CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	A context-consent meta-framework for designing open (qualitative) data studies
Authors	Branney, Peter;Reid, Kate;Frost, Nollaig;Coan, Susan;Mathieson, Amy;Woolhouse, Maxine
Publication date	2019-05-22
Original Citation	Branney, P., Reid, K., Frost, N., Coan, S., Mathieson, A. and Woolhouse, M. (2019) 'A context-consent meta-framework for designing open (qualitative) data studies', Qualitative Research in Psychology, 16(3), pp. 483-502. doi: 10.1080/14780887.2019.1605477
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://www.tandfonline.com/doi/full/10.1080/14780887.2019.1605477 - 10.1080/14780887.2019.1605477
Rights	© 2019 Taylor & Francis Group, LLC. This is an Accepted Manuscript of an article published by Taylor & Francis in Qualitative Research in Psychology on 22 May 2019, available online: http://www.tandfonline.com/10.1080/14780887.2019.1605477
Download date	2026-11-16 09:48:07
Item downloaded from	https://hdl.handle.net/10468/8079

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

A meta-framework for designing open data studies in psychology: ethical and practical issues of open qualitative data sets.

To date, open science, and particularly open data, in Psychology, has focused on quantitative research. This paper aims to explore ethical and practical issues encountered by UK-based psychologists utilising open qualitative datasets. Semi-structured telephone interviews with eight qualitative psychologists were explored using a framework analysis. From the findings, we offer a context-consent meta-framework as a resource to help in the design of studies sharing their data and/or studies using open data. We recommend 'secondary' studies conduct archaeologies of context and consent to examine if the data available is suitable for their research questions. This research is the first we know of in the study of 'doing' (or not doing) open science, which could be repeated to develop a longitudinal picture or complemented with additional approaches, such as observational studies of how context and consent are negotiated in pre-registered studies and open data.

keywords; open science; open data; qualitative methods; replicability debate; archeology

1 Introduction

Whether we know it or not, qualitative researchers in the UK are entangled in changes encouraging us to make our research record accessible to others. There are two versions – the short and long(ish) – of the back-story or contemporary history that have led to our interest in open data and qualitative methods. In this paper, we report a telephone interview study of UK-based psychologists to explore the ethical and practical issues that arise when utilising open qualitative data-sets. Before describing our study, we give UK context by outlining the two histories of open data.

1.1 The short-history: a crisis in experimental psychology that affects us all

The short version is that Western experimental psychology is in a crisis because the revelations of research fraud (Levelt et al., 2012, Hammersley, 1997) led to a wider recognition that questionable research practices are the norm (Neuroskeptic, 2012). This was confirmed in the replication of 100 psychology experiments by the Open Science Collaboration (2015); 97% of the original studies reported significant effects but only 36% were replicated. The replication of a finding is foundational to the hypothetico-deductive method in (experimental) psychology. Yet as a discipline, psychology has valued novel and significant findings over replications. Experimental psychological research has moved to share study design before data collection (pre-registration) and the data is made available for corroboration of findings. Consequently, the discipline of psychology is changing in the UK – British Psychological Society (BPS) accreditation standards have been updated to allow open science dissertations (2016), the Peer-Reviewer's Openness Initiative (Morey et al., 2016) is calling on peer-reviewers to give 'non-comprehensive reviews' where materials and data are unavailable (or their unavailability is unexplained), and the last three BPS annual conferences (2016-18) have featured sections on open science – but qualitative research has

to-date, not been part of the conversation. The two events by the BPS on open science – *Replicability and Reproducibility Debate*, 26th May 2016 and *Moving Psychological Science Forward,* February 2018[i] - made no mention of qualitative methods. Nevertheless, systemic changes in the practice of psychological science will impact everyone. It is no surprise, then that qualitative researchers are increasingly being requested to share their data when submitting papers to journals (see, Branney et al., 2017) and, even if they can legitimately claim that data cannot be shared to protect participants' privacy, these requests can still feel like demands.

1.2 The long(ish)-history: data as common property

There has been an international move towards open data that chimes with the UK neoliberal context, where free-market policies exist alongside (although often against) notions of common property. The contemporary need to actively manage and share research findings was arguably crystallised by geophysicists in the 1950s (Korsmo, 2010). In planning to take synchronous measurements at sites globally 1957-58, the geophysicists anticipated they would need to manage this "torrent of raw data" (ibid., p. IGY58). The Committee on the Availability of Data was established and recommended the creation of World Data Centres so that "data would be collected and made available to any scientist without condition except for the cost of reproduction and mailing" (ibid. p. IGY57). That information should be shared fits with the notion of common property in the UK, which is perhaps best exemplified through the National Health Service, which is (mostly) free at the point of use. This is particularly so for research funded partly or wholly through taxation, which to some extent explains why a government science funder – the Economic and Social Science Council – was key in establishing the Qualidata archive (Corti et al., 1995, Corti & Thompson, 1998, Hammersley, 1997). While not unique to the UK, there is a long tradition

of making census data and nationwide surveys freely available, particularly the British Crime Survey and the General Household Survey.

