Open Research Online

The Open University's repository of research publications
and other research outputs

Ethics issues in digital methods research

Book Section

oro.open.ac.uk

Ethics Issues in Digital Methods Research

Claire Hewson

The Open University

(Pre-publication copy, March, 2015)

INTRODUCTION

This chapter addresses issues of ethical research practice in the context of the recently emerging range of methods which use the internet to support the creation of primary research data – variously referred to as online methods, digital methods, and *internet-mediated research* (IMR) (here, the latter term will be used). Social and behavioural researchers started devising and piloting IMR methods from around the mid-1990s, with surveys, experiments, interviews, and observational studies all being represented in early pioneering attempts (e.g. Bordia, 1996; Hewson, 1994). Since then IMR methods have flourished, expanding in volume, interdisciplinary reach, and range of methodological approaches (as discussed in the introduction to the present collection). In particular, the emergence of 'Web 2.0', as discussed in Chapter 1, has facilitated the recent expansion of unobtrusive methods, including those involving data 'mining' or 'harvesting' (often requiring the use of complex computer algorithms), which can lead to what have become known as 'big data'[i] sets. Such unobtrusive approaches which make use of the digital traces of

peoples' online behaviours (typically, without obtaining consent) have led to debates regarding what is appropriate ethical practice in an IMR context; in particular, a salient issue has been the distinction between what should be considered 'private' and 'in the public domain' in an online context. Obtrusive research methods in IMR (where informed consent *is* obtained) have also created new challenges and debates regarding what is ethical practice (as discussed further below). Some of the difficulties which can arise may not be immediately obvious to researchers accustomed to gathering data using traditional offline methods, creating the need for IMR-specific guidelines which highlight issues and caveats. The ethical issues which emerge in an IMR context may also potentially lead to new ways of thinking about ethics in traditional (offline) research contexts, for example by highlighting ambiguities in what constitutes 'public' and 'private' spaces, or questioning the distinction between 'real' and 'pseudonymous' identities. The impact of such debates in transforming traditional conceptions of ethical research practice remains largely to be seen. A key tenet of the present chapter is that striving to specify a pre-defined set of 'online research ethics rules' is not very useful. Rather, researchers need to carefully assess a number of key considerations, and make decisions, within the context of any particular research project. The present discussion is intended to be of use to IMR researchers, students and Research Ethics Committees (RECs) when planning, designing and assessing an IMR research study. To this end, some practical guidance on best practice procedures across a range of methodological approaches is offered throughout.

ETHICS ISSUES IN INTERNET-MEDIATED RESEARCH

Existing IMR Ethics Guidelines

Early pioneers in the use of IMR methods often focused primarily on methodological issues, rather than ethical considerations (Peden & Flashinski, 2004). However, it soon

became apparent that a number of pressing issues, both in applying existing ethical standards and in resolving novel issues that emerge in an IMR context, required attention. In 1999 the report of a workshop convened by the American Association for the Advancement of Science (AAAS) was published (Frankel & Siang, 1999), which highlighted the following key issues in online research ethics: complexities in gaining informed consent; the use of 'anonymous', or pseudonymous, identities; exaggerated expectations of privacy; the blurred public-private domain distinction. The report also questioned the applicability and interpretation of existing 'human subjects' research guidelines in an IMR context. Taking *autonomy*, *beneficence* (maximising benefits, minimizing harm) and *justice* as basic ethical principles for directing research with human participants, the AAAS document identifies several enhanced risks in IMR: dubious reliability and validity of data; greater scope for leakage of research participants' personally identifiable data; difficulty in implementing robust informed consent and debrief procedures; ambiguities over what is 'in the public domain' online. Since this early discussion paper, several professional bodies have published ethics guidelines for IMR, including The Association of Internet Researchers (AoIR, formed in 1998) (Markham & Buchanan, 2012 [*version 2*]); the American Psychological Association (APA) (Kraut et al, 2002); the British Psychological Society (BPS, 2013 [*version 2*]). Similarly to the AAAS document, the BPS (2013) guidelines identify some basic ethics principles for research with human participants – *respect for the autonomy and dignity of persons*; *scientific value*; *social responsibility*; *maximising benefits and minimising harm* – and discuss particular issues which can arise in adhering to these principles an IMR context. Discussions of ethics in IMR can also be found in journal papers (e.g. Rodham & Gavin, 2006), book chapters (e.g. Ess, 2007), and online resources and guidelines (e.g. see the information available at: *Exploring Online Research Methods:*

