical or realistic, particularly in light of the seemingly public availability of SM data and arguments that users should be cognisant and aware. Vulnerable populations provide an additional level of complexity to this argument, where participants’ understanding of the publicness of ostensibly private spaces online may be compromised (Hendersen et al, 2013).

Vayena et al (2015) have therefore called for a new ethical framework for research utilising SM, particularly in relation to big data, which broadens the concepts of protection and consent, and suggests researchers should be selecting from the “*wide range of procedural, economic, legal, educational, and technical protections that are available*” (p. 437). In other words, other protections beyond direct consent and de-identification need to be considered to ensure that participants are adequately protected in the context of SM research. As with any new development, standards will not emerge overnight. In the meantime then, REC members need to balance these tensions regarding risk and consent when considering ethics

applications for research proposals that include a SM component. This paper will go some way towards illuminating the considerations that RECs are faced with, when reviewing SM research applications.

Methods

Recruitment

Websites of the 20 most research-intensive UK universities (as determined by the UK Research Excellence Framework, 2014) were searched in order to identify relevant contact details of Chairs and/or members of their university-level and departmental/faculty-level Research Ethics Committees (RECs). Where contact details were missing or incomplete, emails requesting these details were sent to relevant information email addresses provided. Specific interest was placed in identifying those RECs who were responsible for reviewing research which adopted social media approaches in health. As such, RECs solely within physical or life sciences were excluded.

Where contact details were identified, emails requesting participation in the project were sent to the Chair of each REC, as well as to REC members whose areas of expertise was health and/or social science (i.e., excluding REC members with expertise solely in clinical, life or physical sciences, as determined by their webpages). Participant information sheets were attached to all requesting emails. Two follow-up emails were sent to non-responders.

In total, 63 individuals were contacted to request participation in the project. Nineteen individuals responded and interviews were arranged. Participants represented 13 UK institutions across 18 different university-level or faculty/department-level (social sciences,

humanities, medical sciences, psychology, or health sciences) RECs. They included nine REC Chairs, one deputy Chair and nine REC members. Ten participants sat on a university-level REC, and 13 participants sat on a departmental/faculty-level REC (four participants sat on both). On questioning about their experience with the ethical review of research which uses social media data (health or otherwise), four stated no such experience (though one of these interviewees had some experience using social media data within their own research); eight noted limited experience; and seven stated having experience.

Interviews

Interviews were conducted by GS either face-to-face, over the telephone or via skype. Interviews lasted between 40-60 minutes, and were digitally recorded. The interview schedule was broad, asking participants about their own experience of using social media data for research, as well as their general use of social media more generally in their professional or personal life. The interview schedule also explored interviewee's views about the ethical issues surrounding the use of social media data for research (in particular to health research); their views about whether such research should require ethical approval, and knowledge about the policies at their own institution in relation to this; their experiences of reviewing such research in a REC capacity and their decision-making in relation to this; any guidelines, training or literature they had used in order to aid their decision-making in this area; and, for those with no experience in this area of ethical review, how interviewees thought they would make decisions about this research in their capacity as a REC Chair/member.

Analysis

Analysis of interview data was approached using inductive reasoning employing the inductive approach of grounded theory (Strauss, 1987; Charmaz, 2006). The analysis (or coding) of data was based on two inter-linked rounds: overview analysis and detailed analysis (Strauss, 1987). Overview analysis consisted of memo-making and broad coding. Extensive memo-making was employed by the interviewer directly after each interview. Broad coding proceeded by scanning the interview transcripts for relevant ideas and themes. This phase was conducted in duplicate by first and second authors, before one theme per author was selected for detailed analysis on the basis of extensive team discussion.

SM experience of participants

Participants' experience was classified into 3 levels; (A) experienced; (B) limited experience; and (C) inexperienced. Levels of experience were determined during the interviews with the participants self-categorising and/or making direct reference to their own perceived level of experience. Participants were noted to have direct experience if they had conducted and/or reviewed more than one SM study (~~n=10~~, columns A-~~BB~~ in Table 1).

Participants were categorised as having limited experience if their practice of conducting and reviewing SM studies had been indirect (through other collaborators or students), or limited to a single study. In addition, if participants owned and used a social media account either personally and/or professionally, this was also counted as indirect experience (~~columns #=~~ 22, columns C-DC-E in Table 1).

