

# **‘Am I not answering your questions properly?’ Clarification, adequacy and responsiveness in semi-structured telephone and face-to-face interviews**

*Qualitative Research*

13(1) 87–106

© The Author(s) 2012

Reprints and permission: [sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)

DOI: 10.1177/1468794112439086

[qrj.sagepub.com](http://qrj.sagepub.com)**Annie Irvine**

University of York, UK

**Paul Drew**

University of York, UK

**Roy Sainsbury**

University of York, UK

## **Abstract**

This article considers spoken interaction in semi-structured qualitative research interviews, comparing those that are conducted by telephone or face-to-face. It draws upon recent empirical research that illuminated some of the differences that may be observed between these two interview modes. Methodological techniques drawn from Conversation Analysis were used to conduct a systematic and transparent comparison of the interview interactions, focusing on the spoken interactional devices that researcher and interviewee employ in order to pursue and maintain a collaborative and comprehensible dialogue. The article begins with an overview of previous discussion on the interactional effects of the telephone in qualitative interviews. Here, we find that while instructional texts have traditionally advised that the telephone mode is not well-suited to the task of qualitative interviewing – primarily because the lack of face-to-face contact is said to restrict the development of rapport and a ‘natural’ encounter – researchers giving personal accounts of conducting telephone interviews tend to offer more nuanced or critical reflections on the extent to which the lack of visual cues affects the interaction in practice. Empirical findings are then presented on: formulation and completion, clarification and comprehension, acknowledgement, interviewees’ checks on the ‘adequacy’ of their talk, and the

---

## **Corresponding author:**

Annie Irvine, Social Policy Research Unit, University of York, Heslington, York, YO10 5DD, UK.

Email: [annie.irvine@york.ac.uk](mailto:annie.irvine@york.ac.uk)

duration of interviews. Key findings were that: completion or formulation of interviewee talk by the researcher was more common in face-to-face interviews; interviewee requests for clarification were slightly more common in telephone interviews; vocalized acknowledgements given by the researcher were less frequent in telephone interviews; interviewee checks on the adequacy of their responses were more common in telephone interviews; and telephone interviews tended to be shorter than those conducted face-to-face. The article discusses possible explanations for the findings that emerge alongside consideration of some potential implications.

### Keywords

Conversation Analysis, interview interaction, mode comparison, qualitative interviews, semi-structured interviews, telephone interviews

## Introduction

This article presents findings from a study that explored interactional difference between semi-structured qualitative interviews conducted by telephone or face-to-face. Methodological text books have traditionally advised that the telephone mode is not well suited to the task of qualitative interviewing (e.g. Gillham, 2005; Legard et al., 2003; Rubin and Rubin, 1995). In particular, the lack of face-to-face contact is said to restrict the development of rapport and a 'natural' encounter (Shuy, 2003), elements that are considered to be important for generating rich qualitative data. However, there are potential advantages of using the telephone for research interviews; for example, savings in time and travel costs and greater anonymity around sensitive topics (Chapple, 1999; Kavanaugh and Ayres, 1998; Sturges and Hanrahan, 2004). Therefore, it is not uncommon to find qualitative studies in various disciplines that have conducted some or all interviews by telephone. However, the empirical evidence base on what consequences the use of the telephone may have for interview interactions and the resulting data is currently underdeveloped. As observed by Sweet (2002: 58), '[t]he telephone has found its way into qualitative research processes as a medium for data collection, but its use has not generated the critical discussion that is merited'.

In this article, we do not rehearse the practical and ethical pros and cons of telephone interviews (for an overview, see Sturges and Hanrahan, 2004). Here we focus in on the *interactional* aspects of research interviews conducted by telephone or face-to-face, drawing upon empirical research evidence to illuminate some of the differences that may be observed between the two interview modes. Rather than comparing the substantive content of the interview conversations, our interest is in the spoken interactional devices that researcher and interviewee employ in order to pursue and maintain a collaborative and comprehensible dialogue. We begin with an overview of previous discussion on the interactional effects and implications of telephone interviews, before outlining the methodology of the present study. Several aspects of our analysis are then presented, including questions of researcher responsiveness, interviewees' attempts to seek confirmation of the relevance of their talk and requests for clarification of questions posed by the researcher, and the duration of interviews. This is followed by a discussion of the potential explanations for and implications of the findings that emerge. Note that this article focuses on the use of 'traditional' telephone interviews, which do not feature any video- or internet-based visual technology.

## Perspectives on interactional difference

This section reviews previous literature that has considered the interactional effects that might arise when qualitative research interviews are conducted by telephone rather than face-to-face. Contributions have come from instructional (mainly introductory) text books and also a number of published researcher accounts of first-hand experiences (Carr and Worth, 2001; Chapple, 1999; Dicker and Gilbert, 1988; Holt, 2010; Stephens, 2007; Sturges and Hanrahan, 2004; Sweet, 2002; Tausig and Freeman, 1988) and overview or review articles (Burnard, 1994; Novick, 2008).

The most fundamental difference between telephone and face-to-face interviews is the absence of a visual encounter. This is thought to affect the interaction in a number of ways. Novick (2008: 395) provides a useful categorization of the types of 'data loss or distortion' that potentially result from the absence of visual cues: first, the loss of nonverbal data, namely body language and facial expressions; second, the loss of contextual data, including the interviewee's physical characteristics and the interview setting; third, the loss or distortion of verbal (spoken) data. However, Novick calls into question the assumption that data loss or distortion in any of these respects is *necessarily* detrimental to the interaction or resulting data. Indeed, while the absence of visual information is typically construed in text books as a disadvantage (e.g. Berg, 2007; Cresswell, 2007; Fielding and Thomas, 2008; Gillham, 2005; Robson, 2002), researchers giving personal accounts of conducting telephone interviews tend to offer more nuanced or critical reflections on how the lack of visual cues affects the interaction in practice.

The brief overview that follows is structured around a number of themes that emerge from the literature, regarding the ways in which conducting qualitative research interviews by telephone may potentially affect the social encounter and spoken interaction. These include effects on: rapport and the 'naturalness' of the interaction; comprehension and the transmission or interpretation of meaning; monitoring of responses and emotions; levels of interest and attention; and the duration of interviews.