http://www.restore.ac.uk/orm/ethics/ethcontents.htm; also the IMR-specific guidelines of various university ethics committees, locatable by searching online). Also noteworthy is the cross-disciplinary journal dedicated to the topic: *International Journal of Internet Research Ethics* (IJIRE, http://ijire.net/). The next section now discusses the main IMR ethical issues requiring consideration, as identified in the existing literature.

Obtrusive IMR methods

Gaining Informed Consent

In situations where participants are actively recruited to knowingly take part in a research study, informed consent is required, and participants must be given the right to withdraw their consent, during or after participation (within a reasonable timescale), and must be suitably debriefed once participation is complete. In IMR ensuring these procedures are implemented effectively can be problematic for a number of reasons. Most fundamentally, reduced levels of researcher control compared with many offline contexts, due to the non-proximal nature of the interaction, can raise various issues. Whilst non-proximal methods are also commonly used in some types of offline research, such as postal surveys, a unique feature of IMR is that it allows far greater levels of interactivity whilst maintaining high levels of anonymity, and reduced identifiability and traceability of participants. Thus, an online experiment could potentially take place in a rich, interactive 3D environment, without the researcher knowing the (offline) identities of participants, or having any means of tracing them after completion of the study – or, perhaps most significantly, if they withdraw and choose to exit part way through. Likewise, online surveys can be (and often are) placed in publically accessible locations for participants to discover and complete whilst remaining totally anonymous (apart from perhaps the logging on an IP address) and untraceable. IMR interview and focus groups involve participants being remotely present,

rather than actually present with the researcher as in offline face-to-face contexts, often interacting only via typed text, using asynchronous or synchronous discussion / chat software. Obtrusive observational studies, such as ethnographic research within a discussion forum or online virtual world, allows participants to remain relatively anonymous, compared with offline face-to-face approaches.

This distinct feature of IMR can lead to the following problems in devising effective informed consent procedures: ensuring, verifying and documenting that participants have actually read, and understood, consent information; ensuring and verifying that participants are eligible to give consent, for example are not underage, or unable to consent for some other reason. Various solutions and good practice suggestions have been offered to address these difficulties. Presenting consent information in a way that is accessible (clear and easy to read), informative, but not over-lengthy, should help encourage engagement from participants. In the case of web-based surveys and experiments, providing an initial page with this information – which should be informative but succinct – is effective. To check for understanding, including a bulleted list of statements to endorse (e.g. using check boxes) is a useful strategy (BPS, 2013). To avoid participants simply going through and ticking 'yes' to all statements, varying the appropriate response (e.g. 'yes', or 'no') can be effective. A final check box indicating explicit consent to take part in the study is a good way of keeping a record that consent has been formally obtained (it can be useful to include this again at the end of a survey or experiment, to double check).

For other obtrusive methods, such as disclosed ethnographic observation within an existing discussion forum, reliably contacting *all* participants to gain consent can be

problematic. Different strategies are possible, including sending an email to all listed members of a discussion forum, or posting an 'opt-in' or 'opt-out' message to the group (it is generally good practice, and netiquette, to go through group moderators before taking any of these actions). If discussion group *archives* are being used, the email method may be best able to catch members who have been inactive for some time, or since left the group (though dormant email accounts, and transmission errors, make this approach not foolproof). Where discussion groups are larger, reliably contacting all contributors becomes increasingly difficult (leading some researchers to argue that undisclosed approaches which waive informed consent are justified: see further discussion below). Creating a dedicated research study site to which potential participants are invited, e.g. by posting adverts to an existing group, can be an option if the research study design permits (obtaining consent on entry to the site, for example). This approach also allows greater researcher control over the security and confidentiality of research data gathered (discussed further below). However, it might not always be appropriate, for example where more 'naturalistic' settings are desired.

