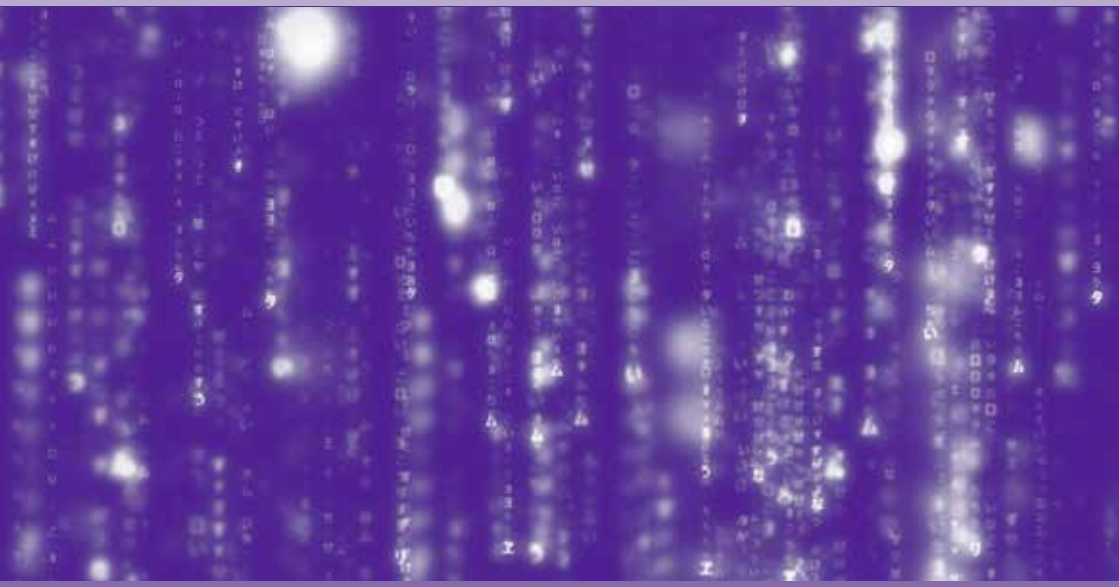
The British
Psychological Society
Promoting excellence in psychology

Ethics Guidelines for Internet-mediated Research



Authors

These guidelines were originally prepared by The Working Party on Internet-mediated Research, convened under the aegis of the British Psychological Society's Research Board in 2014. The guidelines were reviewed in 2017.

Dr Claire Hewson (Editor and Convenor)

Professor Tom Buchanan (Editor)

Dr Ian Brown

Dr Neil Coulson

Dr Gareth Hagger-Johnson

Professor Adam Joinson

Dr Aleks Krotoski

Professor John Oates

Citation: British Psychological Society (2017). *Ethics Guidelines for Internet-mediated Research.* INF206/04.2017. Leicester: Author. Available from: www.bps.org.uk/publications/policy-and-guidelines/research-guidelines-policy-documents/research-guidelines-poli

If you have problems reading this document and would like it in a different format, please contact us with your specific requirements.

Tel: 0116 2254 9568; e-mail mail@bps.org.uk.

Published by The British Psychological Society, St Andrews House, 48 Princess Road East, Leicester LE1 7DR.

Contents

1. Executive Summary

Internet-mediated research (IMR) can raise particular, sometimes non-obvious, challenges in adhering to existing ethics principles. In this document we outline some of the key ethics issues which researchers and research ethics committees (RECs) are advised to keep in mind when considering implementing or evaluating an IMR study. Considering each of the four main ethics principles as outlined in the Society's *Code of Human Research Ethics,* we highlight issues which may need special consideration in an IMR context, using illustrative examples to explain why. These issues include: the public-private domain distinction online; confidentiality and security of online data; procedures for obtaining valid consent; procedures for ensuring withdrawal rights and debriefing; levels of researcher control; and implications for scientific value and potential harm.

Emphasis throughout is on offering advice on how to think about and apply existing ethics principles in an IMR context, while recognising that issues need to be assessed and decisions made within the context of a particular piece of research.

2. Introduction

This document presents guidance on how the *Code of Human Research Ethics* (BPS, 2014) may be interpreted in the context of internet-mediated research (IMR) and what special considerations may apply. It should be considered as supplemental and subordinate to the Society's *Code of Human Research Ethics* and the overarching *Code of Ethics and Conduct* (BPS, 2009). It closely follows the principles and advice offered there, highlighting areas where these may become problematic and require particularly careful consideration in an IMR context.

