The Qualitative Report

Volume 24 | Number 3 How To Article 12

3-18-2019

Using Scribes in Qualitative Research as an Alternative to Transcription

Kim Eaton
University of Western Australia, kim.eaton@research.uwa.edu.au

Werner G. K. Stritzke
University of Western Australia, werner.stritzke@uwa.edu.au

Jeneva L. Ohan
University of Western Australia, jeneva.ohan@uwa.edu.au

Using Scribes in Qualitative Research as an Alternative to Transcription

Abstract

Transcribing qualitative data is resource-intensive. One less intensive alternative is scribing: the documenting of comprehensive notes, including verbatim quotes by an independent observer during an interview. However, the extent to which a comparable thematic analysis can be derived from scribed interview data relative to verbatim transcriptions of these same interviews has not been investigated. Thus, the purpose of this study is to test the number and content of themes derived from interview data, which had been scribed versus transcribed verbatim and to identify the time and cost differences (if any) between obtaining, processing, and analysing scribed data compared to transcribed data. Two modes of scribing were evaluated: in-person (i.e., from notes obtained during live interviews), and from video-recordings of these same interviews. There was high consistency in the number and content of themes (highest at subtheme level) derived from scribed versus transcribed data. Scribing produced significantly less data than transcribing and was economically superior. Thus, in the context of interview-based studies in which common ideas or meaning are sought through thematic analysis, scribing yields a similarly rich set of themes as transcribing, and hence, may offer a valid and feasible alternative when resources are limited.

Keywords

Scribe, Transcription, Interviews, Qualitative

The Qualitative Report 2019 Volume 24, Number 3, How To Article 3, 586-605

Using Scribes in Qualitative Research as an Alternative to Transcription

Kim Eaton, Werner G. K. Stritzke, and Jeneva L. Ohan
University of Western Australia, Perth, Australia

Transcribing qualitative data is resource-intensive. One less intensive alternative is scribing: the documenting of comprehensive notes, including verbatim quotes by an independent observer during an interview. However, the extent to which a comparable thematic analysis can be derived from scribed interview data relative to verbatim transcriptions of these same interviews has not been investigated. Thus, the purpose of this study is to test the number and content of themes derived from interview data, which had been scribed versus transcribed verbatim and to identify the time and cost differences (if any) between obtaining, processing, and analysing scribed data compared to transcribed data. Two modes of scribing were evaluated: in-person (i.e., from notes obtained during live interviews), and from video-recordings of these same interviews. There was high consistency in the number and content of themes (highest at subtheme level) derived from scribed versus transcribed data. Scribing produced significantly less data than transcribing and was economically superior. Thus, in the context of interview-based studies in which common ideas or meaning are sought through thematic analysis, scribing yields a similarly rich set of themes as transcribing, and hence, may offer a valid and feasible alternative when resources are limited. Keywords: Scribe, Transcription, Interviews, Qualitative

Qualitative researchers use a range of methods that facilitate the in-depth exploration of the complexities of human perspectives, constructs, and concepts (Lincoln & Guba, 2003; Yilmaz, 2013). Yet, qualitative research is often prohibitive as it can be laborious, time consuming, and expensive (Neal, Neal, van Dyke, & Kornbluh, 2015; Tilley, 2003). Transcription, the processing of raw interview data into a text-based form, is a major contributor to the resource-intensive nature of qualitative research (Halcomb & Davidson, 2006). Transcription, also referred to as transcribing, can take up to 10 hours per hour of interview recording for a verbatim transcription (Bailey, 2008), and is somewhat more economical at up to 7 hours per hour of recording for non-verbatim transcription (Britten, 1995). Computerised transcription methods (e.g., voice recognition software) only partially remediate the issue, given that errors in punctuation can arise, which impact on transcript comprehensibility (Jarnow, 2017; Johnson, 2011; Perrier & Kirkby, 2013). Further, copious data are produced, which then take time to analyse, increasing with the amount of text (Johnson, 2011; MacLean, Meyer, & Estable, 2004; Tessier, 2012; Tilley, 2003). Given these high resource demands, alternative methods that increase the cost-viability of qualitative research have been sought. One such recently employed alternative is the use of a scribe, a third person within the interview to document extensive notes, with these notes subsequently analysed (e.g., Bex Lempert, 2016; Corrigan, Pickett, Kraus, Burks, & Schmidt, 2015; Mowat, 2012).