Ensuring that consent gained from participants is actually 'valid' raises particular difficulties in IMR, due to the diminished verifiability of participant characteristics. Using sampling methods that are less likely to reach and attract certain groups, such as those under the required consent age, can be useful. A careful risk assessment of the possible harm that may result from access by participants not eligible to give consent should also be carried out. For example, research which is particularly sensitive, or involves adult themes, should be considered higher risk. The BPS (2013) guidelines outline some specific good practice strategies, including asking for age information prior to presenting study materials so participants who do not meet the minimum age requirements can be redirected, and

blocked from re-entering the study (e.g. this can be achieved using cookies). In high risk situations additional safeguard measures can be used, such as verifying participant characteristics (and perhaps also recording consent) offline (BPS, 2013), or using methods which maximise levels of identifiability, such as Skype interviews. However, in some cases it may be decided that the risks are too high to allow the study to be conducted online. Research which involves obtaining parent / guardian consent should not to be ruled out in IMR, but requires especially careful risk assessment and stringent ethics procedures to be in place (see Hessler et al, 2003 for an example).

Withdrawal and Debrief

Implementing effective withdrawal and debrief procedures can be problematic in IMR. In web-based surveys and experiments a participant (typically) interacts remotely with an automated program (or HTML form), without any researcher presence, and often anonymously with no subsequent way of being traced or contacted (this is the case, at least, with many of the publically available web surveys readily accessible at sites such as: http://www.onlinepsychresearch.co.uk [Online Psychology Research UK]). In such situations, there is no way for the researcher to know whether the participant has engaged with the debrief information presented (usually placed on the final page of the survey / experiment, and offering researcher contact information). In cases where a participant withdraws early, ensuring debrief can become even more problematic since exiting by closing a browser precludes the debrief page being presented at all. Providing a clearly visible 'withdraw' button on each page of a survey (or experiment), and urging participants to use this if they wish to exit early can help, and allows redirection to a debrief page (at least this is an option in many of the most popular current online survey software packages). Adding a question asking participants if they are happy for any partial

submitted data to be used is also useful, since using these data could violate withdrawal rights if participants do not want their partial responses to be used; on the other hand, not using partial responses may violate respect for the autonomy of individuals, and the time they have given the study, if they did wish their (incomplete) answers to be used. In IMR it is often harder (than in many offline contexts) for a researcher to ascertain whether participants may have left a study in a state of distress, or discomfort, so paying particular attention to developing robust, reliable withdrawal and debrief procedures is important. Trade-offs can emerge, for example allowing post-study withdrawal requires being able to identify an individual's set of responses, but such traceability can conflict with anonymity requirements. Existing guidelines (e.g. BPS, 2013) offer some advice on the methods available (e.g. using an email address, or password, for identification purposes), and when they might be useful. Different approaches will suit different contexts; for example, if deception is involved, then ensuring debrief becomes more urgent, so maximising traceability at the expense of anonymity might be appropriate.

Even in contexts involving greater levels of researcher involvement, such as online interviews, focus groups, participant observations, and so on, similar withdrawal and debrief issues can arise, since participants can leave mid-study without any explanation. However, in these contexts a researcher is more likely to notice withdrawal when it occurs, and detect any adverse reactions from participants during the study. Ensuring a means of contacting any participants who might disappear in this way is advisable, so that debrief information can be sent, and follow-up questions and checks offered (e.g. regarding a participant's wellbeing). Knowing whether participants are happy for their existing contribution to be used as research data is also important, especially in contexts where removing the contribution of one person (e.g. from a discussion group) can be

challenging, and perhaps detrimental to the quality of the data gathered. Contacting participants in these situations may not always be straightforward, even in cases where an email address has been obtained (see above for barriers to the use of email to contact participants). Membership of existing online groups and virtual spaces is fluid and ever-changing. An avatar in SecondLife might be there one day, and gone the next, similarly to a participant in a discussion forum or mailing list group. The safest strategy is for researchers to use procedures which maximise the identifiability and traceability of research participants; this is especially important for very sensitive research where risk of harm to participants is greater. Setting up a dedicated research site which participants access with a username / password, linked with a valid email address, can be an effective strategy (a temporary email address could be provided, which can be set to forward to a more permanent account, if greater anonymity is a desired).