More recently, the Organisation for Economic Co-operation and Development (OECD; The Oecd Committee for Scientific and Technological Policy, 2004) agreed to work towards making publicly funded archive data 'open' to access. Subsequently, UK funding bodies and research governance processes are changing practices and require researchers to make their data openly available to, for example, avoid over-researching participants, increase impact, and facilitate secondary analyses. The UK Research and Innovation (UKRI; previously Research Councils UK, or RCUK) policy on the open access of publications "considers that the 'content' of a paper includes, but is not limited to, the text, **data**, images and figures (Research Councils Uk, 2013, p. 4, emphasis added). Additionally, there is a concordat on open data between higher education and research funders in the UK in which the Minister of State for Universities and Science explains that, in the context of 'taxpayer-funded' research he sees "open access to research data as a fundamental good" (Higher Education Funding Council for England et al., 2016, p. 2). In line with these changes, governance processes, such as the NHS system of Research Ethics Committees, are considering how they can balance the privacy of participants with the need for openness in "maximising the value of research data collected from public funding" (Bishop, 2016, p. 2). While there have been debates around 'secondary (qualitative) analysis' in the UK (Hammersley, 1997, 2010b, 2010a, Heaton, 1998, Heaton, 2004), there is a qualitative difference because moves towards open science are systemic and will effect all researchers. This means that qualitative researchers applying for funding or ethical review are being asked to consider if and how they will make their findings available to others (and may have to request funding for the costs associated with open data, such as preparing data for a data

repository); and those applying for jobs or promotion may be asked if they have any data archives.

While qualitative research is likely to raise "legitimate sensitives around data derived from potentially identifiable human participants" (Research Councils Uk, 2013, p. 4), these can be negotiated in the context of open science. Indeed, there are a wide range of qualitative studies in the UK available from the UK Data Service. For example, interviews with men about their transition to fatherhood (Henwood, 2012); weekly diaries from over 18 months, supplemented with interviews and focus groups, as a community responded to foot and mouth disease outbreak (Mort, 2006); and transcripts of naturally occuring telephone calls about neighbour disputes (Stokoe & Edwards, 2009). This means that qualitative researchers need to explore the 'legitimate sensitives' raised by their research and how they can negotiate them.

1.3 Aim

The replicability crisis in psychology and the gradual move to re-appropriate raw data as common property means that qualitative researchers in the UK are increasingly facing questions about what this means for their research. Within this context, qualitative researchers in psychology are already doing open science but to date this has been as individuals or small groups. Our aim in this paper is to explore the ethical and practical issues psychologists are having with open qualitative datasets. To achieve this, we conducted a telephone interview study. While the interviews also explored pedagogical issues, this paper focuses on ethical and practical issues.

2 Method

Registered with the Open Science Framework[1], the study aimed to gather views on the pedagogical, practical and ethical use of secondary data in research and teaching from qualitative psychology researchers based in higher education institutions across the UK. Individual semi-structured interviews were conducted by telephone with eight participants and data were analysed using a three-step inductive framework analysis (Branney et al., 2012).

Potential participants were initially identified using a systematic and purposeful search of pedagogical, psychology and qualitative methods journals for UK based authors of research articles that indicated use of secondary datasets and of those supported by the Economic Research Council secondary data analysis initiative[ii]. Criteria for inclusion in the search were:

1. Conducting a study where qualitative data were to be shared;
2. Conducting data analysis from an existing qualitative data set; and/or
3. Integrating a qualitative data set in to a BPS accredited undergraduate or postgraduate teaching course

These authors were then contacted by members of the Qualitative Methods in Psychology (QMiP) committee to invite participation and to snowball recruitment from their networks. The final number of participants was eight. In determining the sample size for this exploratory study, we acknowledge that recruiting qualitative researchers as participants can increase the likelihood of accelerated data collection due to potential for skill equivalence in the context of method and interviewing between the interviewer and participant. Halting

[1] doi.org/10.17605/OSF.IO/Y9VSR

recruitment at a sample size of eight participants reflected participants' ease and willingness to go beyond talking about their own experiences and to integrate insights and experiences about the wider qualitative research 'community', to which they belong. Morse (2001) calls this 'shadowed data', where participants discuss the experiences shared by 'others' in similar positions and/or frames of reference.

The participants, six women and two men, self-identified as psychologists working in qualitative research and/or teaching in a UK university (see Table 1). The interviewees were involved in research in a variety of areas, for example, cancer management, educational psychology, health, parenting and wellbeing. Five interviewees had conducted qualitative analysis of a pre-existing data set either as a researcher or supervisor of a postdoctoral student. All but one had experience of using secondary data sets in teaching. Only two interviewees' experiences were limited to using secondary data sets in teaching although 'Clare's' experiences of conducting secondary analyses was through supervision.

Table 1: Participants

ID	Experience of conducting research where required to make data available for secondary	Experience of conducting analyses on pre-existing qualitative data	Experience of using qualitative data sets in teaching	Level of consent*
P01 Amelia	Yes	Yes	Yes but not on a course accredited by the British Psychological Society	2
P02 Bill	Yes	Yes	Yes	3
P03 Clare	No	Yes (supervision	Yes	3

		of students' projects)		
P04 Gamila	No	No	Yes	2
P05 Emer	No	No (but done something similar)	Yes	1
P06 Fatima	Yes	Yes	Yes	1
P07 David	Yes	No	No	3
P08 Hana	No	Yes (used PhD student's data)	Yes	3

*(0 = wishes to withdraw data completely; 1 = summary of interview only; 2 = summary and textual quotes; 3 = summary, textual quotes and audio-recording

2.1 Ethics

This study followed the BPS Code of Ethics and Conduct (2018) and Code of Human Research Ethics (2014) and was granted institutional approval from Leeds Beckett University's Psychology Research Ethics Committee. Consent was a negotiated process; step 1 involved providing information and discussion with the researcher; step 2 was discussion immediately before interview and audio recording of consent; and (before archiving the data) step 3 involved sending the critical listening summary to the participant and asking them to sign a form giving copyright to the interview. The copyright and consent form included a choice of three options:

1. Critical listening summary, but **no** direct textual quotations nor audio recording
2. Critical listening summary with anonymised textual quotations, but **no** audio recording
3. Critical listening summary anonymised textual quotations **and** audio recording

While offering to take steps to anonymise the summary and transcript – for example, by changing names of people and places – we also highlighted during consent that participants may be recognized by those familiar with their work, particularly the community of psychologists using qualitative methods. Regardless of their level of consent, participants will be referred to with sex-specific first-name pseudonyms.