On the other hand, inexperience consisted of having no experience at all with conducting or reviewing SM studies, and not owning a SM account for personal or professional use (~~n=2~~, column EF in Table 1).

- TABLE 1 HERE -

For the remainder of the paper, when we talk about “experienced” REC members we are referring to those occupying categories A-B; when ~~D~~ were classified as either being “experienced” or we talk about REC members having “limited experience” we are referring to those within categories C-D; and when we talk about the, ~~and a~~ minority of participants exclusively in category E were labelled as entirely ~~as entirely~~ “non-experienced”, we are referring to those within category E¹.

Results

Do no harm in practice: Context, user awareness and de-identification

REC members centred their ethical questioning of SM research applications on the principle of ‘do no harm’. This involved engaging with *“a whole trench of really familiar ethics questions”*; and *“standard worries around confidentiality and consent, and whether privacy has been respected”* (REC 14). In practice, ‘do no harm’ generally required all research participants to be adequately informed about the project, to have consented to the research they are participating in, and to have had their identities protected. These traditional ethical norms (Vayena et al, 2015) were relatively unproblematic when a non-exceptionalist

approach was taken (Gelinas et al, 2017) and SM research proposals had already written the need for SM user consent into the research methodology. However, all REC members were aware that when explicit consent had not been sought from SM users (as is often the case for SM research: see Henderson, Johnson and Auld, 2013), applying the ‘do no harm’ principle was complicated by the presence of SM data being seemingly freely and publicly available. In these instances REC members had a well-developed awareness that SM users could not be assumed to be aware their data was being used for research purposes (Henderson et al, 2013; Moreno et al, 2013; McKee & Porter, 2009). This idea of SM user ‘perceived privacy’ raised the question of whether any research proposal that did not ask for SM user consent was ethically problematic.

To address this, REC members drew on other strategies to help inform their ethical judgments when reviewing SM research. These included looking at the context of the research (i.e., the specific topic under investigation) and “*what actually are the questions that are being asked from the data that you're collecting on people*” (REC 7). More ‘benign’ subject matter was generally required as requiring less scrutiny in terms of consent, because of a perceived decrease in risk to SM participants:

*I still don't like the idea [of not asking for consent], you've got something on me that I didn't consent to, you can sort of argue on principle, that's not a good thing to do, but it's still benign so it doesn't really matter [for example if the topic is shoe preferences]. But it becomes more of an issue when particularly the context that you -
- the theme of your work which is health. (REC 7)*

This was contrasted with more sensitive research topics (health-related, pregnancy blogs, or radical religious views expressed on twitter), where there was more chance of harm being done to the research participant through the collection of SM data. Beyond the principle of ‘do no harm’ that was of central importance to all REC members, varying levels of personal and professional experience shaped the way that interviewees viewed risk and consent in the context of SM research. We go on to explore these differing views in relation to REC member experience of SM research, below.

Inexperienced REC member views: The perceived need for consent

The majority of REC members with limited experience of SM research and/or review who perceived the need for consent for all SM research, spoke about falling back on strong traditional notions of ethical behaviour around consentⁱⁱ, and approaching ethical review as they would for non-SM research, “*look[ing] at it in terms of the usual principles apply*” (REC 15). These REC members felt uncomfortable allowing researchers to use SM data without consent, “*even if it’s for the good of society*” because users would be “*aggrieved*” (REC 7). Indeed, one interviewee likened their concerns to people speaking privately in a public staff canteen and their conversation being publicised out of context. Another interviewee explained how, even on public (non-password-protected) SM sites, researchers need to be sensitive to their use of SM data, and realise that there are ‘people behind the data’ who need to be ‘consented’, and made aware. In these instances the “*general principle*” taken by less experienced REC members was that researchers “*should attempt to inform and ask consent*” (REC 3).