### *Rapport and naturalness*

A prominent theme in the literature is the scope for developing rapport during telephone interviews. Most often, this is perceived as being more difficult to achieve, requiring enhanced skills or efforts. For example, Fielding and Thomas (2008: 253) caution that 'interviewers need very effective communication skills to make the interaction "natural" while keeping an eye on the interview guide and helping respondents stay on topic'. Meanwhile, Dicker and Gilbert (1988: 68) note that:

Once access has been gained to any potential respondent, either for a face-to-face or telephone interview, it is necessary to gain their confidence and support. *This is especially true when using the telephone.* A telephone conversation with an unknown interviewer can create a considerable amount of anxiety for the respondent. [emphasis added]

While Shuy's (2003) overview of the pros and cons of telephone versus face-to-face interviewing is based primarily on considerations for survey researchers, he makes some

notable comments about 'contextual naturalness' that are also relevant to qualitative research. Shuy (2003: 179) contends that, 'face-to-face interaction compels more small talk, politeness routines, joking, nonverbal communication, and asides in which people can more fully express their humanity'. Echoing this, Gillham (2005: 103) asserts that in telephone interviews, we 'lose much of that empathy, the interpersonal chemistry so vital to generating the motivation and interest of a face-to-face interview'. Chapple (1999: 90) suggests that rapport may be harder to establish in telephone interviews because there is reduced potential for 'symbolic exchanges' such as sharing food or drink.

However, although concerns about the more unnatural nature of telephone interview conversations are commonly raised, there is also recognition that in many cultures today, people are well-used to communicating by telephone both informally (e.g. Gillham, 2005) and in more formal settings, for example, therapeutic counselling or specialist helplines. Sweet (2002: 60) believes that qualitative researchers using the telephone need to 'develop new ways to establish and maintain rapport'. She implies that in her study, she managed to achieve this, with interviewees reportedly expressing 'satisfaction' with the telephone interviews. Similarly, despite his initial scepticism about the suitability of telephone interviewing, Stephens (2007) describes his telephone interviews as being 'largely successful', in most cases achieving equally 'friendly rapport' as in face-to-face interviews.

### *Meaning and comprehension*

Hermanowicz (2002: 497) claims that in telephone interviews 'breakdowns in communication easily arise simply by being apart. Consequently the ability to conduct a meaningful conversation is readily compromised'. It has also been suggested that clarifications may be harder to execute over the telephone (Burnard, 1994). As a result of the lack of non-verbal cues, Burnard (1994: 69) contends that 'the interviewer has to pay special attention to the phrasing and clarity of his or her questions'. Likewise, Stephens (2007) found there was a need for questions to be articulated more clearly in the absence of accompanying physical and facial gestures. Shuy (2003: 181) notes that visual signals from the researcher are important in encouraging interviewees to elaborate or clarify what they have said, and suggests that this can lead to more 'thoughtful' responses. Thus, there is an implication that a lack of visual cues can impact upon the quality of the data generated. There are also suggestions that more subtle meanings conveyed by the interviewee may be missed in the absence of visual cues. For example, Fielding and Thomas (2008: 253) note that 'important nuances may be lost because we communicate by body language as well as speech' and Gillham (2005: 103) comments that 'all those non-verbal elements which are a major part of live communication are missing: a layer of meaning [is] stripped out'.

In a comparison of four qualitative semi-structured interview modes (face-to-face, telephone, email and MSN messenger) Opdenakker (2006) draws particular attention to face-to-face interviews' distinct advantage in providing 'social cues' such as voice, intonation and body language. Opdenakker notes that these aspects can 'give the interviewer a lot of extra information that can be added to the verbal answer of the interviewee' (2006: 7). However, like Novick (2008), Opdenakker recognizes that the importance of these non-verbal cues may vary depending on the research objectives.

### *Monitoring responses and emotions*

Being able to see an interviewee may give more ready access to any signs of confusion, reticence, discomfort or distress provoked by a line of enquiry (e.g. Gillham, 2005; Rubin and Rubin, 2005; Shuy, 2003). Commenting on her direct experience, Sweet (2002) found that 'the loss of non-verbal communication made it very difficult to assess silences and nuances within the interview' (p. 62) and that at times when she was uncertain of what was happening in the interviewee's environment, she would ask for repetition or extension of responses. Other researchers assert that it is possible to detect such emotional states through auditory channels. For example, Tausig and Freeman (1988: 424) found that 'careful listening enabled the telephone interviewer to hear affect that was conveyed with the content without the benefit of visual access' by drawing on such cues as tension, anger, sarcasm, curtness, tears, or rapid speech. However, it is possible that tone of voice may be misinterpreted (Chapple, 1999).

### *Interest and attention*

Visual cues supplied by the researcher may serve a role in indicating attention and interest to an interviewee. Stephens (2007: 211) observes that, in face-to-face interviews, the non-verbal cues and 'small utterances' play a larger role than we might anticipate in continually shaping and guiding the interview; as non-verbal cues are unavailable in telephone interviews, researchers may need to use more explicitly verbalized pointers that are 'less frequent but more directive'. Dicker and Gilbert (1988) also draw attention to the fact that 'encouragement' offered to telephone interviewees must come via 'paralinguistic utterances' rather than the more 'unobtrusive' non-verbal forms of communication that could be used in face-to-face interviews. Miller (1995: 36) also notes how she used vocal responses to communicate her 'continued attention and interest' where non-verbal nods or facial expressions might have been used in face-to-face encounters. Similarly, Chapple (1999) notes the importance of *sounding* interested and concerned when conducting telephone interviews.

Reflecting on both her own and her interviewees' experiences of telephone interviews, Holt (2010: 117–118) highlights the challenge inherent in seeking to limit one's own interjections while providing sufficient feedback to assure interviewees that you are 'still present and listening'. She summarizes this as 'the interactional difficulties in a research encounter where the researcher is deliberately reticent and where there are no visual cues to compensate for this reticence'. In common with the researchers cited above, Holt attempted to compensate for this by contributing lots of 'umms and aahs and yes's' but interviewees in her study still said that more input from the researcher would have been helpful to them in knowing how to form their responses.