2.2 Data Collection

Individual semi-structured telephone interviews were conducted using a topic guide. Research has shown (e.g., Holt, 2010, Stephens, 2007) that the absence of non-verbal cues in telephone interviews can minimize power differentials arising from visual perceptions of dimensions such as age, class and gender. When both researcher and participant are familiar with the telephone as an instrument of communication the lack of visual cues can encourage ongoing narration by the participant. We were confident that the interviewers and the academic participants were familiar with using the telephone in a work context and that this familiarity would facilitate a rapport comparable to a face-to-face interview.

The guide aimed to explore interviewees' views on the pedagogical, practical and ethical use of secondary data in their research and teaching and included questions such as 'Do you think there are any pedagogical advantages or disadvantages to the practice of using a pre-existing data set in teaching?'; 'To what extent did the research questions differ in the secondary analysis from those in the primary analysis?; 'Can you tell me about the process of gaining consent for the future use of participants' data?' Interviews were audio-recorded and conducted between June and August 2014. They were carried out by members of the research team recruited specifically for this purpose ('interviewers'). The average duration of the interviews was 48 minutes.

2.3 Data Analysis

A three-step inductive framework analysis was conducted on the data to enable collaboration between analysts and incorporation of the interaction between interviewer and participant (see e.g., Branney et al., 2012). In defining a 'theme' in this framework analysis, we follow Madill, Flowers, Frost and Locke (2018) in taking Braun and Clarke's (2006) four discreet dimensions of a theme as dialectics or two opposing aspects of a pole that are in conversation. In this study, a theme is conceptualized as, first, providing a rich description of the interviews rather than a detailed analysis of one particular aspect (although the use of quotes highlights particular aspects and the two themes prioritise aspects of the data over others); second, a theme is inductive rather than theoretical because they closely resemble the interviews; third, they were semantic in trying to describe what participants said rather than finding latent or interpretative meaning; last, the themes are presented through a critical-realist and humanist approach in which the interviewees accounts were taken at face value although within a wider framework through which the 'doing' of (open) science is socially constructed.

In this study, a 'critical listening summary' was produced by the interviewers as part of the first step. This enabled a focus on the responses to questions in the topic guide and incorporated the interactions between interviewer and interviewee. Interviewees' responses were summarised with illustrative quotes and additional notes on their construction made by the interviewers.

In the second step, interviewers worked with other team members to reach consensus on potential key themes across the data set. This was done by first clustering the summaries and notes produced by the interviewers and then further refining potential themes through re-clustering and discussion of the notes. In the third step the team

members returned to the audio recordings of the interviews and added further detail to the analysis by seeking out aspects of the accounts that illuminated and/or challenged the identified themes. The final set of themes identified pedagogical, practical and ethical benefits and challenges of using and sharing secondary data and how this approach to research is integrated into the teaching of research methods. In the following section we present key themes around practical and ethical issues.

3 Findings

All interviewees talked of being broadly supportive of the principle of open data and open science more generally. This may reflect a self-selection bias with those uninterested or opposed to engaging in this research choosing not to take part in the study. Interviewees variously talked of open data usage as *good practice*, *timely* and *progressive, making best use of the participants' and researchers' time and funders' resources*, as well as *making best use of limited resources when funding is scarce,* and/or *participants are few, over-researched and/or difficult to engage in research*.

Talking through the possibilities for the reuse of data, interviewees considered the potential to conduct studies that combine secondary and primary sources of data with multiple 'big data' secondary data sets, akin to a systematic research review or meta-analysis. This potential development of data sources encouraged the interviewees to consider the need to extend the boundaries of existing research with the use of more novel, under-utilised and efficient methods of data acquisition.

Interviewees defined secondary data either in terms of the purpose of the analysis or in terms of who collected it. In relation to the 'purpose of the research', the (re)analysis was described as *secondary* to the *primary* research aim, with the data being "*re-purposed*" or "*used again for a different research aim, to answer different research questions"* (Clare). In

relation to 'who collected it', secondary data was also described as data collected by '*someone else'* such as when a researcher joined a study team after data collection. In this case the analysis is conducted in line with the original study aims but the researcher has a different (or *secondary*) relationship to the data (and the participants) compared to the *primary* researchers who collected it. For example, Hana was involved in the analysis of data collected by her PhD student; this data had not been utilised by the student or presented in their thesis.

The two ways of looking at secondary data means that the same researcher can use data they collect for a secondary purpose, thereby adding additional value to an already 'rich' dataset. In such cases, interviewees saw the similarities with returning to one's own data with new research questions to support secondary analysis. This shows that some aspects of open data are already commonplace. Finally, interviewees talked of data that is in the public domain, such as archive materials, and company and governmental data not originally produced for research purposes, as being another source of secondary data.