Having said this, there was also a realisation by these particular interviewees that using this

non-exceptionalist approach (Gelinas et al, 2017), i.e., analogising SM research to its non-SM counterpart, did not always work in practice, because of the fast-paced and ever changing nature of SM platforms and associated research methodologies. For some particularly inexperienced RECs there was a degree of anxiety attached to the dynamic and changeable nature of SM as a technology, where uncertainty and concerns about “*known unknowns*” (REC 12) were viewed as something to be worried about in a somewhat reactionary way. It is likely that such uncertainty around not knowing ‘the potential harms until they happen’ resulted in these REC members taking an even more cautious stance in relation to the perceived need for consent.:

The way that it's [SM] being used you would never know what's next. Because when Facebook came in, there was also Myspace and other platforms were available. And people try to understand how to engage with those kinds of platforms and what were the ethics concerned when you were doing that. And then some other social media comes the next day, and it's a completely different way of dealing with it and you don't know. So...you don't know yet what they are or what are the potential harms until harm happens.... (REC 1)

Compounding all the difficulties and uncertainty expressed by less experienced REC members in relation to reviewing studies that included a SM component, were concerns relating to the lack of clear “*guidelines on how you deal with these issues*” that can make it “*really hard*” (REC 18) in the absence of a reliable normative framework. In spite of less experienced REC members struggling with the idea of using a non-exceptionalist approach in a changing research field, they tended to forward a view of SM research that upheld old

notions of the need to gain consent regardless of the practical difficulties obtaining it. This was because these interviewees were guided by a deep sense of responsibility to protect SM users - a view that was also shared by the REC members who were more experienced in both SM research and also review (n=10):

Some people will say okay, it's in the public domain so I can use it as I want. Actually as researchers...[we can] never use anything just as we want. We always have to look at the protection of participants, potential harms, whether or not it's us causing it or them causing it in a way...And if we're manipulating that as researchers, we have certain ethical requirements that we need to adhere to. (REC Member 4)

Here it was vital that SM user expectations (user awareness) were considered as part of best practice ethical decision-making, whereby thinking about “*the likely expectations of the people*” (REC 11) was paramount. For some interviewees, failing to take account of these user expectations meant that SM research was akin to “*covert research*”, that ran counter to traditional ethical norms of research in relation to consent and knowledge of being researched:

There's something about social media that people are less informed about how their data might be used...it feels more of a form of covert research in some ways because people are carrying out their research when other people who use social media may not be aware of it. (REC 19)

Such a view was strengthened in situations where SM data was sourced from password-

protected sites. Here interviewees placed even more responsibility upon themselves to protect SM users and to seek permission to use SM data, or at least make SM users aware that their data could be used for research purposes. In these instances, consent was viewed as being required by the majority of interviewees regardless of experience (n=17):

We certainly discussed the nature of the platform and distinction between public facing things like twitter and closed chat groups, very much an ethics committee would be mindful of certainly because I think that those lead to a difference in how we would view the privacy element of that. (REC 16)

Overall then, for less experienced REC members consent was seen as a compulsory element for any SM research, rather than something that could be applied more selectively depending on the perceived risk to SM users.

Inexperienced REC member views: Fair game

The alternative view taken by a minority of REC members with limited experience of conducting and/or reviewing SM research was that most SM users were aware that their data could be used for research purposes:

I think people are cognizant of the danger that they're using that kind of technology....that they might put up or post or upload to those sites, the way those sites can be used, co-opted by different people for different purposes... (REC 8)

These REC members (n=7) forwarded an almost laissez-faire view and saw SM data as “*fair game*” for research purposes where “*if you're prepared to put the stuff in the public domain, then it's in the public domain*” (REC 2). This view was particularly prominent for those with very limited experience of SM research that amounted to indirect experience of research and review of a single SM based study.

...if there is something that I can just access to somebody's data just by googling it, and then I don't need -- or a researcher doesn't need consent from the participant because already it's readily available. Why bother? (REC 1)

Alongside this “*fair game*” notion of SM data, there was an expectation that users *should* be aware of the public nature of online platforms, and it was not for REC members to be responsible for those who are not. This attitude was especially true if SM users were posting on a public (non-password-protected) site, such as twitter:

The question is, “are they responsible for not knowing or should someone be telling them?” And I suppose I tend to think that people have a certain responsibility for themselves...we don't remain children throughout our lives and the stuff that people put on social media I think sometimes is very self-destructive and stupid but they can't really say they didn't know it's public information. And if you don't want stuff to be made public then don't make it public. (REC 10)

The responsibility of the REC member was therefore generally seen as protecting against harm by solely ensuring that the potential for identifying SM users from their data was

minimised within both large-scale quantitative studies and small scale qualitative studies. The use of SM data without consent was therefore perceived as ethically unproblematic, especially in public forums, so long as users remained de-identified:

...what are the risks to the individual if this data is collected on them that they haven't consented to. Actually the risks are minimal, because, you can't trace it back to who they are...(REC 7).