### *Interview duration*

Finally, it has been argued that telephone interviewing is more demanding and fatiguing than face-to-face communication (Gillham, 2005; Shuy, 2003; Tausig and Freeman, 1988) and so interviews are likely to be shorter. Gillham (2005: 105) contends that it is a challenge for the interviewer to 'maintain the level of sensitive attention necessary for a

productive interview' and asserts that telephone interviews are more arduous and are therefore inevitably shorter than those conducted face-to-face:

It is extremely hard work to keep going. Because interviewer and respondent have only vocal communication to go on, it requires, if anything, even more concentration than a normal interview. And related to that, irrespective of level of structure, an endurable length of time is less, usually much less than with a face-to-face interview. (Gillham, 2005: 103)

Gillham advises that *at most* we should expect participants to tolerate just half an hour of being interviewed by telephone. However, those providing empirical accounts (Grant, 2011; Sturges and Hanrahan, 2004; Sweet, 2002) have reported that telephone interviews do not differ notably in length or depth from those conducted face-to-face.

The body of literature cited above provides some useful insights into the kinds of interactional difference that may arise when conducting qualitative research interviews by telephone rather than face-to-face. However, few researcher accounts provide detailed descriptions of the analytic approach through which they arrived at their conclusions and text books invariably provide their cautionary advice on telephone interviews with no reference to empirical evidence. Moreover, the majority of researchers report only on their use of the telephone, rather than presenting direct comparisons of interviews conducted via different modes within the same study (notable exceptions being Stephens, 2007; Sturges and Hanrahan, 2004; and Sweet, 2002).

Sturges and Hanrahan (2004: 107) note that 'those who are interested in mode comparisons of qualitative interviews will find relatively little empirical work in this area. Attending more closely to the matter of *quality* as well as quantity of evidence in this area, Novick (2008: 394) offers the following assessment:

Although authors have reported that qualitative telephone interviews may have many advantages and yield high quality data, there is relatively little formal evidence to be found regarding their merits and shortcomings vis-à-vis face-to-face interviews.

In sum, direct, detailed and empirically based comparisons of the telephone and face-to-face interview modes seem, to date, to be lacking in the qualitative methodological literature. We know very little about the way mode affects the spoken interaction and the resulting implications for our data. In this context, a research project was devised that sought to address the question of mode-related difference in a more robust and systematic way than has previously been attempted. Our methodological approach is presented in the following section.

## Method

### *Data set*

The mode comparison study was small-scale and exploratory, using 11 semi-structured qualitative research interviews that had originally been conducted for an earlier study on mental health and employment (Irvine, 2008). For practical reasons, two different

interview modes had been used in the original study: some interviews were conducted by telephone and some were conducted face-to-face.

A subset of these interviews offered a unique opportunity to conduct a systematic comparison of mode-related difference. Three particular strengths of this data set were that (i) the interviews were conducted by the same researcher and so any effects of different personal interviewing styles were minimized; (ii) interviewees were allocated to interview mode by the researcher (largely based on geography) and so the influence of any conscious preference among interviewees for one mode or the other was minimized; and (iii) the interviews were all carried out prior to devising the present mode comparison study, so there was no deliberate attention on the part of the researcher to variation in interviewing practice in each mode at the time the interviews were conducted.

The subset of data used in the mode comparison study comprised five face-to-face and six telephone interviews, totalling just under 17 hours of audio data. Interviewees included males and females, aged between 32 and 54 years. The interviews were conducted in English, in which all interviewees and the researcher were fluent.

The key aims of our study were to contribute to a more evidence-based understanding about what differences interview mode might produce in the interaction that takes place during a qualitative research interview and to consider what implications any such differences might have for the research data. The intention was not to prove or disprove the validity or effectiveness of the telephone mode for qualitative research interviews, but to illuminate whether and how the two modes differ and, in turn, to raise questions about what additional considerations might be necessary if researchers choose to conduct qualitative interviews by telephone.

### *Analytic approach*

The study was primarily interested in the possible *interactional* differences that may (or may not) exist between telephone and face-to-face interviews, rather than the *substantive content* of the interviews. As such, the study drew upon some of the techniques offered by the method of Conversation Analysis. Conversation Analysis examines the structures and sequences of spoken interactions and seeks to uncover the various practices that speakers use to accomplish social actions through talk. The approach to analysis centres on the close examination of collections of audio (and sometimes video) data alongside detailed transcriptions, to identify recurring patterns and structures in interaction.<sup>1</sup>

The 11 interviews were transcribed to a sophisticated level of detail using the Jeffersonian transcription system (Jefferson, 2004b), which represents features such as intonation, changes in volume, speaker overlap, intake and exhalation of breath, pauses and their duration. A guide to the conventions used in this transcription system is included as Appendix. Producing the transcripts in this level of detail allowed for a close examination of precisely what took place in each spoken interaction and for the fine-grained comparison of interactional difference across modes. By drawing upon techniques and concepts provided by Conversation Analysis (such as turn taking, overlap and acknowledgement) we were able to address the question of mode effects in a more detailed and systematic way than has previously been attempted.



However, it should be emphasized that this study *drew upon the techniques of* Conversation Analysis, rather than having truly *conducted* Conversation Analysis on the data. Conversation Analysis itself focuses on the design and sequencing of turns at talk and their interactional consequences to a much greater degree than was done in the present study. The present analysis remained at the level of classification and quantification of *types* of turn or interactional phenomena in a similar manner to the work of Clayman et al. (2007), Heritage and Roth (1995) and Heritage et al. (2007).

It should also be noted that the present analysis focused specifically on the structure or ‘mechanics’ of the interview interaction, rather than the ways in which meaning or identity may be constructed through dialogue. The aims of the present study were therefore in contrast to several previous analyses that have applied the techniques of Conversation Analysis to research interview data (e.g. Baker, 2003, 2004; Rapley, 2001; Rapley and Antaki, 1998; Roulston, 2001; Roulston et al., 2001; Wooffitt and Widdicombe, 2006) but have something in common with the work of ten Have (2004) and Mazeland and ten Have (1996).

Through consultation of the existing literature and exploration of the data, a wide range of potential avenues for inquiry presented themselves. These related to: question construction; pauses and silence; turn transitions; speaker overlap; elaboration and digression; repetition and formulation; duration; floor-holding; comprehension; empathy; emotion; delicacy; and the overall ‘shape’ of the interview, among many others. Within the scope and scale of the study, not all of these proved to be practicable lines of enquiry. However, several themes of interest were pursued and this article presents some of the key findings that emerged.