Ensuring Confidentiality

In both obtrusive and unobtrusive (see below) IMR methods there are a number of threats to the confidentiality of participants' data which go beyond what is normally present in offline research (Reips & Bufadi, 2012). Risks include hacker access, transmission errors (e.g. if using email), and third party control (e.g. if using server-hosted online study software solutions). Traditional offline approaches which store data on local media storage devices, or in hardcopy form (e.g. pen and paper questionnaires) are not subject to these risks. Most crucially, in IMR careful measures are needed to minimise the risk of participants' *personally identifiable* data being accessed by an unauthorised third party. One established way of maintaining confidentiality is to make sure all individual responses collected are anonymous, so that even if the data are accessed (e.g. by third party providers, or hackers), responses cannot be traced back to the individuals who produced

them. However, in IMR potentially identifying information is more likely to be automatically recorded as part of the data collection process; for example online survey software often stores information about IP addresses alongside responses (though in some packages, such as SurveyMonkey, this feature can be turned off), and posts to online discussion groups typically are accompanied by a username associated with an email address. Compare this with a laboratory-based experiment, where participants take part on a local computer provided by the researcher and are assigned a 'code name / number' which is only linked with their responses by a separate coding sheet, used only by the researcher. The IMR researcher thus needs to carefully consider additional risks, solutions and precautionary measures, following the guidance available in existing texts (e.g. Hewson et al, 2003).

Stripping away any potential personally identifying information from data sets at greater risk of third-party access is one important safety measure, and can work well when the data sets, and any original source material from which they were drawn (e.g. a dedicated, private discussion forum), are fully under the control of the researcher. In cases where data is sourced from or stored using third-party services, such as hosted survey or experiment software, or public online discussion forums, additional complications emerge. Here, the researcher is unable to access and remove personally identifying information from the original data sources, so particularly careful consideration must be given to dissemination practices, due to the durability, traceability and searchability of these original sources. For example, publishing verbatim quotes from a public online discussion forum could lead to them being traced back to source, viewed in context, and individual authors being identified, posing a serious potential threat to participant confidentiality. The problem is further compounded if the researcher publishes the name of the original source

(e.g. website address or social network site [SNS] name) alongside the reported research findings. For this reason, it has been suggested that source material information should not be given, and that quotes should not be used, or at least should be carefully paraphrased (BPS, 2013). Objections to these recommendations refer to such constraints sometimes being detrimental to research goals. On balance, risks associated with traceability and identifiability of participants must be weighed against the research benefits of disclosure during dissemination.

The traceability of individual's data should also be considered in terms of the local equipment individuals use to participate in a study, since computers store traces of the activity that occurs on them. Interacting with a web browser can leave traces, including the text responses given to survey questions, on the local computer which was used; this means other users of the same computer equipment may be able to retrieve and view these responses. Researchers should assess the extent to which it is their duty to inform participants (e.g. in the debrief, or consent, information) that this is the case, and offer advice and guidance on how to remove such traces, where this is possible. Finally, the leakage of researcher interpretations of participants' responses is another consideration to be taken into account – given the enhanced accessibility of published research reports online these days, there is greater risk that a research participant, or someone who knows them, might locate such interpretations. This means that extra care might be needed to protect the personal identities of research participants in published reports of findings (and this applies equally to offline studies whose findings are disseminated in ways that make them more publically accessible).