Despite the recent use of scribes in qualitative research, the extent to which a similar thematic analysis is derivable from scribed data relative to verbatim transcription has not been established. Further, although one key criticism of transcription is that it is resources-intensive, there has been no test of how economical scribing is, relative to transcription. Thus, the aim of

this current study is to address these issues by examining the extent to which thematic analysis of interview data scribed in-person (and from video-recording) yields comparative results to data transcribed verbatim. The time and labour costs involved in scribing versus transcribing is also compared.

To address the problems of time, labour, and cost, some have argued in favour of foregoing the transcription process altogether (Bentley et al., 1988; Kieren & Munro, 1985; McNall & Foster-Fishman, 2007). Early alternatives focussed on coding from the raw data itself (i.e., the electronic recording). However, despite evidence that coding directly from the recording kept researchers close to their data—an essential element of analysis—and was faster than coding from transcription (Crichton & Childs, 2005; Gravois, Rosenfield, & Greenberg, 1992), reliability was often compromised (Lapadat & Lindsay, 1999). For example, compared to coding from transcription, coding from the recording resulted in a 13-34% loss of data and marked inconsistencies in code assignment (Kieren & Munro, 1985). Further, without a written record of the data, an audit trail can be difficult to establish (Halcomb & Davidson, 2006). Audit trails are essential for demonstrating credibility and trustworthiness of the evolution of the overall thematic construct (Koch, 2006). For these reasons, qualitative methodologists recommend processing raw data into a text-based form prior to analysis (Tessier, 2012).

Bearing in mind the necessity of text-based datasets, one proposal has been to use field notes scribed during the interview (Kieren & Munro, 1985). Field notes consist of the researcher's documented ideas regarding the interview, and comments on context and the conversation exchange itself (Phillippi & Lauderdale, 2017). The benefit of field notes is that they are contemporaneous and can be elaborated on with subsequent viewing of the electronic recording (Bentley et al., 1988; Tessier, 2012). Although some use field notes to supplement recordings and/or their transcriptions to aid in the interpretation and generation of meaning (Halcomb & Davidson, 2006), others rely solely on extensive field notes as the data corpus, replacing verbatim transcription entirely (e.g., McNall & Foster-Fishman, 2007).

The collection and analysis of field notes is a feature of rapid evaluation and assessment method (REAM) studies (McNall & Foster-Fishman, 2007). Generally, REAM projects aim to be fast and selective in data acquisition and analysis (Beebe, 2014; Manderson & Aaby, 1992). To do so, case summary templates with subject headings are created a priori (Beebe, 2014). During interviews with and/or observations of participants, field notes are recorded directly onto these templates under a corresponding heading (Beebe, 2014). These notes are subsequently coded and then compiled into a data matrix (Averill, 2002) for cross-case evaluation (Beebe, 2014). In the McNall and Foster-Fishman (2007) variant of REAM, field notes were not assigned to pre-prescribed headings; rather, codes and then themes were assigned a posteriori to field note collection. The authors argued that this method was reliable and valid as it met the Guba and Lincoln (1989) adequacy criteria of trustworthiness and authenticity, such as credibility, transferability, confirmability, and fairness. How robust the thematic analysis of field notes is, compared to verbatim transcription, has yet to be established.