Moreover, for some 'pretty liberal' (REC 10) interviewees, this risk-based approach to ethical review applied to research proposals sourcing SM data from a private (password protected) forum:

If you said, "look it's crucial we understand what people are discussing in a particular chat room and we're only interested in what they say, not who they are and the interactions that they're having and the way in which they interact" I suppose I would err on the side of being pretty liberal for that as long as the investigator was simply an observer or something....I mean it's not like they're breaking into someone's house and overhearing their conversation. They're actually free to join, they're free to listen.
(REC 10)

Here then, REC members with limited or no experience conducting and/or reviewing studies that included a SM component, tended to forward a view SM research that emphasised a fair game view that resulted in the perception of a reduced need for consent overall. It is worth noting however that less experienced REC members sometimes held both traditional notions

of the need to consent alongside fair game views of publically available data (e.g. REC 7). This highlights the complexity and uncertainty of ethical decision-making in relation to risk and consent in the online context.

Experienced REC member views: The nuanced middle ground

For the more experienced REC members with direct professional experience of conducting SM research (n=4), the fair game argument of publically available data was considered a “naïve view” (REC 16) and a “dangerous thought” (REC 14), which overlooked the complexities of user awareness and the responsibility of researchers to do no harm as a minimum standard. However, the need to always seek consent was also problematic for these REC members due to pragmatic considerations about study design where seeking consent was not always desirable or possible. Rather, more nuanced views that were simultaneously less cautious in relation to consent, whilst being more mindful of participant protection (as opposed to the fair game argument) tended to be expressed by these REC members.

While a variety of views were expressed by more experienced REC members, those who upheld a strong sense of responsibility in terms of protecting participants explained that whilst in an ideal world they would prefer researchers to ask for SM user consent directly, they understood the impracticalities presented by SM research in relation to this aim.

Methodology was an important factor in determining when it was appropriate to resort to notions of negotiated consent from SM users. This negotiated level of consent was perceived appropriate for studies using large-scale data sets because they presented a much lower level of risk:

If we're data mining 10,000 Twitter feeds...I don't think it's really necessary because we're not identifying anybody, we are aggregating those data, we're not disaggregating them. So, again, it's the actual research practice that's involved...no, we don't need to really -- asking a person's consent....it's not even their data that we're accessing, it's what their data does which is contribute to a spike in the use of a particular hashtag...it's not even the content of what they said. (REC 13)

Alongside this, interviewees were 'realistic' about the practical difficulties of gaining consent for SM research. For example, instances in which researchers are harvesting "millions of tweets" (REC 17) consent was not seen as being either practical or possible; nor where SM users are chatting on an anonymous chatroom; or had left the discussion:

But the informed consent process obviously, that will be very, very difficult to do. So you have to be realistic as well and I think, well, how would you email millions of people and would then do an opt out or opt in process or what -- so I think, in that case, in the past, we have said, well, okay, as long as nobody is harmed by this research in anyway and it's not against any terms and conditions then, it's probably okay. (REC 17)

In addition, dilemmas about altering user behaviour through attempting to obtain consent from participants were raised by interviewees:

And so the question arose whether that method of data collection met criteria for informed consent...they feared that if they announced themselves on this

chatroom site that people wouldn't be candid...so I thought, it's quite a dilemma. (REC 5)