## Findings

We now present findings of the analysis on five interactional aspects of the research interviews: formulation and completion; clarification and comprehension; acknowledgement tokens; adequacy checks; and interview duration.

### *Formulation and completion*

In interviews conducted in both modes, the researcher sometimes completed an interviewee’s utterance for them or helped them to find a word they were struggling with. In conversation analytic terms, this is described as a ‘collaborative completion’ (Lerner, 1996, 2004). There were also occasions where the researcher rephrased what an interviewee had just said in order to show understanding, a phenomenon described as ‘formulation’ in the conversation analytic literature (Drew, 2002; Heritage, 1985). We acknowledge that formulations and completions do not serve identical interactional functions. A more nuanced analysis would require the separate consideration of these two phenomena. However, for the present purposes, the two have been grouped as representing forms of affiliative action that show understanding of what the interviewee is ‘getting at’. Some examples are given in Box 1, below:



## Box 1 Formulations and completions

## Extract 1.1 Face-to-face interview

- Int:* Imagine coming into work every day. You're feeling .hh nau:seous which i:s a typical (0.2) side effe[ct of these pills
- Res:* [Mm hm
- Int:* so you're feeling (.) ill .H feeling nauseous .hhh >you're feeling very < tired .hh you're feeling very erm anxious and shaky because one of the: .hh effects it has on the central nervous system is to actually make you quite .hh you know
- Res:* → °s-° agita[ted kind of thing [yeah]
- Int:* → [Agitated [yeah] absolutely .hh er::m and so: you're having to cope with all of that: .hhh and come into wo:[rk
- Res:* [Mm hm
- Int:* and just pretend that (0.6) everything's fine

## Extract 1.2 Face-to-face interview

- Int:* I'm: conscious that I've not (0.4) addressed things in a- in: the same way that I would have done [in the past.
- Res:* [mm hm
- Int:* I've not (0.8) gone looking for problems. (0.2)
- Int:* [huhhh. [huh huh huh [.hh
- Res:* [R: ight ye[ah [yeah .h[hh
- Int:* [er::m (.) so I've kinda \*n::ot\* (0.2) just done what I've been asked to do but more (0.4) towards that
- Res:* Mm hm (0.2)
- Int:* type of approach than:
- Res:* Yea[h
- Int:* [mhh (.) than the (0.2) the: er the: go getting type stuff.
- Res:* → Yea:h not making work for yourse[lf.
- Int:* → [Not making [work
- Res:* [huh
- Int:* for yourself.

Taken together, there was a total of 33 instances of formulation or completion. These occurred more often in face-to-face interviews than telephone interviews. Of 33 instances overall, 26 occurred in face-to-face settings (across all five of the interviews in this mode) while just seven occurred in telephone interviews (in four of the six interviews in this mode).

### *Clarification and comprehension*

Interviewees sometimes sought clarification from the researcher. This was done in a number of ways, with subtly different objectives. Most clarification requests related to the specific terms or referents, or to the broader parameters, of a researcher's question. Some examples of this type of clarification request are given in Box 2, below:

#### Box 2 Requests for clarification of question referents or parameters

##### Extract 2.1 Telephone interview

- Res:* Do you ↑want to just start off (0.2) with just giving me a bit of background to yourself to sort of set the scene perhaps. (0.9)
- Int:* To myself huh huh [↑huh huh
- Res:* [huh huh
- Int:* → Er::m (0.4) what you mean like with (0.6) school and → work an::d you [ (know like work) ]
- Res:* [yea::h and your sort of s- ]
- Int:* and stuff [ yeah ]
- Res:* [household] and things like that.

##### Extract 2.2 Telephone interview

- Res:* Do you think it would be useful if there were (0.6) there was someone else a- around at work that you could talk to. (0.2)
- Int:* → .hhhh (0.2) what a designated person
- Res:* Mm: (0.4)
- Int:* .tch ↑er::m (0.2) .HHH (0.9) no because you- I think to talk to people I think you have to have rappo:rt

Overall, far more of these types of clarification request arose in telephone interviews ( $n = 31$ ) than in face-to-face interviews ( $n = 6$ ). However, breaking down the clarification requests by individual interview revealed that one telephone interviewee in particular accounted for the majority of this difference. This interviewee asked very frequently for questions to be clarified (20 times during a 65 minute conversation). This was alongside a number of other features of hesitancy, uncertainty and self-doubt

displayed by this individual during the interview. Taking this into account, the difference between modes reduced substantially. However, there remained slightly more of these requests for clarification of question parameters in the telephone interviews, as compared to face-to-face.

Other types of clarification request, much less common in the data, included the interviewee making an explicit request that the researcher repeat the question, rephrase the question, or reiterate the question (after having strayed somewhat off track). Instances of each in this data set numbered only three or four, and so do not lend themselves to further discussion.

As well as clarification requests, there were a few instances in the data of what we might consider to be 'substantive' misunderstanding. Ten instances were found in the data set, distributed evenly across the two interview modes.

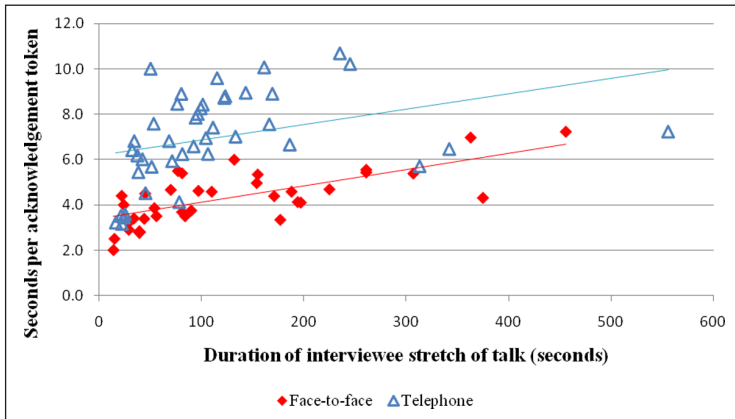
### Acknowledgement tokens

Acknowledgement tokens (also known as *receipt tokens*) include such things as: *yeah*, *mm hm*, *uh huh*, *mm::*, *okay*, *right* and *oh*. These are utterances that take a turn at speaking but do not take over the conversational floor (Drummond and Hopper, 1993). As discussed in the literature review above, such vocalized tokens may be important in signalling the researcher's interest and attention during interviews, particularly where visual options are unavailable.