3.1 Constructing Themes

In relation to the use of secondary datasets suited to qualitative analysis, two issues central to the principles of professional ethics and conduct emerged. Context and consent were identified as two descriptive themes through the Framework Analysis. The two themes overlap considerably, so rather than being distinct they are differentiated by their perspective; context largely took the perspective of the researcher (our interviewees) collecting and analysing data whereas consent foregrounded the level of informed consent that participants in primary studies could give.

3.2 Open Data Context and Researchers

Interviewees talked of a lack of information about the circumstances under which the research data was originally collected. For example, in reference to researchers conducting secondary analyses, Clare commented that they

> *"miss a layer of interpretation that you get being there in the moment or being able to listen to the audio […] so you become one step removed. The interviewer's influence and how the interview was co-produced may be lost" (Clare)*

In particular, issues of 'place' and the original context surrounding the research process were regarded as pertinent issues for the re-use of open data. While this primarily related to information that would be important for the analysis or interpretation of the data, it also related to the context in which research occurred. For example, Fatima said that particular styles of interviewing and the questions asked lend themselves more to specific types of data analysis (e.g. Interpretative Phenomenological Analysis) as opposed to others.

Context related to the perceived (in)completeness of the open data and lack of associated supporting information. Interviewees were concerned that not knowing whether they had sufficient information risked imposing unintended meanings onto the data. For instance, where data has been transcribed, the method of transcription could vary in where the presentation of pauses, intonation, etc. were inserted. For example, Amelia said

> *"Somebody might be laughing their socks off at something you think is quite serious!" (Amelia)*

Even when original audio and video recordings were available, researchers may lack the time to work through them in sufficient detail due to funding deadlines or not including sufficient time in the original research design for this additional work. Emer talked of having access to the audio recordings but predominantly relying on the written transcripts because of insufficient time to listen to the recordings given the vast number of interviews conducted in the research (approximately 100)[iii].

Interviewees worried that without access to the original research context underpinning the elicitation of the data they may misinterpret or misrepresent their participants. Indeed, Emer described feeling ill-informed about the conception of the research, sampling of participants and collection of data even though she was working with the study team. In relation to her experiences of analysing interview data that she did not collect herself she suggested that there would not be the same level of engagement with the data, and she would lack detailed knowledge about the interview as an interactive process.; the transcripts provided a record of *what* was said, but not *how* things were said.

Amelia elaborated that while there is the context of data, methodological details may be missing from open data such as style of questioning, and level of expertise of the interviewer:

> *"The challenges are, if you don't have access to the audio you can really misunderstand. Erm, it can be quite frustrating if the interviewer doesn't follow up some stuff that you'd have liked to have been followed, or…hasn't interviewed very well" (Amelia)*

It is possible that even those involved in primary data collection may overlook this information when returning to it but when working with secondary analysis the challenges

of accessing and considering context are arguably greater. As Amelia pointed out there is also the context of participants' lives, which no study fully captures, and as Emer put it, there are "*layers*" of context.

The issue of context for secondary data was also seen in more productive terms where interviewees considered the possibility that open data could be *sources* of context. That is, they could turn to open data to access layers of information about different topics, particularly historical or longitudinal ones, through the questions being asked of the data. Such changes may be seen in the conduct of science and the development of the discipline of psychology over time. Emer commented that open data sets would offer a useful source for gaining an historical perspective on her area of research which concerns parenting advice and practices.

In considering the wider context of research, interviewees talked of funders. Where funding is limited or unavailable, open data provides one way of making best use of the time and effort put into collecting data. Clare discussed the use of public funding and ensuring it is used to its fullest potential:

> *"Of course there's the advantage of, erm, kind of, public funding being used to its fullest potential as well, so money that's been pumped into one study goes a bit further if you use the study again, and, erm, other advantages are of course it saves me time as a researcher 'cos I don't have to conduct these interviews again, and erm…it saves the environment slightly because of all the travel costs as well"* (Clare)

Also, Fatima said that she wouldn't get any research done due to the difficulties in obtaining funding if it were not for the opportunity to re-use data from her PhD.

Amelia highlighted the way in which data is collected depends on the design of the study. How interviews are transcribed, for example, depends on the type of qualitative analysis being employed, which means that two studies on the same topic may be unable to use each other's data because of the approach each has taken. Additionally, conducting a study that intends to make its data open requires planning to ensure that materials are collected in a way amenable to archiving to suit funder requirements while also ensuring that the research integrity of the primary research team can continue to be upheld beyond the life cycle of the original phase of data collection. Many, if not most, qualitative researchers may, for example, rely on a variety of notes in their reflexive journal for their analysis and will have to consider if and how this type of 'data', instrumental to the research analysis, should contribute to the open data archive.

The issue of power located within the theme of 'open data context' allowed us to explore the relationships between the broader community of researchers. For example, Emer, talked of fearing negative peer scrutiny in sharing their data, particularly when it was perceived that others could listen in to how they asked questions and responded to participants in interviews. In a similar vein, Amelia said

> *"It can be a bit embarrassing how badly I interview [laughter] no, I don't think I'm that bad [laughter]…but there is that thing like 'oh goodness I didn't really say that did I? [laughter]" (Amelia)*

Interviewees also considered their power in asserting ownership of the data they collected and negotiating access with other researchers. This included whether secondary studies should seek permission from the primary study researchers and indeed the primary

participants. Amelia reported analysing data from a colleague's research and signing a confidentiality agreement to view the data. Fatima wondered if 'secondary consent' forms that set out the limits of use of the data to other researchers would become a necessity. Such forms may ask for explicit permission for data use in further analysis, possibly by other researchers and with different research questions.