REC member 17 explained how choosing to not ask for consent in these instances was not necessarily ethically problematic, so long as there was no potential for the SM users to be harmed during the research process. Indeed, when 'secured consent' (Hutton & Henderson, 2015) was unavailable at the point of data collection due to the need for unobtrusive researcher presence that didn't "*destroy the research*" (REC 11), or impracticalities obtaining consent directly, best practice was framed by interviewees as including safeguards that could: provide information to participants on the research being conducted; disclose researcher presence within forums; and enable retrospective consent and opt-out processes for research participants. In addition, negotiated consent processes were positioned as safeguards that could produce varying levels of research integrity and participant protection. For example, asking the moderator of a forum for consent was viewed as being more ethically rigorous than a terms and conditions disclosure that represents little more than a superficial box ticking exercise, as demonstrated by the following two quotes that are illustrative of each approach to consent respectively:

But I think guidelines particularly around where you don't need to get individual consent. I mean there might be situations where you can get group consent which is something I am very, very comfortable with particularly in anthropology, or you know or get a gatekeeper's consent. (REC 9)

And;

And, in contrasting a more rigorous approach to consent with a less rigorous tick box for opting out of the research; ...the question is how in some cases it's going to be impossible to get anything like traditional voluntary informed consent in which case you'll have to forgo the benefits that the research data could have accrued or you find some sort of quasi consent process, you know, involving some kind of dodgy opt out type, small print type thing which solves the conscience of the researcher but isn't consent in any kind of substantive sense. (REC 14)

More experienced REC members also explored the potential for terms and conditions disclosures to operate on a more meaningful level when combined with other safeguards such as anonymization. Such ethical strategies were described as representing a “*belt and braces approach*” (REC 19) that could be employed to further minimise risk to participants:

...we took an ethical stance on the project that I'm talking about by using discussion forums from organizations where they make it clear in the terms and conditions that data is publicly available, you don't have to register to read it. And also that the research might or the data may be used by third-parties for research purposes. So, we took a decision to go with that and so, that combined with the anonymisation is almost like a belt and braces approach. So, we've done everything we can to mitigate the risks of harm coming to individuals but we can't stop it completely.

Relatedly, more experienced REC members emphasised that terms and conditions consent within SM platforms need to find ways to “*make it clear...that data is [to be considered as] publically available*” (REC 19) [so as to avoid privacy violations through what has been](#)

~~regarded as ‘improper access’ by some authors exploring SM research (Zimmer, 2010). –so as to avoid the privacy violations through improper access that have been highlighted in the literature (Zimmer, 2010).~~ Best practice safeguards represent how REC members negotiated consent in relation to the idiosyncratic nature of SM data and the practical difficulties in obtaining permission from SM users. Far from being dismissive of the ethical complexities when reviewing SM research, these interviewees were just keener to engage with new ideas and ethical approaches to help facilitate a research project being conducted ethically without the necessary burden of obtaining traditional consent:

I think we need to be creative...I mean for any type of research we see that consent has a very clear historical route...but now it just doesn't serve the purpose...I wonder whether we should in some cases get rid of the informed consent altogether (REC 18)

Indeed, as REC 8 stated, the discussion should perhaps be less about “*whether or not you should be using public domain information*” and more about “*how you actually handle that kind of information, that kind of knowledge in a way that is responsible*” – a point echoed by more experienced REC members and best practice proponents (Vayena et al, 2015) alike. And whilst one experienced REC interviewee took a particularly principled stance in relation to consent where the “*right not to be researched and the right not to be observed*” (REC 13) was emphasised, this position did not preclude research in the absence of consent for this REC member, but rather necessitated the need for a contextualized, process-based and “*very refined ethical strategies and ethical guidance which lets people know how and why you might use those data*” (REC 13). This emphasises the need for more participatory practices which this interviewee noted should ideally reflect more sustained versions of consent rather than research practices where ethics and consent are seen as a ‘one-off’ (Hutton &

Henderson, 2015):

And I think there's an additional step here, which is ethics are seen to be a one-off process... I don't think that we consider enough the role of participants in informing the development of ethical strategies...Involve your participants, ask the people whose data you were using....ethics are practices of power, I think they're relational, I think they are things that we do with others and they shape the relationships that we have with them. (REC 13)

Overall then, for experienced REC members consent was not a compulsory necessity of good ethical practice, but rather the decision to require consent was balanced with the perceived risk and harm to SM users produced by the idiosyncratic qualities of each study.