Other researchers have investigated the specific functions of certain tokens, for example *Oh* (Heritage, 1984) and *Mm* (Gardner, 1997) and also the variation in the functions performed by different tokens such as *uh huh*, *mm hm* and *yeah* (Drummond and Hopper, 1993; Jefferson, 1993). This literature demonstrates that the various types of acknowledgement token are by no means functionally equivalent. However, for purposes of the present study, we did not differentiate the specific functions of each form of acknowledgement token. While we recognize that by grouping all acknowledgement tokens together, we overlook their nuanced interactional functions, what our analysis of researcher acknowledgement tokens allowed us to assess was the frequency with which the interviewee received *some form of* vocalized indication that what they were saying was being heard, understood, or otherwise 'received' by the researcher.

To measure the frequency of response tokens, we took seven stretches of interviewee talk from each interview (the three longest stretches plus four other randomly selected stretches of varying lengths) and counted the number of acknowledgement tokens uttered by the researcher during each stretch. We therefore examined 77 stretches of talk in total (seven from each interviewee).<sup>2</sup>

First, we counted up the total number of all acknowledgement tokens and calculated the frequency with which these occurred, from which we established that in any given stretch of interviewee talk, an acknowledgement token was given (on average) every  $x$  number of seconds. This provided a measurement of how often the interviewee received *some signal* of the researcher's continued attention. Acknowledgement tokens sometimes appeared in combinations, for example "yeah (.) mm hm". A decision was taken to break up such combinations into two (or more) components. The basic components considered were: *yeah*, *mm hm*, *mm::*, *okay* and *right*.



**Figure 1.** Frequency of acknowledgement tokens (all types).

Figure 1 shows the results of the analysis comprising all acknowledgement tokens. For each stretch of interviewee talk considered, the horizontal axis shows the duration (in seconds) and the vertical axis shows the number of seconds (on average) between each researcher acknowledgement token during that stretch of talk – in other words, an acknowledgement token was vocalized *every x seconds*.

Figure 1 shows that acknowledgement tokens were vocalized by the researcher with lesser frequency in telephone interviews than in face-to-face interviews. For stretches of interviewee talk lasting below 200 seconds (approximately 3½ minutes), acknowledgement tokens were typically vocalized every 3–5 seconds in face-to-face interviews but every 6–9 seconds in telephone interviews.

Secondly, we counted up just the number of ‘continuers’. Continuers might be considered the most neutral type of acknowledgement tokens, such as *mm hm*, *uh huh* or *mm::*. *Mm hm* and *uh huh* have been shown to be largely functionally equivalent (Drummond and Hopper, 1993), as they indeed seem to be in the present data. Regarding *mm::* we listened carefully for inflection and included in our calculations those that appeared to be serving the same function as *mm hm* but not those where the inflection demonstrated more emphasis, surprise, empathy, etc.

The calculation of a continuer every *x* number of seconds indicated how frequently, in a stretch of interviewee talk, the researcher gave what might be considered the most neutral form of receipt, simply indicating comprehension, continued attention and inviting the interviewee to continue speaking (see Gardner, 1997: 133). Our analysis of continuers alone provided a similar pattern to that of the full range of acknowledgement tokens.<sup>3</sup> Continuers were typically vocalized every 6–9 seconds in face-to-face interviews, while in telephone interviews, they were typically vocalized around every 9–14 seconds.

The overall finding, then, is that the researcher did less in the way of vocalized acknowledgement during telephone interviews. Acknowledgement tokens as a whole, and also continuers specifically, were typically vocalized by the researcher with lesser frequency in telephone interviews than in face-to-face interviews.

## Adequacy checks

In classifying the various types of turn taken by interviewees, it emerged that interviewees sometimes checked with the researcher that what they were saying was adequate. This could be in two senses, including whether they had said *enough* (i.e. sufficiency) and whether what they were saying was *along the right lines* (i.e. relevance). Some examples of adequacy checks are shown in Box 3, below:

### Box 3 Adequacy checks

#### Extract 3.1 Telephone interview

- Int: I was very very close to my dad .hhh[hh and I-  
 Res: [Mm:  
 Int: I lost him.  
 (.)  
 Res: Mm [hm  
 Int: [Like we'd moved up to the house me and the  
 two children  
 Res: Yeah.  
 Int: And then like three years ago just like I lost my  
 ↑da:d so that was like ano↑ther .hhhh  
 (0.8)  
 Int: I just- I take rejection very very badly.  
 (0.6)  
 Res: Y:eah.  
 Int: → Am I not answering your ques[tions properly  
 Res: [No no absolutely  
 it's fine

#### Extract 3.2 Telephone interview

- Int: [.hhh ] a::nd so my particular background I:  
 I feel has always been a pr- a pressure  
 Res: Mm hm.  
 Int: from one sort or another .hh[hh  
 Res: [Mm hm.  
 (4.4)  
 Int: → I'm sorry does that answer you:r [your question  
 Res: [Yeah  
 Res: yeah yeah .hh erm and you m- you mentioned at-  
 at college then was th- (.) th- there was another  
 sort of more severe episode.

Although instances in this data set were rare (11 in total), all but one of the six telephone interviewees did some kind of explicit adequacy check at least once during their interview, while this was only the case for one of the five face-to-face interviewees.

### *Interview duration*

Despite much variation in individual interview length, telephone interviews were typically, and on average, shorter than those conducted face-to-face. To the nearest minute, the average (mean) duration of all 11 interviews was 90 minutes. The average duration across the five face-to-face interviews was 101 minutes, while the average duration across the six telephone interviews was 21 minutes shorter at 80 minutes. These findings are thus in contrast with those of Sturges and Hanrahan (2004) and Sweet (2002), who report little difference in interview duration across modes.

Further analysis (see Irvine, 2011, for details) revealed that the shorter duration of telephone interviews was a result of the *interviewee* speaking for less time, rather than a proportional reduction in talk from both parties. In face-to-face interviews, interviewees tended to 'hold the conversational floor' for around 87–90 per cent of the time while in telephone interviews they typically held the floor for somewhere between 76–84 per cent of the time, with relatively greater occupation of 'airtime' by the researcher during telephone interviews. Additionally, in telephone interviews, interviewees generally 'held the floor' for shorter stretches at a time between questions, promptings or other substantive input from the researcher.