The primary function of this document is to help researchers and RECs plan and evaluate research proposals, and to help with the process of ethical decision making in the context of specifying and implementing appropriate IMR research designs. It is not intended to provide a 'rule book' for IMR. It should be recognised that technologies, their social uses and the associated implications for research may change rapidly over time and new considerations will become salient. This requires a return to 'first principles' and an informed application of general ethics principles to the new situation. This document deals with some of the issues one may need to think about.

The *Code of Human Research Ethics* outlines the four main principles underpinning the ethical conduct of research:

1. Respect for the autonomy, privacy and dignity of individuals and communities;
2. Scientific integrity;
3. Social responsibility; and
4. Maximising benefits and minimising harm.

3. Internet-mediated Research

Advances in technology extend opportunities for psychological research. Such technological advances may also introduce additional, and sometimes non-obvious, complexities around adherence to ethics principles. This is particularly true in the case of internet-mediated research.

The term 'internet-mediated research', as used in this document, covers a wide range of quantitative and qualitative approaches to research involving human participants. IMR can be broadly defined as any research involving the remote acquisition of data from or about human participants using the internet and its associated technologies. As with traditional approaches, internet-mediated projects may adopt a variety of research designs. Their focus may be on obtaining quantifiable measurements (e.g. as in surveys or many types of experiment), or on obtaining rich, meaningful, elaborate narratives, as is often desired in qualitative approaches. They may be reactive (where participants interact with either the materials, as in an online survey, or the researcher, as in online interviews).

Alternatively, they may be non-reactive where data about individuals are collected unobtrusively (e.g. analyses of 'found text' in blogs, discussion forums or other online spaces, analyses of hits on websites, or observation of other types of online activity such as search engine histories or digital traces stored as a by-product of mobile app usage).

The boundaries between IMR and other designs can be blurred where research includes elements of both face-to-face observation/interaction and remote data collection. However, the key point is that the design normally involves acquisition of data from or about individuals in the absence of face-to-face co-presence. This restricts the researcher's capacity, in contexts where a participant is actively aware of and knowingly participating in a study (i.e. reactive contexts) to monitor, support, or even terminate the study if adverse reactions become apparent. Coupled with the greater scope for carrying out quite complex interactive procedures in IMR

with no direct face-to-face presence (e.g. in experimental designs), this makes IMR methods quite distinct from many offline methods where there is no face-to-face presence (such as postal surveys).

These key features of IMR can raise a number of ethics issues which need careful consideration. Additionally, very often research participants will be located in one or more different countries, so a project may span multiple nations, cultures and legal jurisdictions.

Different types of IMR design raise different ethics considerations. While many of these issues are dealt with in the *Code of Human Research Ethics* and are not unique to IMR, in this context they may create special considerations around the way the general principles should be interpreted and applied. For example, the extent to which the research can be thought of as occurring within a private or public domain, given that those boundaries are often blurred online, may be difficult to decide. As noted above, level of risk to participants may be difficult to control in some IMR designs, given researchers' lack of direct oversight over participants' behaviour, mood or identifiability. This, along with the ubiquity/accessibility of the internet and the data on it, may have implications for procedures around valid consent, withdrawal and debriefing as well as protection of participants. Additionally, emerging new methods may give rise to novel ethics issues and this point should be kept in mind when considering the novel methodological opportunities afforded by IMR.

A summary of the main ethics issues for researchers and RECs to consider when designing, implementing or assessing an IMR study can be found in Table 1.

Table 1: Summary of the main ethics issues to consider when designing, implementing or assessing an IMR study.

Principle	Considerations
Respect for the autonomy, privacy and dignity of individuals and communities	*Public/private distinction* – The extent to which potential data derived from online sources should be considered in the public or private domain; *Confidentiality* – Levels of risk to the confidentiality of participants' data, and how to minimise and/or inform participants of these risks, particularly where they may potentially lead to harm; *Copyright* – Copyright issues and data ownership, and when permission should be sought to use potential data sources; *Valid consent* – How to implement robust, traceable valid consent procedures; *Withdrawal* – How to implement robust procedures which allow participants to act on their rights to withdraw data; *Debriefing* – How to implement robust procedures which maximise the likelihood of participants receiving appropriate debrief information.
Scientific integrity	*Levels of control* – How reduced levels of control may impact on the scientific value of a study, and how best to maximise levels of control where appropriate.
Social responsibility	*Disruption of social structures* – The extent to which proposed research study procedures and dissemination practices might disrupt/harm social groups.
Maximising benefits and minimising harm	*Maximising benefits* – How each of the issues mentioned above might act to reduce the benefits of a piece of research, and the best procedures for maximising benefits; *Minimising harm* – How each of the issues mentioned above might lead to potential harm, and the best procedures for minimising harm.