Discussion

This study demonstrates that REC members hold broad conceptualisations of risk and consent in SM research, and it goes on to empirically study how these conceptualisations are brought into the ethical decision making process via the views of individual REC members with differing and varied opinions of how consent and risk are to be applied in the SM context (Samuel, 2017). While REC members struggle to reconcile the competing arguments that are presented by SM, this study shows that they are aware of them, and they actively incorporate them into their ethical decision-making. It has also explored how these views are related to the REC members' level of experience, showing discernible trends that were related to REC members' professional experience of SM research and/or review.

Whereas differences in views between REC members are an expected part of group dynamics, and are a strength of academic groups assessing ambiguous or untested criteria (Derrick, 2018), for SM these views are further complicated by the differing understandings of whether virtual data is considered public or private. In this study, the level of REC member experience of reviewing and also conducting SM research shaped their perception of risk, and the necessity of applying different levels of consent (traditional or negotiated). REC members with less experience of SM research and review had polarised views in terms of when consent was or was not required, supporting Vayena et al's (2015) observation that perceptions of privacy are often "based on a strict binary conception of identifiability or public availability" (Vayena et al, 2015; p. 435). In this conception, consent becomes the go-to standard of ethical accountability if either identifiability or public availability are in any doubt. Correspondingly, de-identification becomes the go-to ethical strategy if REC members doubted the extent to which data was publically available or the necessary level of consent was impractical or not possible. For Vayena et al (2015), binary conceptions of identifiability or public availability, as discussed above, are inadequate for SM research because de-identification can often represent an insufficient protection in the presence of technological advances that make re-identification a relatively simple task. As such, these authors suggest that users are often inadequately protected when de-identification is the go-to standard of protection and also when an approach "based solely on notice and consent where subjects have "limited opportunities to withhold, revoke, or modify consent" (Vayena et al, 2015; p.431) is taken.

In contrast, more experienced REC members were better able to negotiate consent, basing their ethical decision-making on more nuanced conceptions of risk in the online context where data was neither 'fair game', nor off limits without consent. This more nuanced approach in relation to risk and consent emphasised responsible usage of SM data. As such it

is in line with the recommendations of Hutton & Henderson (2015) where perceived privacy of SM users can be negotiated via data anonymization and the application of best practice when secured consent (Hutton & Henderson, 2015) is unavailable, either because of the practical difficulties obtaining it, or because of the advantages to be gained from forgoing user permission. In addition, some of these REC members highlighted participatory practices and dynamic forms of consent (Hutton & Henderson, 2015) that could respond to the individual context of each project to provide more meaningful, responsive and democratic approaches to participant permission.

This more nuanced approach to risk and consent is also more in line with Vayena et al's (2015) recommendation that we broaden conceptualisations of old ethical norms for SM research and that "researchers and review boards should be encouraged to incorporate systematic risk-benefit assessments [of SM research] and new procedural and technological solutions from the wide range of interventions that are available" (p 421). However, whilst more experienced REC members' approach to ethical decision-making was more nuanced, and moving more towards Vayena et al's (2015) recommendations, they still tended to fall back on binary notions of identifiability and consent in the context of public un/availability, i.e. although they considered publically available data as being in need of refined ethical strategies to allow its responsible use, they did not always have a clear idea about what these strategies beyond identification and consent may be. Rather they tended to articulate that consent didn't "serve the purpose" in the contemporary context of SM, and that a "belt and braces" approach combining ethical safeguards in order to minimise risk to participants was therefore required. This upholds Vayena et al's (2015) call for a new ethical framework based on "a modern understanding of privacy" (p.435) focusing on the increased potential to learn about others that SM provides and a reduced reliance on the binary requirements of consent

and de-identification.