## **Discussion**

To summarize the findings presented above, in this comparison of telephone and face-to-face interviews:

- completion or formulation of interviewee talk by the researcher was more common in face-to-face interviews;
- interviewee requests for clarification were slightly more common in telephone interviews;
- vocalized acknowledgements given by the researcher were less frequent in telephone interviews;
- interviewee checks on the adequacy of their responses to the researcher's questions were more common in telephone interviews; and
- telephone interviews tended to be shorter than face-to-face interviews.

The findings of this small-scale, exploratory study must be treated as tentative and non-generalizable. However, they provide a number of emerging themes that could be investigated further in a larger-scale study. In this section, we discuss our findings and consider some potential explanations and implications.

First, our finding that there were more instances of researcher completion and formulation in face-to-face interviews may partly lie in the extent to which rapport is achieved in each mode. Collaborative completion is an affiliative action (Lerner, 1996, 2004) and

it may be that in the face-to-face interviews, the more social nature of the encounters engendered a relationship whereby the researcher sensed more scope to collaborate in the co-production of data. To some extent, the matter of audibility – having to keep quiet in order to hear what the interviewee is saying – may also have contributed to a lesser tendency for the researcher to offer completions or formulations during telephone interviews.

Turning to clarification and comprehension, in this data set, the telephone mode did not appear to lead to increased difficulties in substantive understanding. Thus, on the present evidence, Hermanowicz's (2002) claim that 'breakdowns and misunderstandings in communication easily arise' when using the telephone seems possibly overstated. However, to a small extent, telephone interviewees made more frequent requests for clarification or reiteration of questions. It has been suggested that a poor quality telephone connection can lead to increased fatigue and distraction caused by 'mental detours' while puzzling over unclear words (Stofega, 2006; Vandermate, 2004). Vandermate (2004) states that 'phone conversations are more difficult than face-to-face conversations because the brain must filter and replicate less than stellar aural information, recovering data that drops from a normal phone's audio range, and reassembling words and phrases to fit the meaning of what the caller is saying'). While it should be noted that these claims come from telecommunications providers (who may have a vested interest in promoting more sophisticated equipment), this notion of more 'effortful' listening offers a plausible (partial) explanation both for our findings on clarification requests and the finding that telephone interviews tended to be shorter, on average, than face-to-face interviews.

An alternative explanation for differences in the frequency of clarification requests may lie in the way that the researcher formulated her interview questions. It was not within the scope of the present study to explore question formulation in detail. However, this is an important and under-investigated area, the advancement of which would be of value to the qualitative research community. Possible questions for further study might include: How are questions *actually* produced in real-life qualitative interview interactions (as opposed to the written topic guide or the advice of instructional texts)? Why might the interviewer produce a question in a particular way? How does question construction affect the response provided? What type and range of utterances do the work of prompting or probing?<sup>24</sup>

Our finding that acknowledgment tokens were uttered less frequently by the researcher in telephone interviews seems counterintuitive, given that the absence of visual cues would suggest a tendency towards *greater* use of verbal responses. The lesser frequency of acknowledgment tokens in the present data suggests that they were not 'replacing' such things as nods or smiles (see Miller, 1995). One possible explanation for the lesser frequency of acknowledgement tokens on the telephone is that the researcher was having to listen harder and so withheld her own utterances to avoid interfering with audibility.

Another possible explanation is that the researcher *was indeed* paying less attention during telephone interviews. A perceived practical advantage of telephone interviews, noted by Chapple (1999), Stephens (2007) and Sweet (2002) is the ability to take notes unobserved by the interviewee. Sweet (2002: 60) notes that this avoids 'causing visual distraction and disrupting the interview process'. The researcher who conducted the



interviews in the present study also considered this ability to take notes discreetly as an advantage. However, it is possible that she was less successful than she thought in simultaneously maintaining engagement with the interviewee while making notes and memos. The greater scope for (unobserved) distraction or diversion while ostensibly listening to the interviewee may also have inadvertently affected the researcher's responsiveness.

This relative lack of 'feedback' may in part explain the tendency for telephone interviewees to make more explicit checks that what they were saying was 'adequate' (sufficient or relevant) to the researcher's needs. Taken alongside the inherent absence of visual cues, we might infer that telephone interviewees received less unprompted confirmation of the adequacy of their talk from the researcher, hence their resorting to explicit checks. An alternative or additional explanation for the increased use of adequacy checks by telephone interviewees may lie in the broader social context of the encounter. In comparison to the face-to-face research encounters taking place in people's homes or workplaces, telephone interviews in this data set involved much less social interaction pre- and post-interview. As such, we suggest that telephone interviews may have had a relatively more 'businesslike' atmosphere. Thus, it is possible that telephone interviewees were more conscious of the task at hand and their role in the interaction as information providers.

Holt (2010) makes some insightful comments in this respect. For her telephone interviews, Holt set out to use a narrative (rather than semi-structured) approach. In subsequently consulting her interviewees on their experiences of being interviewed in this way, Holt found that interviewees would have appreciated more direct prompting and steer from the researcher. She suggests that interviewees' expectations were driven by their prior conceptions of a research interview, whereby greater researcher directivity was expected. Holt recognizes that the format of the narrative interview may have been particularly unfamiliar to interviewees and so their preference for more structure in this instance 'may not be specific to the use of the telephone'. However, she also notes her interviewees' 'preoccupation about how they *should* perform' in the interviews, particularly with regard to whether they were talking too much (Holt, 2010: 118). Here, there are parallels both with the present findings on adequacy checks and the findings that interviewees tended to speak less overall and for shorter stretches at a time, when interviewed over the phone.

Our findings on adequacy checks suggest that, to some extent, telephone interviewees were less confident that the information they were providing was meeting the researcher's needs. On the one hand, the fact that interviewees resorted to explicit checks indicates that they found interactional strategies through which to resolve their uncertainty. However, this possible sense of inadequacy in fulfilling the 'role of interviewee' may have had an impact on their lasting impression of their 'performance' in the interview and hence on their enduring feelings about the experience of participation in the interview (and the research project overall). For qualitative researchers, this may constitute an important ethical consideration, particularly when working with vulnerable populations or sensitive research topics.