4. Ethics Guidelines for Internet-mediated Research

We now consider each of the four principles, as outlined in the *Code of Human Research Ethics* (2014) and highlight issues which should be given especially careful consideration in an IMR context.

Principle 1: Respect for the Autonomy, Privacy and Dignity of Individuals and Communities

The *Code of Human Research Ethics* (2014) highlights several key considerations related to this principle, including: *valid consent, withdrawal, confidentiality, anonymity, fair treatment,* and *rights for privacy.* In an IMR context, the issue of privacy is especially problematic and needs additional careful consideration due to the unclear status of different sources of online information which may serve as potential research data (e.g. discussion forum posts, social media site activity). Furthermore, the facilitation of unobtrusive collection of very large data sets, involving the traces of people's online behaviours, in IMR enhances possibilities for analyses which may reveal individuals' personal characteristics, or even identities, which they may have assumed to be private (e.g. by aggregating information from linked data sets, or using statistical techniques to predict sensitive personal attributes).

Closely linked with privacy considerations are issues of anonymity and confidentiality. Also closely linked with considerations of privacy are valid consent, including when researchers should strive to ensure this has been obtained and how to properly gain it, and withdrawal; particularly how to properly implement robust procedures for this in IMR. We now discuss these key issues related to this first principle, offering examples and illustrations.

Privacy online

The *Code of Human Research Ethics* notes that, unless consent has been sought, observation of public behaviour needs to take place only in public situations where those observed 'would expect to be observed

by strangers' (p.25), essentially vetoing observation in public spaces where people may believe that they are not likely to be observed. In an IMR context, the distinction between public and private space becomes increasingly blurred, however. For one thing, much internet communication is conducted in both a private (e.g. the home) and public (e.g. open discussion forum) location simultaneously.

Secondly, in this new medium it is not always easy to determine which online spaces people perceive as 'private' or 'public'; where they might be happy to be observed, or otherwise. To complicate things further, a communication perceived as private at the time might become public at a much later date, should the archived information become publicly accessible (e.g. a posting to a locked social media account, that becomes public when privacy settings are changed as has happened on occasion in the past). While much internet communication is often effectively public through greater visibility, traceability and permanence, it is not always apparent whether this makes it ethically acceptable to use such data freely for research purposes. This accessibility and permanence of the traces of people's online activities, behaviours and interactions raises issues in IMR which are not present in the same way in offline face-to-face contexts which are generally more transitory. Researchers should be aware that participants may consider their publicly accessible internet activity to be private despite agreeing to the terms of the web service providers' End User Licence Agreements or indeed that the communication may have been private when it was first conducted, even if it is now publicly available.

Opinions differ on whether materials posted in so called 'public' (perhaps best thought of as 'readily accessible by anyone') online spaces (e.g. social networks, synchronous and asynchronous discussion groups, etc.) can automatically be classed as public activity. When there is a level of ambiguity concerning whether data are 'in the public domain' or not, researchers should particularly consider the extent to which undisclosed observation may have potentially damaging effects for participants, before making decisions on whether to use such data and whether gaining valid consent is necessary. It is important to note that analysis of online discussions or

other activities is not precluded, but it should be carefully considered in light of the ethics concerns highlighted here. A discussion group moderator or list owner may often provide a good point of contact for advice on the best ways to research existing online groups.

Where it is reasonable to argue that there is likely no perception and/or expectation of privacy (or where scientific/social value and/or research validity considerations are deemed to justify undisclosed observation), use of research data without gaining valid consent may be justifiable. However, particular care should be taken in ensuring that any data which may be made accessible as part of the research remains confidential (often achieved by ensuring anonymity, since dissemination of research findings is, generally speaking, inevitable) – see the further discussion of the possible threats to anonymity and confidentiality in IMR under Principle 4 below. As the chance of violations of anonymity and confidentiality that could harm participants within a given research methodology increases, arguments that valid consent is not necessary are weakened. Essentially, a key principle in IMR (as well as offline methods) is to ensure that ethics procedures and safeguards are implemented so as to be proportional to the level of risk and potential harm to participants.