Unobtrusive IMR Methods

The Public-Private Domain Distinction Online

The issue of when the traces of online activity left behind by individuals should reasonably be considered 'in the public domain', and thus arguably available for use as research data without gaining informed consent, has been and remains deeply controversial (e.g. see Brownlow & Dell, 2002). A number of factors and considerations come into play in attempting to answer this question, and as always, decisions must be guided by the particular research study context. As with obtrusive methods, key factors to consider in unobtrusive IMR are: risks of harm to participants caused by the research, e.g. through making personally identifiable information (particularly where this is sensitive, such as illegal activities) more likely to be discovered and disseminated, and likely benefits of the research. A key consideration, which relates to the principle of *autonomy*, is ascertaining users' own *perceptions* and *expectations about* what is public and private online (BPS, 2013; Markham & Buchanan, 2012). However, user perceptions still remain largely unknown. A study which set out to investigate privacy perceptions, using an experimental paradigm, is reported by Hudson & Bruckman (2004). They entered chat rooms and either posted a message alerting participants that they wished to log their discussions (with opt-in or opt-out conditions) or simply lurked (remained present not posting anything). They received hostile responses in both cases, but were kicked out *less often* when they simply entered and lurked, leading them to argue that non-disclosure is justified in order to be able to carry out valuable research. However, other researchers have reported different, more positive, outcomes when entering chat rooms to observe discussions and gather research data, both in 'lurking' contexts (e.g. Rodino, 1997), and where permission was requested (e.g. Madge & O'Connor, 2002). Further research is clearly needed on this topic, since knowing what the privacy expectations of users of online interactive spaces are (which likely vary individually and culturally) is fundamental to reaching conclusions on

the public-private distinction debate (Ferri, 2000). It is worth noting that, strictly speaking, anything covered by copyright law is *not* in the public domain (BPS, 2013), which would render readily 'accessible' content such as Facebook posts, Tweets, and various other social media sources, not 'publically available' for use as research data. In terms of ethical research practice, a more common conception is that anything which is accessible to anyone, without the need for explicit permission, is 'in the public domain'. Legalities such as copyright and data protection laws are important to keep in mind, but may not necessarily dictate what constitutes good ethical research practice (see later example below).

People's awareness of the extent to which their online activities are logged and made available for third-party access needs consideration. If individuals are unaware that their contributions (e.g. to a discussion group), or activity traces, can be accessed and viewed, they may engage in disclosures and / or behaviours that they otherwise would not submit. Furthermore, traces that were once private can later become public (BPS, 2013), without a person's knowledge or permission. Another important point is that given the nature of online communications, and the settings in which they take place – for example, simultaneously in a private space such as a bedroom, and in a public online discussion forum – they may be perceived as both 'private' and 'public' at the same time. This might lead to greater candour and disclosures than individuals might engage in if in an offline fully public setting. All the aforementioned considerations complicate the issue of when it is acceptable, ethically, to use online information that is readily available. In particular, with all the 'unknowns' about user perceptions, knowledge, expectations, and so on, assessing likely *levels of harm* in doing so can be difficult. Fundamentally, deciding what is 'in the public domain' is complex and contextual, and in ambiguous cases researchers must

exercise their judgement to assess a range of factors (as mentioned above, e.g. sensitivity of the data) to make appropriate decisions which serve to minimise risks and maximise benefits. As noted in both the BPS (2013) and Markham & Buchanan (2012) guidelines, a context-dependent, bottom up, flexible approach is appropriate when dealing with ethics decisions in IMR, in general, and this is the approach advocated here.

Practical considerations are also important – for example, it may often not be practical (or indeed possible) to reliably contact all individuals who have contributed information to a potential data source (e.g. as in an archived discussion forum), due to the fluid nature of individuals' presence in online spaces (as well as dormant emails addresses, transmission errors, and so on). The *scientific value* of a piece of research may also be compromised by disclosure, due to interfering with the authenticity of participants' behaviours, and thus reducing ecological validity (as reported by Reid, 1996, in an observational study in a MUD [Multi User Dungeon] environment). Finally, the principle of *social responsibility*, particularly avoiding disrupting existing social structures, becomes relevant here – one argument against disclosure in observational research online being that it may harm group members by fostering suspicion and mistrust.