Creating open data means that researchers can perceive that they are 'giving up' power of ownership, which may mean, for example, they are no longer the only ones who can analyse and benefit from the data, such as for publications, grant applications and promotions. Indeed, this may mean a professional split between those who collect data from those who may go on to re-analyse it.

3.3 Open Data Consent and Participants

All interviewees talked about open data in relation to participants' consent. Understanding consent as an ongoing process rather than a one-time mutually informed agreement was problematic for the use of open data. This issue required interviewees to re-examine their professional responsibilities and to consider whether they would be maintained to the highest standard if their own qualitative data moved to open access. Their concerns extended to their own experience of using secondary data.

Consent could also be seen as productive within an informed process of consent because it makes open data possible, enabling its use for possible future analyses not anticipated at the time of study design or data collection. David gave an example of consent being given by participants for their data to be used 'for research purposes', without specifying that, at a future date, it may be other researchers making use of that data. Similarly, Amelia, talked about a basic consent form she had used for years and noted that

"Several years ago I added a criteria to it for consent which …which is along the lines of 'I consent to this material being shared with other researchers on the condition that my anonymity is maintained' – and nobody has ever said 'no' to that one" (Amelia)

She then goes on to say that this allows her a lot of freedom to collaborate with others in future work, something she's very keen on.

This idea of seeking permission for some yet unspecified other form of publication presents problems for some of the interviewees. Emer expressed concern that even if consent for future generic reuse of their data is given, this does not necessarily follow that there is consent for a future specific study. Instead, participants could be given the option of consenting to the study in question and to making their data open for other research.

Interviewees highlighted the importance of thinking about consent in relation to data management at the study design stage, particularly because what can be offered in terms of data sharing and anonymization will depend on the research aims and approach and resources available. Hana was concerned that participants' consent was based on whether they trusted the researcher; the rapport and professional standards which the researcher conveyed, rather than the details and mechanisms of the study and the data it relied upon. This highlights the work required in gaining permission, such as showing participants examples of how their data might be shared and/or getting consent in clear stages so that they have time to reflect on their involvement.

Participants described potential scope for nuanced and differential consent rather than the currently commonly used limited consent of merely taking part/not taking part/withdrawing. For example, a participant may consent to a study but decline sharing

their data; or they may want to share their data in one format but not another; or participants may ask , during an interview, for example, to exclude something they just said (regardless of whether they consent to open data). Clare explained:

> *"We actually video our interviews, we video and audio record them and then participants have a choice over whether they want us to use the video extracts or audio or just written clips and they can be completely anonymised as well" (Clare)*

Interviewees talked about an implied or explicit contract with their participants being both productive and limiting. These responsibilities were talked about as 'productive' by the interviewees because of an imperative to do 'justice' to the time and effort of taking part in research. As Hana put it,

> *"it just seemed wrong not to do something with it. These women…some of these women are no longer with us either; they had given up their time, a very traumatic time of their life to take part in these focus groups" (Hana)*

Recognising the 'limiting' aspect of the contract, interviewees were wary to protect against misappropriation of data, such as how quotes being taken out of context in the popular media to assign the participant a certain membership category that might pathologise them, such as 'illegal immigrants'. This misappropriation could also include other research where the approach may be counter to the implied or explicit contract. Clare, for example, talked of the importance of faithfully representing participants' accounts, which meant that she would be keen to have an input to additional studies to ensure this responsibility is met. However, counter to this sentiment was her acknowledgement that she

has no say over who gets to access and re-use the data, and/or how it is used because her university regards it as intellectual property.

4 Discussion

Using a framework analysis, our aim in this study was to explore the ethical and practical issues psychologists have with what we have referred to in this paper as 'open data'; that is, with increasing expectations of data accessibility. We identified *context* and *consent* as overlapping descriptive themes that were differentiated respectively by whether they took the perspective of the researcher or of the participants in primary studies, which mirrors debates on secondary qualitative analysis (Hammersley, 1997, 2010b, 2010a, Heaton, 1998, 2004).

The hypothetico-deductive research cycle is a useful way of exploring how context and consent relate to open science. From an EU funded project to normalize open science, Figure 1, shows how science can be 'open' at each stage of the research cycle. Considering context and consent throughout this cycle, concerns were raised that by using data for a different purpose to that for which it was collected threatens a distortion or misinterpretation of participants' data, to which they may not have given fully informed consent. There is concern that secondary data analysts will not necessarily know how the data was collected and may be bringing different techniques to its processing.

The 'context' in which the primary data is collected forms part of the epistemology of any qualitative study. It is only by considering the coherence between the method, theory and researcher engagement brought to the research process of the study that its trustworthiness can be evaluated. The positivistic emphasis on 'replicability' of studies carries an inherent assumption of an objective stance being taken by the researchers, and a

quest to replicate results as well as process. In qualitative studies however, the centrality of the researcher means that it is the awareness and consideration of reflexivity that is of greater importance than seeking consensual outcomes. Mauthner, Parry and Backett-Milburn (1998, p. 742) argue that full data sets are made up of both contextual 'background' data, and data collected during interviews. Further, they add that to regard contextual data as distinct from interview data is a 'false distinction'. Contextual data will vary across all qualitative studies, even different researchers using the same method will construct a different context for the data collection and analysis, its interpretation and presentation (Frost, 2016). Therefore, when considering the reuse of data, whether by the original data collector or by subsequent researchers, it is essential to consider the contexts of both original collection and its reuse. In the present study, the researchers' reflexive engagement with the process along with that of the participants has been made as clear as possible to potential audiences and future users of the data. By considering the historical context of the data alongside its contemporary context when it is reused allows a fuller and development of a more dynamic picture of the topic.