A further consideration in relation to the use of data deemed to be in the public domain concerns legislative aspects. While personal web pages may appear to be public documents, copyright remains with the author or web hosting company, and indeed many authors ask to be informed if a link is made to the page. In a similar vein, ownership of 'public' content published on social network sites (updates, chat logs, photos/videos, links, reports from activity elsewhere on the web, etc.) often remains with the web service provider, as does the ownership of the 'private' communications between members that are mediated by the web service. Under these circumstances, it may be prudent to consider whether there are multiple entities from whom permission to use online data should be sought (e.g. individual user and web service provider).

While it may in many cases seem impracticable or unnecessary to always gain explicit permission from data owners (e.g. a website

company), these legal aspects should be kept in mind since in some contexts they may be important in protecting both participants and researchers. Strictly speaking, for a document or online trace to be 'in the public domain' it must not be protected by copyright law.

Valid consent

Valid consent should be obtained where it cannot be reasonably argued that online data can be considered 'in the public domain', or that undisclosed usage is justified on scientific value grounds (as set out in the *Code of Human Research Ethics*). Assuring that the principle of participation on the basis of valid consent is fully complied with can raise particular issues in IMR. Obtaining a record of valid consent arguably requires verifying certain relevant characteristics of the person providing it (e.g. to determine that they meet any necessary age requirements). This can be more difficult to achieve in an IMR context than in situations where there is direct face-to-face contact with participants – see the further discussion of this point under Principle 4 below.

Establishing that participants have properly engaged with valid consent procedures in IMR is not always easy, particularly for anonymised questionnaires. As with paper questionnaires, completion of a questionnaire may often be seen as a proxy for valid consent. Provided that an information sheet describes the purpose of the study beforehand and the true nature of the questions that follow, valid consent can arguably be assumed if the questionnaire has been completed, though it is recommended good practice to include a check box (for example) in response to an explicit consent statement (offered both at the start and the end of the procedure). Use of radio buttons or check boxes can also be an effective strategy for allowing participants to indicate that they have read and understood key aspects of the consent information (e.g. their withdrawal rights, how information will be disseminated). Counterbalancing how 'I agree' statements have been worded may help encourage participants to read the information (i.e. to avoid making it easy to simply tick all boxes and proceed). Though care should also be taken not to 'over

complicate' consent procedures online, so that participants who do clearly wish to proceed and participate in the study can easily do so. Overly lengthy consent information pages are more likely to be quickly skimmed, or not read at all. As in all research environments, special care needs to be taken when seeking valid consent from groups whose members may be vulnerable to coercion. Procedures (perhaps necessarily offline) will often need to be used to obtain parent/guardian consent before conducting research with underage or vulnerable participants online.

It is important in IMR, as in any research, that participants providing valid consent are given sufficient details about the study, and the nature of their participation, as well as possible associated risks. Not all of these risks are obvious in IMR, as they can be different to risks that might normally be present in offline contexts. One such risk relates to the levels of researcher control over confidentiality of data, particularly during the data gathering process (for example, where data is stored on the server of a third-party software provider). While it is normal practice (offline) to assure participants of the confidentiality of their personally identifiable information, in IMR the risks for violating this principle can be greater. Researchers need to be aware that it is impossible to maintain absolute confidentiality of participants' personal information gathered online because the networks are not in the control of the researcher. Situations where data are collected in IMR with no potentially identifying information attached are not common. For example, even an IP address stored alongside online survey responses may be linked to an identifiable individual (see the further discussion under Principle 4 on potential risks to anonymity and confidentiality).

Researchers need to consider ways in which participants are properly informed about how the data they provide are electronically stored and transported, particularly where risks are higher (e.g. standard e-mail is a relatively insecure transmission method). Further, participants should be informed about the possibilities for breaches of confidentiality through the use of search engines and the accrual of data from multiple sources. For example, published anonymised