The above discussion has indicated the complexity of making decisions about when it is ethically appropriate to gather data unobtrusively, without disclosure, from online sources. An illustration of the contrasting approaches that have been taken can be found in comparing two very similar ethnographic studies of pro-anorexia support groups. Fox, Ward & O'Rourke (2005) contacted moderators and group members of a pro-anorexia website to fully disclose their intentions to follow discussions, using participant observation methods, and use these as research data. Permission was granted and they report

gathering valuable data using this approach. Brotsky and Giles (2007), on the other hand,

argue that disclosing their intentions as researchers in a similar study (also following

discussions on a pro-anorexia website) would have compromised the integrity of the

research. Hence, they went undercover, one researcher posing as a plausible persona

within the group for some time before going on to gather data from members' posts which

was then used as research data (and published, including using quotes from group

discussions). The latter approach is highly controversial (indeed, the author has

encountered reactions of outrage when reporting on this study in seminar sessions).

Nevertheless, the authors present clear arguments for adopting this stance, and acquired

REC approval for the study. Interestingly, moderators may even sometimes block attempts

to disclose research intentions to discussion group members, as reported by Tackett-

Gibson (2008) who intended full disclosure (to observe online communities engaged in

exchanging drug use information), but was given permission by group moderators only to

lurk and carry out observations unobtrusively, and access stored discussion archives. This

relates back to the principle of social responsibility and avoiding disrupting existing social

groups. As noted earlier, it is good practice to contact moderators before proceeding with

an observation study of an online group, and they may well have particular insights into

whether disclosure is likely to disrupt, damage, or benefit a particular group.

Ensuring Confidentiality

The topic of confidentiality in IMR was discussed above in relation to obtrusive methods.

As noted, the particular issues with internet data collection, in comparison with many

offline methods, relate to storing data securely, maintaining participant anonymity (so

individual responses are not personally identifiable), and using dissemination procedures

that do not entail a high risk of allowing data to be traced back to source (such as an

active online discussion forum used as a site for ethnographic research). In unobtrusive methods, where informed consent is typically not obtained (as discussed above), ensuring data confidentiality arguably becomes most crucial. Thus, when informed consent is gained participants can be warned of any risks, and may be willing, for example, to agree to anonymised quotes being disseminated, or in some cases to waive anonymity (e.g. some activist or marginalised groups are keen for their voice to be heard). However, when traces are 'harvested' as data without an individual's consent, making sure that these data cannot be linked back to an identifiable person becomes especially crucial (e.g. BPS, 2013). In cases where large, quantified, aggregate data sets are gathered, such as from Google searches, webpage browsing activity, shares on Twitter, SNS links, etc. risks of data becoming personally identifiable are typically at their lowest. As long as the researcher takes steps to strip away any potentially identifying information (such as IP addresses, email addresses, usernames, etc.) then storing and disseminating such large data sets poses minimal risks to the individuals who provided the data.

Risks will often be higher in qualitative research studies, however, particularly those which source linguistic data from online discussion forums, and in these cases the aforementioned considerations (such as whether to reveal names and locations of sources, or use verbatim quotes) require careful attention. Again, a range of contextual factors will come into play in assessing risks and making appropriate, ethically sound decisions; these include the sensitivity of the research topic, and data, how vulnerable the individuals involved are, how beneficial the research is, and so on. The core principle repeated in the BPS (2013) guidelines is pertinent here, which is that a researcher should assess all key relevant factors, and "ensure that ethics procedures and safeguards are implemented so as to be proportional to the level of risk and potential harm to

participants." (BPS 2013:8). As mentioned earlier, sometimes legal requirements and good ethical practice may seem to be in conflict; for example, copyright legislation may dictate attributing authorship for any online documents sourced (e.g. personal blogs or homepages), which might be in conflict with requirements to maintain anonymity and confidentiality to protect individuals. Some further discussion of legal issues in IMR can be found in BPS (2013) and Markham & Buchanan (2012).