We may also ask further questions about why telephone interviews tended to be shorter and what the implications of this may be for our research. As we discuss in more

depth elsewhere (see Irvine, 2011) we found that, to a moderate degree, the shorter length of telephone interviews could be accounted for by a reduction in coverage of themes as the study progressed. However, the principal explanation appeared to lie in a tendency for telephone interview participants to provide relatively less detail or elaboration. As noted earlier, previous authors have suggested that reduced rapport or increased demand/effort may lead to telephone interviews being shorter or less rich in detail. Studies seeking relatively simple or descriptive data may not require a level of rapport that encourages extended reflective accounts; greater *quantity* of data may not necessarily imply greater *quality* of data for a given analytic purpose. On the other hand, for some studies, such in-depth material could be crucial in thoroughly addressing research questions.

## Conclusion

Shuy (2003: 191) asserts that ‘for researchers to know exactly what they may be missing or gaining when they carry out telephone interviews rather than in-person interviews, they need reliable information based on detailed comparisons between the two modes’. The present study has sought to contribute to meeting this need and has shown that, when data is considered in close detail, certain differences between interview modes *do* emerge. As has been noted, our small-scale, exploratory investigation has only been able to provide some initial suggestions as to what difference it might make if we conduct qualitative research interviews by telephone rather than face-to-face, all of which are tentative and none of which can be generalized. However, we hope that our study has raised the profile of this important methodological question and we would invite and encourage other researchers to join us in furthering the inquiry.

In expanding the knowledge base in this area, it would be valuable for any subsequent studies to include data from interviews conducted by a range of researchers and on a number of different topics involving interviewees with different backgrounds and characteristics. And as has been recommended by Holt (2010), a valuable component to any future research on qualitative mode effects would be the inclusion of interviewees’ views on their experience.

## Funding

The study was funded by the Economic and Social Research Council (ESRC) grant no. RES-000-22-3668.

## Notes

1. For an introduction to the method of Conversation Analysis see Drew (2005) and Sidnell (2007).
2. Because there is one more telephone interviewee, slightly more instances come from telephone interviews (42) than face-to-face (35).
3. For reasons of space, the graph showing continuers alone has not been presented here.
4. We are grateful to colleagues who took part in a practitioner workshop as part of the present study, in particular Professor Jonathan Potter and Dr Tim Rapley, for highlighting the value of further research into question construction and the ‘practical enactment’ of an interview guide.

## References

- Baker C (2004) Membership categorization and interview accounts. In: Silverman D (ed.) *Qualitative Research: Theory, Method and Practice*. London: Sage, 162–176.
- Baker CD (2003) Ethnomethodological analyses of interviews. In: Holstein JA and Gubrium JF (eds) *Inside Interviewing: New Lenses, New Concerns*. Thousand Oaks, CA: Sage, 395–412.
- Berg BL (2007) *Qualitative Research Methods for the Social Sciences*. Boston, MA: Pearson Education Inc.
- Burnard P (1994) The telephone interview as a data collection method. *Nurse Education Today* 14: 67–72.
- Carr ECJ and Worth A (2001) The use of the telephone interview for research. *Nursing Times Research* 6(1): 511–524.
- Chapple A (1999) The use of telephone interviewing for qualitative research. *Nurse Researcher* 6(3): 85–93.
- Clayman SE, Heritage J, Elliott MN, et al. (2007) When does the watchdog bark? Conditions of aggressive questioning in presidential news conferences. *American Sociological Review* 72: 23–41.
- Cresswell JW (2007) *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. Thousand Oaks, CA: Sage.
- Dicker R and Gilbert J (1988) The role of the telephone in educational research. *British Educational Research Journal* 14: 65–72.
- Drew P (2002) Comparative analysis of talk-in-interaction in different institutional settings: a sketch. In: Glenn P, LeBaron C and Mandelbaum J (eds) *Excavating the Taken-for-Granted: Essays in Language and Social Interaction*. Mahwah, NJ: Lawrence Erlbaum, 293–308.
- Drew P (2005) Conversation Analysis. In: Fitch KL and Sanders RE (eds) *Handbook of Language and Social Interaction*. Mahwah, NJ: Lawrence Erlbaum, 71–102.
- Drummond K and Hopper R (1993) Back channels revisited: acknowledgement tokens and speakership incipency. *Research on Language and Social Interaction* 26(2): 157–177.
- Fielding N and Thomas H (2008) Qualitative interviewing. In: Gilbert N (ed.) *Researching Social Life* (3rd Edition). London: Sage, 245–265.
- Gardner R (1997) The conversational object mm: a weak and variable acknowledging token. *Research on Language and Social Interaction* 30(2): 131–156.
- Gillham B (2005) *Research Interviewing: The Range of Techniques*. Berkshire: Open University Press.
- Grant A (2011) Fear, confusion and participation: incapacity benefit claimants and (compulsory) work focused interviews. *Research, Policy and Planning* 28(3): 161–171.
- Heritage J (1984) A change-of-state token and aspects of its sequential placement. In: Atkinson JM and Heritage J (eds) *Structures of Social Action*. Cambridge: Cambridge University Press, 299–345.
- Heritage J (1985) Analyzing news interviews: aspects of the production of talk for an overhearing audience. In: Van Dijk T (ed.) *Handbook of Discourse Analysis* (Volume 3). London: Academic Press, 95–117.
- Heritage JC and Roth AL (1995) Grammar and institution: questions and questioning in the broadcast news interview. *Research on Language and Social Interaction* 28(1): 1–60.
- Heritage J, Robinson JD, Elliott MN, Beckett M and Wilkes M (2007) Reducing Patients' unmet concerns in primary care: the difference one word can make. *Journal of General Internal Medicine* 22(10): 1429–1433.
- Hermanowicz JC (2002) The great interview: 25 strategies for studying people in bed. *Qualitative Sociology* 25(4): 479–499.