verbatim quotes may be traced to the discussion forum archives from which they originated, where they are likely to be linked to an individual's identity (discussion group posts might be permanently archived). A researcher should be clear about the extent to which their own collecting and reporting of data obtained from the internet might pose additional threats to privacy over and above those that already exist, and whether this might expose participants to potential harm of any sort. Any additional risk may need to be conveyed to participants (particularly where these risks are higher), whilst also taking all reasonable precautions to reduce levels of risk and safeguard the confidentiality of data. Also, participants may not be fully aware of the degree to which their discussion group posts are already available to public scrutiny, so making this clear in valid consent information may be appropriate. As noted above, issues of confidentiality and anonymity are intricately linked, the anonymising of data typically being a way of ensuring confidentiality. Where data are particularly sensitive and/or more difficult to anonymise (e.g. data using detailed personal narratives) then risks to confidentiality increase. Here again the principle of proportionality of consent procedures to level of risk applies; where threats to confidentiality are greater, it might be argued that participants should be carefully informed of the nature of these risks. Similar issues may emerge in the context of considerations relating to the sharing of full research data sets, such as when deposited in research data archives, and should be similarly considered.

Deception

For some research designs, it may be necessary to withhold relevant information or disguise the research question(s) before data gathering (e.g. to avoid contaminating the data and jeopardising the validity and scientific value of a study). This may arguably be seen as involving some level of deception of participants which (depending on context) can raise additional ethics concerns. In face-to-face research, such ethics issues are typically addressed by debriefing participants about the true nature of the research at the end of the study. This respects the dignity of persons by explaining why the study was conducted in this way, and reassuring participants. In IMR there

is an additional risk: that participants may not participate for the full duration of the study and may not be exposed to the debriefing information that could otherwise provide important safeguards. RECs should balance the scientific value of any withholding of information or deception against the risk that participants may discontinue before the disclosure and debriefing (relatively easily done in an online survey), and any likely harm that could emerge in such cases.

Withdrawal

An important element of valid consent is ensuring that participants are aware of the extent of their right to withdraw from participation in a study, and also their right to withdraw data post-participation. Any necessary time limits on data withdrawal should be made clear at the point of valid consent, and any requests from participants to remove their data which are in accordance with these rights should be complied with. In IMR, a number of points should be noted in relation to ensuring that a participant's right to withdraw is not violated. Two key factors, which make IMR approaches rather different to many traditional offline contexts, should be borne in mind:

(a) the typical lack of face-to-face presence between researcher and participants; and

(b) the automated collection of data during the research process.

Together, these factors compound the risk that participants might decide to withdraw from a study without this being obvious to the researcher, and after partial (or even complete) data have already been submitted and stored. Quantitative survey, questionnaire and experimental contexts are prime candidates for this potential risk. For example, a participant may decide to exit a survey or experiment part way through, and do this by closing their web browser. In such situations it may not be clear whether the participant intended to withdraw their valid consent for the use of any data already stored. To use any such partial data could thus violate a participant's withdrawal rights.

Essentially, in IMR such difficulties need to be anticipated, and withdrawal procedures made clear and robust as possible. For

example, displaying a clearly visible 'exit' or 'withdraw' button on each page of a survey or experiment is often good practice. Clicking this would ideally lead to a debrief page and perhaps also a statement asking participants if they require their data to be withdrawn, or whether their partial data can be used (this relates also to the principle of scientific value). Problems will still arise in situations where a participant chooses to exit by closing their browser window, however. Also, some situations make it difficult to implement the 'exit' procedures recommended here (e.g. off-the-shelf online survey software solutions may often not incorporate this functionality). A button at the very end of a study confirming consent to use the data or partial data submitted could help here; arguably, if this has not been verified by a participant then their data should not be used.

The issue of participants wanting to withdraw their data after completing the study must also be considered. Ensuring that participants have been provided with clear instructions and correct contact information, in case they decide at a later point to withdraw their data, is essential. If there is a necessary time limit within which participants can reasonably request that their data are withdrawn then this should be stated clearly in the valid consent information. This time limit should not be unreasonable or aimed at restricting the right to withdraw. However, it is reasonable in cases where, for example, aggregate data may be produced, analysed and then prepared for publication. The possibility of retrospective withdrawal may also require additional mechanisms for storing large data sets in ways that would identify (to the researcher only) individual contributions to the research. This point may not be unique to IMR, but the solutions for tracing individual data that have otherwise been stored anonymously may differ somewhat for this format. Participants could be issued with ID codes to use to identify their contribution, thus allowing their data to be withdrawn if requested retrospectively. Care needs to be taken to ensure that such mechanisms are in accordance with current data protection legislation.