Pluralistic qualitative researchers commonly adopt this stance, recognizing the value to the study's quality and evaluation of making as clear as possible all the contributing factors to the 'background data' so that it can be incorporated into the overall data analysis (Frost et al., 2010). By interrogating the data from different perspectives, whether they be methodological, or analytical (Clarke et al., 2015) meanings within it can be of use to different stakeholders, assumptions brought by different methods can be used to ask different questions of the data, and interpretations that are meaningful to a diverse audience can be made. Tensions between differing epistemologies are not regarded as

problematic but as offering insight from different perspectives to the complexity of human experience and meaning-making.

Whilst it may be that data used in a secondary capacity cannot answer questions that were not originally asked of it (Mauthner et al., 1998), pluralistic researchers also argue that adopting varied perspectives on interrogating data can enable new questions to be raised about the research focus. New meanings found in the data are regarded as pointing to new understandings about a topic rather than as discardable 'divergences' from a consensus. Tensions in epistemological differences are regarded as valuable sources of insight. By adopting a critical realist view of the themes identified in this study and a social constructivist view of the way the study was conducted we have allowed for the 'false distinction' (Mauthner et al., 1998, p. 742) to be blurred so that the reflexive engagement of the original data set is included with the interview data. Regarding the re-use of data more widely, considering context in this way serves to underpin the quality of both original research and any subsequent work from it.

The storage of data and its 'results' raised questions about confidentiality and anonymity, and over the long-term allows for obscuring of original meanings given in response to questions focused on a particular issue or asked from a particular perspective. Publication of studies based on secondary data analysis may overtake intentions of the primary researchers, effectively using their data to carry out research they had planned to do themselves. For the interviewees in this study it raised questions of usage rights, author credit, data storage and publication. Being qualitative researchers, the interviewees in this study identified the added layers of ethical sensitivity that their research requires, including consent as fully informed and as an ongoing process rather than a one-off tick box task.

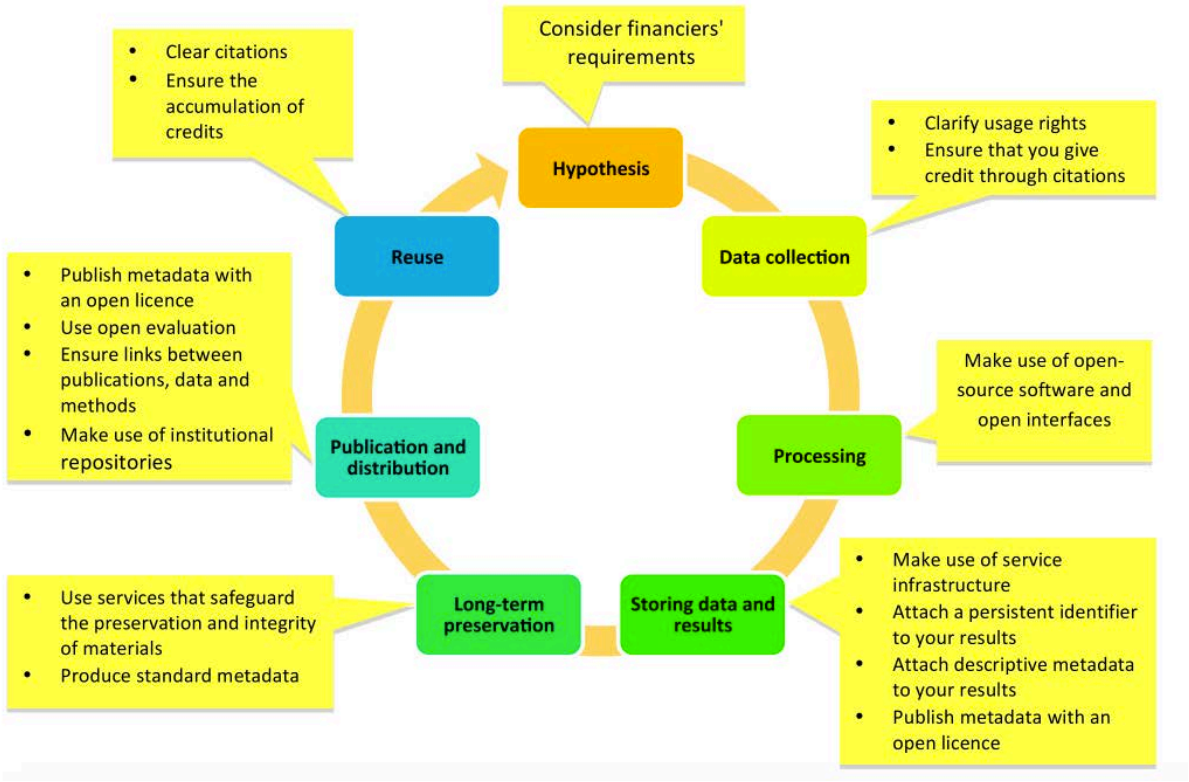


Figure 1. Promoting openness at different stages of the research process (from https://www.fosteropenscience.eu)