- Holt A (2010) Using the telephone for narrative interviewing: a research note. *Qualitative Research* 10(1): 113–121.
- Irvine A (2008) Managing mental health and employment. Department for Work and Pensions Research Report No. 537. Leeds: Corporate Document Services.
- Irvine A (2011) Duration, dominance and depth in telephone and face-to-face interviews: a comparative exploration. *International Journal of Qualitative Methods* 10(3): 202–220.
- Jefferson G (1993) Caveat speaker: preliminary notes on recipient topic-shift implicature. *Research on Language and Social Interaction* 26(1): 1–30.
- Jefferson G (2004) Glossary of transcript symbols with an introduction. In: Lerner GH (ed.) *Conversation Analysis: Studies from the First Generation*. Amsterdam: John Benjamins, 13–31.
- Kavanaugh K and Ayres L (1998) 'Not as bad as it could have been': assessing and mitigating harm during research interviews on sensitive topics. *Research in Nursing and Health* 21: 91–97.
- Legard R, Keegan J and Ward K (2003) In-depth interviews. In: Ritchie J and Lewis J (eds) *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: Sage, 138–169.
- Lerner G (1996) On the 'semi-permeable' character of grammatical units in conversation: conditional entry into the turn-space of another speaker. In: Ochs E, Schegloff EA and Thompson SA (eds) *Interaction and Grammar*. Cambridge: Cambridge University Press, 238–276.
- Lerner G (2004) Collaborative turn sequences. In: Lerner G (ed.) *Conversation Analysis: Studies from the First Generation*. Philadelphia, PA: John Benjamins, 225–256.
- Mazeland H and ten Have P (1996) Essential tensions in (semi-)open research interviews. In: Maso I and Wester F (eds) *The Deliberate Dialogue: Qualitative Perspectives on the Interview*. Brussels: VUB University Press. Online version published in 1998. Available at: <http://www.pscw.uva.nl/emca/ET.htm>
- Miller C (1995) In-depth interviewing by telephone: some practical considerations. *Evaluation and Research in Education* 9(1): 29–38.
- Novick G (2008) Is there a bias against telephone interviews in qualitative research? *Research in Nursing and Health* 31: 391–398.
- Opdenakker R (2006) Advantages and disadvantages of four interview techniques in qualitative research. *Forum: Qualitative Social Research* 7(4). Available at: <http://www.qualitative-research.net/index.php/fqs/article/view/175> Accessed 2 October 2009.
- Rapley M and Antaki C (1998) What do you think about ...? Generating views in an interview. *Text* 18(4): 587–608.
- Rapley TJ (2001) The art(fulness) of open-ended interviewing: some considerations on analysing interviews. *Qualitative Research* 1(3): 303–323.
- Robson C (2002) *Real World Research*. Oxford: Blackwell.
- Roulston K (2001) Data analysis and 'theorising as ideology'. *Qualitative Research* 1(3): 279–302.
- Roulston K, Baker CD and Liljestrom A (2001) Analyzing the researcher's work in generating data: the case of complaints. *Qualitative Inquiry* 7(6): 745–772.
- Rubin HJ and Rubin IS (2005) *Qualitative Interviewing: The Art of Hearing Data* (2nd Edition). Thousand Oaks, CA: Sage.
- Shuy RW (2003) In-person versus telephone interviewing. In: Holstein JA and Gubrium JF (eds) *Inside Interviewing: New Lenses, New Concerns*. Thousand Oaks, CA: Sage, 175–193.
- Sidnell J (2010) *Conversation Analysis: An Introduction*. Oxford: Wiley-Blackwell.
- Stephens N (2007) Collecting data from elites and ultra elites: telephone and face-to-face interviews with macroeconomists. *Qualitative Research* 7(2): 203–216.
- Stofega W (2006) Can you hear what I mean? Polycom delivers HD Voice. Available at: <http://white-paper.techworld.com/telephony-voip/2888/idc-report-can-you-hear-what-i-mean-polycom-delivers-hd-voice/?otc=75&otd=TW> Accessed 12 August 2010.

- Sturges JE and Hanrahan KJ (2004) Comparing telephone and face-to-face qualitative interviewing: a research note. *Qualitative Research* 4(1): 107–118.
- Sweet L (2002) Telephone interviewing: is it compatible with interpretive phenomenological research? *Contemporary Nurse* 12(1): 58–63.
- Tausig JE and Freeman EW (1988) The next best thing to being there: conducting the clinical research interview by telephone. *American Journal of Orthopsychiatry* 58(3): 418–427.
- ten Have P (2004) *Understanding Qualitative Research and Ethnomethodology*. London: Sage.
- Vandermate J (2004) Telephone voice quality breaks the bandwidth barrier. *Enterprise Networks and Servers*, July. Available at: <http://www.enterprisenetworksandservers.com/monthly/art.php?792> Accessed 12 August 2010.
- Wooffitt R and Widdicombe S (2006) Interaction in interviews. In: Drew P, Raymond G and Weinberg D (eds) *Talk and Interaction in Social Research Methods*. London: Sage, 28–49.

## Appendix

### Transcription conventions

<i>Int:/Res:</i>	Speaker labels ( <i>Int:</i> = Interviewee; <i>Res:</i> = Researcher)
° °	Encloses talk which is produced quietly
* *	Encloses talk which is produced in ‘creaky’ voice
<u>underline</u>	Underlining used to mark words or syllables which are given special emphasis of some kind
CAPS	Words or parts of words spoken loudly marked in capital letters
s:::	Sustained or stretched sound; the more colons, the longer the sound
.hhh	In-breath (more .hhhh to indicate longer in-breath)
[ ]	Encloses talk produced in overlap i.e. when more than one speaker is speaking
(word)	Parentheses indicate transcriber doubt
cu-	Cut-off word or sound
(0.6)	Silence in seconds
(.)	Silence of less than two tenths of a second
↑	Marks high pitch
> <	Marks speeding up delivery (in talk between the facing arrows)

### Biographical notes

Annie Irvine is Research Fellow in the Social Policy Research Unit (SPRU) at the University of York. Her research is focused on health, disability, employment and social security benefits. She has particular interest in mental health and employment and in qualitative interview methods.

Paul Drew is Professor of Sociology and the Director of the Centre for Advanced Studies in Language and Communication at the University of York. As well as basic research into ordinary (social) interaction and communication, he has extensive experience of research into medical, legal and other workplace interactions (including contract research for the Metropolitan Police Service).

Roy Sainsbury is Professor of Social Policy and a Research Director in the Social Policy Research Unit (SPRU) at the University of York. His research interests and experience include employment, social security, disability, health and well-being. Recent work has included an evaluation of Pathways to Work and an investigation of the relationship between mental health and employment.