In qualitative approaches, different issues may arise in relation to ensuring participants' withdrawal rights. For example, in an online focus group it is unlikely that a researcher would remain unaware of a

participant's wish to withdraw, but extracting the contributions from one individual from the data set may prove challenging (e.g. other group members may refer to them, or their comments, so simply deleting all the text they submitted may not be sufficient). These issues are not specific to IMR, however. Unobtrusive approaches require particularly careful consideration in relation to withdrawal issues. Although on first impression it might seem that withdrawal issues are not relevant where participants have not given consent in the first place, the possibility and repercussions of individuals finding their contributions (e.g. discussion forum posts, or social media activity) have been used in a research project, and taking issue with this, must be assessed. The enhanced potential the internet offers for the swift, broad-reach publication of research data and findings, as well as the enhanced access to traces of personal (online) activity for use as potential research data, may increase such possibilities beyond what is normally the case in offline research. Again, the level of risk must be assessed for any proposed research design, and appropriate ethical measures taken proportional to this level of risk. On a legal note, should a person find out that their online posts or traces of activity have been accessed, stored and used as research data, they are likely to have rights under the Data Protection Act to stop these data being processed if they could be linked to them personally. In many cases it is very unlikely that a person will ever find out that their online posts have been used for research purposes. However, this does not preclude the responsibility of the researcher to ensure that maximal anonymisation procedures are implemented (for example, researchers may consider paraphrasing any verbatim quotes so as to reduce the risk of these being traced to source, and participants identified). Here again, the principle of proportionality becomes pertinent: considerations of the level of risk/harm must be weighed up against scientific value, the quality and authenticity of reports of research findings, and possible practical issues too.

Principle 2: Scientific Integrity

In relation to this principle, the *Code of Human Research Ethics* notes the importance of ensuring that a research project meets the criteria of 'quality, integrity and contribution'. A noteworthy issue here in IMR is levels of control: in an IMR context the lack of direct physical proximity

may impact on levels of control over and knowledge of participant behaviours, characteristics and research procedures. Such lack of control may have an impact on the validity of a piece of research, its findings and conclusions (for example, particularly in experimental designs where tight control over variables is crucial to validity). Related to this point, the *Code of Human Research Ethics* highlights the potential for harm to arise from the dissemination of inaccurate or misleading information (such as invalid research results and conclusions).

Levels of control

The typical greater degree of 'distance' from participants in IMR can lead to difficulties in maintaining levels of control over research procedures and environment. This may be manifested in not being able to control (or verify):

(a) who has access to participate (as discussed above, and again under Principle 4);

(b) the environmental conditions under which participants are responding (e.g. are they watching television at the same time);

(c) participants' feelings, reactions, responses to the research process; and

(d) variations in the research procedure due to different hardware and software configurations.

Points a, b and d are especially relevant to issues of scientific integrity (*Code of Human Research Ethics*, p.9–10) and are discussed here. Points a and c relate closely to issues of harm and are discussed further below in relation to Core Principle 4: maximising benefit and minimising harm (*Code of Human Research Ethics*, p.11–12).

Regarding variations (between participants) in the participation context and procedural aspects of a study, the key issue is that a lack of control may result in variations occurring that might lead to invalid data and conclusions. This concern is especially pertinent in research designs where tight control over such variations is essential. For example, in a perception experiment it may be crucial to tightly control stimulus presentation parameters

(luminosity, hue, size, etc.); in a memory experiment it may be essential to prevent participants from going back using their browser 'back' button and viewing previous pages. Repeat submissions may also seriously undermine the validity of a piece of research. Data forensics can help in detecting multiple submissions from the same participant (by checking the IP address, browser and operating system information, pattern of responses, etc.). Many commercial online survey platforms incorporate such checks and attempt to prevent multiple submissions as a matter of course. The levels of control required, and those able to be achieved, must be considered for any specific study design when deciding whether an IMR approach can be utilised. Maximising levels of control is possible (e.g. there are now various tools available for implementing IMR studies which adhere to standards which can, for example, control presentation formats between different browsers). Control in terms of knowing who has participated – such as being able to verify crucial demographic information – is also relevant to data validity (e.g. in studies looking at gender differences), but can be hard to verify in practice.