Best Practices

Participant protections in the context of ostensibly publically available data are less obvious on both an ethical and practical basis. [In order to avoid privacy violations based on the improper access that has been raised as a concern by some SM researchers \(Zimmer, 2010\)](#)~~#~~ ~~order to avoid privacy violations based on improper access~~, it is vital that RECs are able to gain a tacit understanding of the complexities of SM research where privacy, user awareness and consent combine to obfuscate the core ethical principle of ‘do no harm’. This will involve REC members gaining important experience in SM research and review, so they are better able to align with the modern understanding of privacy and adequate approaches to protecting participants that are emphasised by Vayena et al (2015) and also Henderson & Hutton (2015). More experienced REC members came close to this ideal, but they still fell short in terms of moving beyond traditional perceptions of de-identification and consent as representing adequate safeguards to mitigate risk – particularly in the context of ‘low-risk’ aggregated data. Creating sustainability and establishing trust (Vayena et al, 2015; p. 435) in research that includes a SM component requires a broad toolkit of protections that acknowledge ethics as practices of power and take into account the unique challenges to privacy that SM research creates. Such protections should focus less on the binary constructs of consent and de-identification, and more on the qualitative nature of risk within individual studies, to protect participants and to facilitate research that makes good use of the opportunities that SM affords.

Research agenda

Challenges to the ethical conduct of SM research will continue to evolve rapidly as technology moves forward and as privacy concerns and research regulatory structures change on a global basis. Research should continue to monitor and evaluate risks within the online context, and to seek procedural solutions to the protection and confidentiality of data, respect for the individual, and the personal privacy of those whose data are captured for research purposes. This research should also explore different, innovative types of consent which might be more relevant for SM users and researchers who wish to use their data.

Educational implications

These results have implications for all parties involved in research that utilises SM data including RECs, researchers, research supervisors and research subjects. In particular, providing REC members who may be unfamiliar with SM on a personal or professional basis with a tacit understanding of consent, risk and participant protections in the online context would be a useful way to reduce the opacities associated with the ethics of SM research.

Endnotes

ⁱ It is worth noting that not all SM research goes through ethical review due to perceptions of SM research as a low risk context where the notion of harm is mitigated by the lack of physical interaction.

ⁱⁱ Here we talk about do no harm and consent/identification – but there was also a lot of ethical discussions about the legalities, data storage etc. that are not discussed here.

References

Acquisti, A., & Gross, R. (2006, June). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *International workshop on privacy enhancing technologies* (pp. 36-58). Springer, Berlin, Heidelberg.

Formatted: Italian (Italy)

Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday*, 11(9).

Beninger, K., Fry, A., Jago, N., Lepps, H., Nass, L., & Silvester, H. (2014). Research using social media; users' views. *NatCen Social Research*, 1-40.

British Psychological Society. (2007). Guidelines for ethical practice in psychological research online. British Psychological Society, Leicester: UK. Retrieved from http://www.bps.org.uk/sites/default/files/documents/conducting_research_on_the_internetguidelines_for_ethical_practice_in_psychological_research_online.pdf

British Psychological Society. "Code of Human Research Ethics". (2010). Retrieved from http://www.bps.org.uk/sites/default/files/documents/code_of_human_research_ethics.pdf

Boyd, D., & Crawford, K. (2011, September). Six provocations for big data. In *A decade in internet time: Symposium on the dynamics of the internet and society* (Vol. 21). Oxford: Oxford Internet Institute.

Buchanan, E. A., & Hvizdak, E. E. (2009). Online survey tools: Ethical and methodological concerns of human research ethics committees. *Journal of Empirical Research on Human Research Ethics*, 4(2), 37-48.

Charmaz, K. (2014). *Constructing grounded theory*. Sage.

Carter, C. J., Koene, A., Perez, E., Statache, R., Adolphs, S., O'Malley, C., ... & McAuley, D. (2016). Understanding academic attitudes towards the ethical challenges posed by social media research. *ACM SIGCAS Computers and Society*, 45(3), 202-210.

Custers, B., van der Hof, S., & Schermer, B. (2014). Privacy expectations of social media users: The role of informed consent in privacy policies. *Policy & Internet*, 6(3), 268-295.

Formatted: Italian (Italy)

Formatted: German (Germany)

[Derrick, G. \(2018\) Derrick GE \(2018\) The Evaluators' Eye: Impact assessment and academic peer review. Palgrave MacMillan. London, UK.](#)

Ess, C. (2002). Ethical decision-making and Internet research: Recommendations from the aoir ethics working committee. *Readings in virtual research ethics: Issues and controversies*, 27-44. Retrieved from <http://aoir.org/reports/ethics2.pdf>

[Gelinas, L., Pierce, R., Winkler, S., Cohen, I. G., Lynch, H. F., & Bierer, B. E. \(2017\). Using social media as a research recruitment tool: Ethical issues and recommendations. *The American Journal of Bioethics*, 17\(3\), 3-14.](#)

Formatted: German (Germany)

Gymrek, M., McGuire, A. L., Golan, D., Halperin, E., & Erlich, Y. (2013). Identifying personal genomes by surname inference. *Science*, 339(6117), 321-324.