CASE STUDY: MANIPULATING FACEBOOK USERS' MOODS

In a recent study on 'emotional contagion', Kramer, Guillory & Hancock (2014) examined the effects on individual Facebook users' moods of manipulating the 'emotional valence' of content on their News Feeds. The published findings from this research led to controversy over the ethical soundness of the methods employed, and a good deal of media attention and public comment (generally highly critical). For example, The Guardian reports: "Lawyers, internet activists and politicians said this weekend that the mass experiment in emotional manipulation was "scandalous", "spooky" and "disturbing"." [Guardian, June 2014]). What the researchers did was to take advantage of the way in which users' Facebook News Feeds are filtered, using a ranking algorithm, to select and present a subset of all the possible News Feed posts available (which Facebook uses to try to show users the content they will find most relevant and engaging). In an experimental manipulation, a random selection of users' (close to 700,000) News Feeds were adjusted to reduce the amount of emotional content presented; the study reports finding that reducing positive emotional expressions led people to produce fewer positive and more negative posts, and reducing negative expressions led to the opposite effect. In an editorial response published by the journal soon after the study appeared in print, it was recognised that the study had raised questions about the principles of 'informed consent',

'opportunity to opt out', and personal privacy. Legally, the use of Facebook users' personal data is sanctioned by the Data Use Policy which they agree to upon signing up. However, as noted earlier, adherence to legal requirements does not necessarily translate into ethical research practice.

A number of key points about the study are worth highlighting: Facebook New Feeds are routinely manipulated, so the present intervention was an extension of this practice; the effect sizes were extremely small (but nevertheless statistically significant due to the sheer sample size); the published findings used aggregate, fully anonymised data. These features could be used to justify the study as very low risk, in terms of potential harm to participants, and it is unlikely that any of the randomly selected Facebook users suffered any significant disruption or harm as a result of the study intervention. However, the publication and dissemination of the findings (including the media hype regarding the implications) led to a good deal of public anxiety and outrage at the findings and their implications. Interestingly, whilst REC (typically known in the United States as *Institutional Review Board*) approval for the study was sought, it was exempted from the review process since it used archival data collected by Facebook for internal purposes (it is not clear to what extent the research team had input into this process, however). This IRB decision could be challenged, depending on the details of the research design process, and the nature of the collaboration between Facebook and the research team. Such relatively novel digital methods research contexts remain a challenge for IRBs, in the absence of a set of established ethics guidelines. Clearly, there are presently ambiguities regarding how to deal with the research use of data collected online by private entities such as Facebook[ii].

CONCLUSIONS

This chapter has highlighted and discussed a number of key ethical issues identified as requiring extra care and attention in digital methods research, some of which may not be immediately obvious to the researcher accustomed to conducting research in offline contexts. Issues can emerge in properly applying existing ethical guidelines in IMR, to ensure robust and rigorous procedures to enable adherence to established ethics principles, as well as in tackling new challenges that emerge within an online context. Key issues include: deciding what is public and private online; devising robust informed consent, withdrawal and debrief procedures; and protecting participant confidentiality. As researchers continue to explore, pilot and debate IMR procedures and their ethical implications, principles and guidelines will evolve, and new procedures and ways of thinking about ethics will emerge, potentially impacting upon the way researchers think about research ethics in offline contexts. Digital methods are still in their relative infancy, and so are the ethical guidelines that have been developed to date. As highlighted by the 'emotional contagion' Facebook study, the issues can be varied and complex. What is institutionally sanctioned may still be seen as ethically problematic. Furthermore, researchers may disagree on key points, such as whether publically accessible traces of online behaviour can be considered available for use as research data without gaining consent from those who produced them. What does seem apparent is that no set of 'online ethics rules' for IMR is likely to be forthcoming; rather, flexibility in applying some broader key principles within the context of particular research study designs is required. Changing online landscapes, practices, new research data and researcher experiences will continue to shape the evolution, revision and refinement of ethical principles and guidelines for digital methods research.

Keywords for index: ethics; digital methods; online; guidelines; internet-mediated research (IMR); obtrusive methods; unobtrusive methods; informed consent; public; private; privacy; debrief; withdrawal; confidentiality; anonymity; surveys; experiments; interviews; focus groups; observation; discussion group; archives; risk of harm; traceability; software; deception; principles; copyright; dissemination; chat rooms; lurking; scientific value; social responsibility; ethnography; deception; Facebook; SNSs; trade-offs.

REFERENCES

BPS (2011). Code of human research ethics. *British Psychological Society.* INF.180/04.2011.

BPS (2013). Report of the working party on conducting research on the Internet: Ethics guidelines for internet-mediated research. *British Psychological Society*. INF206/1.2013. Available: http://www.bps.org.uk/system/files/Public%20files/inf206-guidelines-for-internet-mediated-research.pdf.