In general, high levels of control over the details of procedural variables (e.g. calibration of presentation parameters such as screen brightness, font size, etc.) will typically be less important for qualitative approaches such as online interviews, and these may thus be less susceptible to the issues raised above. However, it should also be borne in mind that these contexts can often involve more sensitive topics, and thus the need for control in verifying identity must be carefully assessed. Unobtrusive approaches (e.g. analysing server web logs, and other online sources non-reactively) are less likely than obtrusive approaches to be subject to concerns over lack of control, except perhaps for control over security of any data gathered, so as to protect the personal identity of those who contributed to it.

Principle 3: Social Responsibility

The *Code of Human Research Ethics* raises several key points in relation to the issue of social responsibility, including maintaining respect for and avoidance of disrupting social structures, and carefully

considering consequences and outcomes of a piece of research. In relation to the first point, IMR which proposes to make use of existing online social groups (e.g. social networking sites, discussion forums, multi-user virtual environments, etc.) must bear this issue in mind.

The issue of public/private domain distinction online (discussed above, under Principle 1) becomes relevant: intrusions from researchers into spaces considered private by their users may be invasive, unwelcome and socially irresponsible. Where the scientific value of such research is considered very high, this may lead to a researcher needing to make decisions about whether joining a group without disclosure as a researcher (i.e. undisclosed observation) might be most appropriate, in order to avoid disruption and potential harm (e.g. to group levels of trust and cohesion). Thus, this issue interacts with that of valid consent, and the individual research context will need to be considered to decide what is most appropriate.

In relation to the latter point, the enhanced scope for (often automatic) widespread dissemination of and access to data generated in IMR must be considered. For example, a researcher may make use of a 'research blog' as a forum for field notes (e.g. as might be done in an ethnographic study); this could very quickly lead to the dissemination – to a large number of readers – of information about the study, the data collected and, potentially, the participants taking part. Likewise, a researcher participation study in an open discussion forum has similar dissemination potential (discussion forum archives are often readily available for anyone to search, by topic, and view).

Indeed, even the seemingly unproblematic highlighting of the mere existence of a discussion forum in a quiet corner of the web somewhere may be unwelcome to its users. Such issues have relevance to considerations of harm, as discussed elsewhere in this document (particularly under Principle 4 below). It is not necessarily the interventions themselves that are potentially harmful, but their possible scope for compromising the anonymity/confidentiality of participants. Researchers should consider such potential unintended consequences.

Principle 4: Maximising Benefits and Minimising Harm

This principle embodies many of the key points and issues already raised, including ensuring scientific value (maximising benefits) and taking steps to protect participants from any adverse effects arising from the research. Such steps may include gaining valid consent, ensuring anonymity and confidentiality (to minimise harm) and maintaining appropriate levels of control over the research process (to help maximise benefits and minimise harm). As already noted, a lack of control can lead to issues in verifying identity (e.g. determining whether a participant is the minimum age required to give informed consent or detecting multiple submissions). It seems reasonable to propose – so as to not be overly restrictive – that in relation to issues of verifying identity (e.g. restricting participation), a researcher should carefully weigh up any potential harmful effects should a person below the required age (for example) endeavour to and succeed in taking part. Again, the key principle of making ethics checks and procedures proportional to the assessed risks and potential for harm emerges. In high risk situations, researchers should consider whether their research is actually suited to IMR. For example, where research deals with sensitive or adult themes and the age of the participant cannot easily be verified online or under-16s prevented from participating, researchers should consider whether their research is better suited to a face-to-face presentation. In low risk situations it may often be sufficient to take a range of steps which can help minimise the likelihood of successful participation by excluded individuals, such as taking participants who enter age details within a certain range to an exit page from which they are unable to re-enter (even if they attempt to return and re-enter with different age information).

A lack of control may also prevent the researcher from monitoring participants' reactions and behaviours. For example, this may jeopardise the ability to detect when a participant has withdrawn, and thus properly present debrief information. In relation to this point, deception (by the researcher) raises potential for harm in particular, and in an IMR context the lack of direct contact with research participants can mean extra care is needed if deception

is being proposed. Difficulty in monitoring and responding to participants' (potentially negative) reactions to research procedures, compared with proximal, face-to-face contexts, also creates scope for harm. In general, research involving sensitive topics or procedures might be best avoided where levels of control are low and risk is potentially high. Such IMR contexts where levels of control (over who participates, and knowledge of their reactions) are at their lowest would be, for example, an open web-based survey. Procedures such as online real-time interviews, on the other hand, would perhaps offer the greatest levels of control in IMR. The dimension of levels of control must thus be considered in the context of the specific research methods and context.