As context and consent were key issues, we want to offer them up as a meta-framework (see Figure 2) for thinking through open data that can be used alongside other guides, such as codes of conduct and research ethics. Context and consent provide a set of questions that will depend on whether it is primary (collecting data that could potentially be open) or secondary (a study that could potentially use open data), although we can envisage studies that utilise both. In primary studies, researchers will need to consider what information can and/or should be collected about the context and, given the research aims, whether it is appropriate to use the resources available in this way. Taking an interview study, for example, context could be elaborated through participants' and researchers'

biographies and their reflections on interactions with the project (e.g. response to an interview or launch event). Context will also extend beyond the study – such as popular media and Government responses– and may or may not need collecting. In relation to consent, primary studies will need to consider what data could be collected and the stakes and accountabilities in such data. For example, what are the stakes of the participant and researchers in an interview and how could consent be negotiated? Or how do researchers want to be accountable for their research and/or their participants and each other? This may mean that the potential for coercion, should be considered in relation to both researcher-participants relationships and between researchers, such as in team projects where junior researchers are employed on a contract basis after funding is agreed. In relation to accountabilities, researchers will need to consider the responsibilities they want, or are obliged to take up, for the topic and/or their participants and fellow researchers, and how to negotiate consent around them. Where research is common fodder for popular media, researchers may have to consider the risks to data sharing and how to explore these risks with their participants and study partners. Secondary studies may need to conduct an archeology of context and consent in the primary study or studies, triangulating what is within the data archive with what can be found through other sources. As research questions (given the hypothetico-deductive research cycle in Figure 1) are likely to be formed before such an archeology is conducted, this means there is a risk of investing time in a project before learning whether the context and/or consent is consistent with the aims and approach of a secondary study. It may be important, therefore, for such archaeologies to be introduced as pilot or feasibility studies, so that secondary research can be developed in conversation with what is, and is not, available.

	Primary study	Secondary study
Context	What information can and/or should be collected about the context of this study? Given the research aims of the primary study, is it reasonable to use resources to collect this information?	What information is available about the context of the study? Is this information sufficient to allow secondary study to achieve its aims?
Consent	What data are we collecting and what are the stakes (e.g. participant or researcher) and accountabilities (e.g. researcher's commitment to participants to avoid sensationalizing of topic) in this this data? How can this data be shared or archived and what options are available (e.g. video, audio and/or transcript of video)? How can consent be negotiated with participants?	What did participants consent to in the future use of the data from the primary study? Is this consent consistent with the secondary study?

Figure 2. Context and consent meta-framework for open data

This paper provides a research resource for understanding the roles of qualitative methods in the (Western) discipline of psychology and for researchers working in the UK context of open science. While the literature is growing (Fecher & Friesike, 2014), this is the first empirical study we know of that explores researchers' experiences of doing (or not doing) open science. This study utilized telephone-interviews, a resource-lite method, that could be repeated to explore these ethical and practical issues further. This could be complemented by observational and participative studies examining open science. As this study found context and consent to be important, we recommend exploring how ethics and practical issues are negotiated in pre-registered studies; what, and how is, data being shared in public repositories; and if, and how, open data is used in 'secondary' studies.

In concluding, we want to share an anecdote. In the introduction, we mentioned the BPS events on open science and noted that they made no mention of qualitative methods. We omitted the panel discussion, *Moving Psychological Science Forward*, which was part of the fringe programme of the 2018 annual conference. One of us (Branney) was invited as a panelist to include qualitative methods in the discussion; in informal conversations before the event, a common response to hearing that Branney would talk about qualitative methods was that surely open science was not relevant because privacy concerns would prevent data sharing. We have already given examples of qualitative data that has been shared publicly (Henwood, 2012, Mort, 2006, Stokoe & Edwards, 2009) that shows it is possible to negotiate 'legitimate sensitives' (Research Councils Uk, 2013, p. 4) in ways that allow data sharing. Far from a superficial rejection of open data to protect privacy, our findings show a nuanced and considered engagement with both the pitfalls and possibilites. Just as our participants showed a concern about context and consent, we offer context and consent as a framework to think through open data in designing research.

While our focus has been on qualitative data, we would argue that context and consent will be important for a much broader range, including mixed qualitative-quantitative and purely quantitative studies.

5 References

Bishop, L. 2016. *Research ethics and data sharing* [Online]. UK Data Service. Available: https://www.ukdataservice.ac.uk/media/604362/bishop_rec_uel_20apr16.pdf [Accessed 7/6/2018 2018].

Branney, P, Witty, K, Bagnall, A-M, South, J & White, A 2012, 'Straight to the gp; that would be where i would go': An analysis of male frequent attenders' constructions of their decisions to use or not use health-care services in the uk, *Psychology & Health*, 27**,** 865-880, http://dx.doi.org/10.1080/08870446.2011.636443.

Branney, P, Woolhouse, M & Reid, K 2017, The 'innocent collection of details' and journal requests to make qualitative datasets public post-consent: Open access data, potential author response and thoughts for future studies, *QMiP Bulletin***,** 19-23,

Braun, V & Clarke, V 2006, Using thematic analysis in psychology, *Qualitative Research in Psychology*, 3**,** 77-101,

British Psychological Society 2014, *Code of human research ethics*, Leicester, UK, British Psychological Society.

British Psychological Society 2016, *Standards for the accreditation of undergraduate, conversion and integrated masters programmes in psychology*, Leicester, British Psychological Society.

British Psychological Society 2018, *Code of ethics and conduct*, Leicester, UK, British Psychological Society.

Clarke, NJ, Willis, MEH, Barnes, JS, Caddick, N, Cromby, J, Mcdermott, H & Wiltshire, G 2015, Analytical pluralism in qualitative research: A meta-study, *Qualitative Research in Psychology*, 12**,** 182-201, http://dx.doi.org/10.1080/14780887.2014.948980.