Brotsky, S. R. & Giles, D. (2007). Inside the "Pro-ana" community: A covert online participant observation. *Eating Disorders: The Journal of Treatment & Prevention 15*(2). 93-109.

Buchanan, E. & Ess, C. (2009). Internet research ethics and the Institutional Review Board: Current practices and issues. *Computers and Society 39*(3). 43-49.

Brownlow, C., & O'Dell, L. (2002). Ethical issues for qualitative research in on-line communities. *Disability & Society, 17*(6), 685-694.

Ess, C. (2007). Internet research ethics. In A. Joinson, K. McKenna, U. Reips & T. Postmes (Eds.), *Oxford handbook of internet psychology*. Oxford University Press.

Ferri, B. (2000). The hidden cost of difference: Women with learning disabilities. *Learning Disabilities: A Multidisciplinary Journal*, *10*(3), 129-138.

Fox, Nick, Katie Ward & Alan O'Rourke. 2005. Pro-anorexia, weight-loss drugs and the internet: an "anti recovery" explanatory model of anorexia. *Sociology of Health & Illness 27*(7). 944–971.

Frankel, M., & Siang, S. (1999). *Ethical and legal issues of human subjects research on the Internet - Report of a workshop.* Washington, DC: American Association for the Advancement of Science. Retrieved from: http://www.aaas.org/spp/sfrl/projects/intres/report.pdf (January, 2013).

Guardian (2014) "Facebook reveals news feed experiment to control emotions" [online]. The Guardian, 30th June, 2014. Available: http://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds.

Hessler R. M., Downing J., Beltz C., Pelliccio A., Powell M., & Vale W (2003). Qualitative research on adolescent risk using email: A methodological assessment. *Qualitative Sociology, 26*(1), 111–124.

Hewson, C.M. (1994). Empirical evidence regarding the folk psychological concept of belief. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society,* 403-408. Georgia, USA. ISBN 0-8058-1803-0.

Hewson, C.M., Yule, P., Laurent, D. & Vogel, C.M. (2003). *Internet Research Methods: A Practical Guide for the Social and Behavioural Sciences (1st edition).* Sage: London.

Hudson, J. & Bruckman, A. (2004). "Go Away": Participant objections to being studied and the ethics of chatroom research. *The Information Society, 20*(2), 127-139.

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J. & Cooper, M. (2004). Psychological research online: Report of board of scientific affairs' advisory group on the conduct of research on the Internet. *American Psychologist, 59* (4), 1-13.

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

Markham, A. & Buchanan, E. (2012). Ethical decision-making and internet research. Recommendations from the AoIR ethics working committee (version 2). Available: http://www.aoir.org/reports/ethics2.pdf.

Peden, B. F. & Flashinski, D. P. (2004). Virtual research ethics: A content analysis of surveys and experiments online. In E. Buchanan (ed.), Readings in virtual research ethics: Issues and controversies. Hershey, PA: Information Science Pub.

Reid, E. (1996). Informed consent in the study of online communities: A reflection on the effect of computer-mediated social research. *Information Society, 12*, 169-174.

Rodham, K., & Gavin, J. (2006). The ethics of using the internet to collect qualitative research data. *Research Ethics Review, 2*(3), 92-97.

Rodino, M. (1997). Breaking out of binaries: Reconceptualizing gender and its relationship to language. *Journal of Computer-Mediated Communication, 3*(3). Retrieved from: http://jcmc.indiana.edu/vol3/issue3/rodino.html (January, 2013).

Tackett-Gibson, M. (2008). Constructions of risk and harm in online discussions of ketamine use. *Addiction Research & Theory, 16*(3), 245–257.

---

[i] A term originating in computer science, meaning data sets which are so large that they are difficult to process computationally.

[ii] Interested readers should be able to find further discussions by conducting an appropriate Google search; the first author of the paper, Adam Kramer, has also posted a personal response on his own Facebook page (accessible at the following link, at the time of writing: https://www.facebook.com/akramer/posts/10152987150867796).