Threats to anonymity/confidentiality (see earlier comment regarding the relationship between the two) are also relevant to this principle. In some cases it is not always apparent how traceable online data can be. As noted above, researchers should be aware that in IMR it can be relatively easy to trace quotes which have been published from source material (e.g. as often used in conversation or discourse analysis) to individuals' original postings, using search engines, and that this may compromise their anonymity and hence confidentiality. Serious consideration should be given to whether publishing such traceable quotes requires specific valid consent from the individual, and it should be avoided in any cases where possible consequential risk and harm to participants is non-trivial. Some researchers have addressed this issue by suggesting paraphrasing or combining quotes used in publications, and this could be considered if it is consistent with the research design.

Similarly, publishing the name or address of the website or discussion forum from which data were gathered may compromise the anonymity of individuals or have a negative effect on an online community. Where there is such a risk, it may be argued that this identifying information should not be published alongside any analysis of communication sourced from that site. In some cases it may be clear that the risk of potential harm is low (e.g. large, ubiquitous social network sites; quantitive aggregate data analysis).

Additionally, some groups, such as political activists, may welcome the publishing and dissemination of their discussions (though this does not necessarily mean that there are no risks for group members in doing so). In less clear-cut situations, researchers considering naming the location from which their source material was drawn should discuss it with moderators or other gatekeepers of those web services, and take their insights into consideration. The pseudonyms used by posters to web services (communication forums, blogs, chat rooms, social networks, etc) should be treated with the same respect as a researcher would treat a person's real name.

The lack of researcher control over confidentiality of participants' identifiable data (despite taking all possible precautions) was mentioned above under Principle 1. For example, law enforcement agencies may subpoena research data. E-mailing research participants can also be problematic. When recruiting participants by e-mail researchers need to be aware that the security of unencrypted e-mail is low, and e-mail content can be inadvertently disclosed on the internet, local and other computers. Therefore, even the common practice of e-mailing research participants can, in principle, be problematic.

Psychologists risk breaching participant confidentiality if they use non-secure e-mail in research or practice, and participants themselves may often be unaware of these risks. Additionally, e-mail content may be stored by web hosting companies on several computers.

There are potential threats to anonymity in some IMR contexts, for example, where mechanisms for paying participants may use information which makes participants personally identifiable, such as an e-mail address. Even for IMR questionnaires where data are anonymous in the sense of not containing names, addresses or other direct identity information, researchers should be aware that there may be residual risks that participants can nevertheless be identified. As with questionnaires administered in face-to-face studies, combinations of demographic variables may permit identification (e.g. area, income, occupation, age). RECs may sometimes request a complete listing of the data that are to be gathered from each

participant, such as a set of survey questions or an interview schedule (e.g. with particularly sensitive research) so as to be able to make an informed judgement about such risks. In some contexts, though not all, it may arguably be appropriate and necessary for a researcher to provide this. Researchers should also remain aware that despite research data sets appearing to contain no personally identifying information, it is possible that personal identities may be revealed by comparing data sets with other linked sources which do contain such information. Also, additional sensitive information about individuals may be inferred from data sets. This may be achieved, for example, by interrogating traces, such as web browsing history or social media site usage, to statistically predict sensitive personal characteristics (e.g sexual orientation, or political views) that have not otherwise been disclosed by individuals. Whilst in many cases the risk of leaking personally sensitive data in this way may be very low, researchers should be mindful of these possibilities and take steps to properly assess and reduce any such risks. This will involve properly anonymising research data sets, and carefully considering dissemination and data sharing practices.

Conclusion

In closing, the following points bear repetition. First, the normal principles of ethical research with human participants apply to internet-mediated research, and the basics of ethical practice are not changed. However, the implications of these principles for practice may differ in IMR contexts, and aspects of online environments may make particular issues salient in ways they have not been in traditional research. Certain ethics principles may be more or less salient in different types of research design, and the procedures researchers put in place should be proportional to the likely risk to participants. When planning IMR one should take into account both the existing methodological literature and the fundamental principles of research ethics.

The British Psychological Society was founded in 1901 and incorporated by Royal Charter in 1965.

The British Psychological Society
St. Andrews House, 48 Princess Road East, Leicester LE1 7DR, UK
Telephone 0116 254 9568 Facsimile 0116 247 0787 E-mail mail@bps.org.uk Website www.bps.org.uk

Incorporated by Royal Charter Registered Charity No 229642

INF206/04.2017