Hammersley, M. (2009). Against the ethicists: on the evils of ethical regulation. *International Journal of Social Research Methodology*, 12(3), 211-225.

Hargittai, E., & Litt, E. (2013). New strategies for employment? internet skills and online privacy practices during people's job search. *IEEE security & privacy*, 11(3), 38-45.

Hargittai, E., & Marwick, A. (2016). "What Can I Really Do?" Explaining the Privacy Paradox with Online Apathy. *International Journal of Communication*, 10, 21.

Hedgecoe, A. (2017). Scandals, ethics, and regulatory change in biomedical research. *Science, Technology, & Human Values*, 42(4), 577-599.

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

[Homer, N., Szelinger, S., Redman, M., Duggan, D., Tembe, W., Muehling, J., ... & Craig, D.](#)

Formatted: German (Germany)

W. (2008). Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS genetics*, 4(8), e1000167.

Hutton, L., & Henderson, T. (2015, May). " I didn't sign up for this!": Informed consent in social network research. In *Proceedings of the 9th International AAAI Conference on Web and Social Media (ICWSM)*.

Jones, C. (2015). Ethical issues in online research, British Educational Research Association on-line resource. Retrieved from <https://www.bera.ac.uk/wp-content/uploads/2014/03/Ethical-issues-in-online-research.pdf>

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business horizons*, 53(1), 59-68.

Keim-Malpass, J., Albrecht, T. A., Steeves, R. H., & Danhauer, S. C. (2013). Young women's experiences with complementary therapies during cancer described through illness blogs. *Western journal of nursing research*, 35(10), 1309-1324.

McKee, H., & Porter, J. E. (2008). The ethics of digital writing research: A rhetorical approach. *College Composition and Communication*, 711-749.

Markham, A., & Buchanan, E. (2012). *Ethical decision-making and internet research: Recommendations from the aoir ethics working committee (version 2.0)*.

Metcalf, J., & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society*, 3(1), 2053951716650211.

Moreno, M. A., Goni, N., Moreno, P. S., & Diekema, D. (2013). Ethics of social media research: common concerns and practical considerations. *Cyberpsychology, Behavior, and Social Networking*, 16(9), 708-713.

Morrison, A., McMillan, D., & Chalmers, M. (2014, October). Improving consent in large scale mobile HCI through personalised representations of data. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (pp. 471-480). ACM.

Formatted: German (Germany)

Narayanan, A., & Felten, E. W. (2014). No silver bullet: De-identification still doesn't work. *White Paper*.

Norval, C., & Henderson, T. (2017). Contextual Consent: Ethical Mining of Social Media for Health Research. *arXiv preprint arXiv:1701.07765*.

Park, Y. J. (2013). Digital literacy and privacy behavior online. *Communication Research, 40*(2), 215-236.

Paechter, C. (2013). Researching sensitive issues online: implications of a hybrid insider/outsider position in a retrospective ethnographic study. *Qualitative Research, 13*(1), 71-86.

Sage Bionetworks. (2017). Retrieved from <http://sagebase.org/governance/pcc-toolkit-frequently-asked-questions/>

Samuel, G. (2017). "The Danger of Lurking": Different Conceptualizations of "User Awareness" in Social Media Research. *The American Journal of Bioethics, 17*(3), 25.

Solove, D. J. (2012). Privacy self-management and the consent dilemma.

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge University Press.

Vayena, E., Gasser, U., Wood, A. B., O'Brien, D., & Altman, M. (2016). Elements of a New Ethical Framework for Big Data Research.

Vayena, E., Mastroianni, A., & Kahn, J. (2013). Caught in the web: informed consent for online health research. *Sci Transl Med, 5*(173), 173fs6.

Williams, M. L., Burnap, P., & 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, 0038038517708140*.

Formatted: Italian (Italy)

World Medical Association. (2001). World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization*, 79(4), 373.

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