Corti, L, Foster, J & Thompson, P 1995, Archiving qualitative research data, *Social Research Update*, 10**,** 1-6,

Corti, L & Thompson, P 1998, Are you sitting on your qualitative data? Qualidata's mission, *International Journal of Social Research Methodology*, 1**,** 85-89, http://dx.doi.org/10.1080/13645579.1998.10846865.

Fecher, B & Friesike, S 2014. Open science: One term, five schools of thought, *In:* FECHER, B. & FRIESIKE, S. (eds.) *Opening science*: Springer.

Frost, N 2016, *Practising research: Why you're always part of the research process even when you think you're not*, Palgrave.

Frost, N, Nolas, SM, Brooks-Gordon, B, Esin, C, Holt, A, Mehdizadeh, L & Shinebourne, P 2010, Pluralism in qualitative research: The impact of different researchers and qualitative approaches on the analysis of qualitative data, *Qualitative Research*, 10**,** 441-460, http://dx.doi.org/10.1177/1468794110366802.

Hammersley, M 1997, Qualitative data archiving: Some reflections on its prospects and problems, *Sociology*, 31**,** 131-142,

Hammersley, M 2010a, Can we re-use qualitative data via secondary analysis? Notes on some terminological and substantive issues, *Sociological Research Online*, 15**,** 1-7, http://dx.doi.org/10.5153/sro.2076.

Hammersley, M 2010b, Reproducing or constructing? Some questions about transcription in social research, *Qualitative research*, 10**,** 553-569, http://dx.doi.org/10.1177/1468794110375230.

Heaton, J 1998, Secondary analysis of qualitative data, *Social Research Update*, Working Paper No. 22,

Heaton, J 2004, *Reworking qualitative data*, Sage.

Henwood, K 2012, Masculinities, identities and risk: Transition in the lives of men and fathers, 2000-2009 (dataset), *UK DATA SERVICE*, SN: 6984, http://dx.doi.org/10.5255/UKDA-SN-6984-1.

Higher Education Funding Council for England, Research Councils Uk, Universities Uk & Wellcome Trust 2016, Concordat on open research data. London: UK Research & Innovation,

Holt, A 2010, Using the telephone for narrative interviewing: A research note, *Qualitative Research*, 10**,** 113-121, http://dx.doi.org/10.1177/1468794109348686.
Korsmo, FL 2010, The origins and principles of the world data center system, *Data Science Journal*, 8**,** IGY55-IGY65,
Levelt, WJ, Drenth, P & Noort, E 2012, Flawed science: The fraudulent research practices of social psychologist diederik stapel. University of Tilburg,
Madill, A, Flowers, P, Frost, N & Locke, A 2018, A meta-methodology to enhance pluralist qualitative research: One man's use of socio-sexual media and midlife adjustment to hiv, *Psychology & health***,** 1-20, http://dx.doi.org/10.1080/08870446.2018.1475670.
Mauthner, NS, Parry, O & Backett-Milburn, K 1998, The data are out there, or are they? Implications for archiving and revisiting qualitative data, *Sociology*, 32**,** 733-745,
Morey, RD, Chambers, CD, Etchells, PJ, Harris, CR, Hoekstra, R, Lakens, D, Lewandowsky, S, Morey, CC, Newman, DP, Schönbrodt, FD, Vanpaemel, W, Wagenmakers, E-J & Zwaan, RA 2016, The peer reviewers' openness initiative: Incentivizing open research practices through peer review, *Royal Society Open Science*, 3, http://dx.doi.org/10.1098/rsos.150547.
Morse, JM 2001, Using shadowed data. Sage Publications Sage CA: Thousand Oaks, CA,
Mort, M 2006, Health and social consequences of the foot and mouth disease epidemic in north cumbria, 2001-2003, *UK DATA SERVICE*, SN: 5407, http://dx.doi.org/10.5255/UKDA-SN-5407-1.
Neuroskeptic 2012, The nine circles of scientific hell, *Perspectives on Psychological Science*, 7**,** 643-644,
Open Science Collaboration 2015, Estimating the reproducibility of psychological science, *Science*, 349**,** aac4716, http://dx.doi.org/10.1126/science.aac4716.
Research Councils Uk 2013, Rcuk policy on open access. London: Research Councils UK,
Stephens, N 2007, Collecting data from elites and ultra elites: Telephone and face-to-face interviews with macroeconomists, *Qualitative Research*, 7**,** 203-216, http://dx.doi.org/10.1177/1468794107076020.
Stokoe, E & Edwards, D 2009, Identities in neighbour discourse: Community, conflict and exclusion, 2004-2006, *UK DATA SERVICE*, http://dx.doi.org/10.5255/UKDA-SN-6259-1
The Oecd Committee for Scientific and Technological Policy 2004, Science, technology and innovation for the 21st century. Meeting of the oecd committee for scientific and technological policy at ministerial level, 29-30 january 2004 - final communique, *In:* ORGANISATION FOR ECONOMIC CO-OPERATION & DEVELOPMENT (ed.).

[i] See https://www1.bps.org.uk/what-we-do/bps/governance/boards-and-committees/research-board/research-board-resources/replicability-and-reproducibility-debate/replicability-and-reproducibility-debate and https://www.bps.org.uk/news-and-policy/moving-psychological-science-forward-videos-replication-event-now-online

[ii] See https://esrc.ukri.org/research/our-research/secondary-data-analysis-initiative/

[iii] Note that this participant agreed that summaries of the interview could be used and not verbatim quotes