Interviewer-produced field notes, however, are fundamentally disadvantaged given that the interviewer needs to simultaneously engage in the process of the interview and in making field notes. Interviewer note-taking potentially disrupts the interview, compromising both the notes and the interview exchange (Barker, Pistrang, & Elliott, 2016; Sturges & Hanrahan, 2004). Moreover, it may result in "thin" datasets, replete with missing data and an underrepresentation of participant voices that could render findings incomplete, simplistic, or inconclusive (Hamo, Blum-Kulka, & Hacohen, 2004; Kieren & Munro, 1985; Tessier, 2012). Thus, although interviewer-produced field notes address the need for a written (yet reduced) record of the interview, they may not provide the detail necessary to complete a trustworthy and credible analysis.

To limit the intrusive effects of note-taking during interviews, an alternative is to use a third person (a scribe) to perform note-taking duties (Seaman, 1999). Unlike the interviewer, scribes are independent of the interview, and can document notes, verbatim quotes, and contextual detail. Recently, Bex Lempert (2016) used scribes during interviews and focus groups within a prison setting to record written data, including quotes. The scribes were essential, as electronic recording was prohibited, precluding verbatim transcription of the recording. Corrigan et al. (2015) also used scribes to document extensive notes during interviews and focus groups in a community setting. The scribed notes were subsequently analysed using a grounded theory approach, obtaining thematic saturation. Thus, by using a scribe independent of the interview process, Corrigan et al. (2015) obtained a comprehensive text-based dataset sufficient to complete thematic analysis without the need for verbatim transcription.

In sum, over the past two decades, researchers have made increasing attempts to reduce data collection and analysis burden and increase the cost and time feasibility of qualitative research completed via coding and/or thematic analysis. To do so, alternatives to transcribing qualitative data have been proposed and are frequently used. One method, the use of a scribe to document extensive notes is already being used by researchers to expediently obtain a manageable dataset and to document data in research situations in which electronic recording is prohibited and/or transcription unavailable. However, there has been little exploration as to whether this method achieves similar analytic outcomes in terms of the number and content of themes derived through thematic analysis. Importantly, scribed data have not been directly compared to verbatim transcriptions, and hence the degree to which scribing offers savings in resources, given the need to hire and train a second individual to be present for interviews, is unknown. Thus, in this study we examine what, if anything is the impact of using scribed data—relative to verbatim transcription—to derive themes and subthemes. First, we compare the number and content of themes and subthemes derived from data scribed during interviews (and from video-recordings) to those derived from the verbatim transcription. Then, we conduct an economic analysis (i.e., time and cost) of scribing in comparison to transcription.

The Researchers

Kim Eaton is a clinical psychologist registrar working with children and adolescents and their families. She has recently completed her PhD studies, of which, two chapters constituted qualitative studies. It was during these studies, completed as a student, and in and previous qualitative work undertaken in community- and tertiary- treatment settings that she became increasingly aware of the demands of verbatim transcription and the need for a rigorous alternative to expedite the process. This was particularly the case given that student and community projects are often minimally funded and time-limited. However, in finding limited research or robust evidence supporting alternatives, the impetus for the current study emerged. Werner Stritzke is a clinical psychologist who worked initially for many years as a counsellor with juvenile delinquents in Germany and with abused and neglected children in the United States. Following his PhD, his research in the area of substance use and misuse, and more recently in suicide risk and resilience, has employed a wide range of methodological approaches to suit the particular research questions or populations studied. These include innovative explicit and implicit assessment methods, experimental and longitudinal designs, and qualitative approaches. Jeneva Ohan is a senior lecturer and registered psychologist at the University of Western Australia. She teaches child and adolescent assessment and interventions to students in clinical psychology. Like her clinical experience, her research is in the area of parent and child mental health, mental health service use, and the stigma that these families

encounter. She has used a range of research methods and study designs, from experimental to qualitative and naturalistic, in her research studies.

Method

Ethics approvals for the initial qualitative study and the subsequent methodological comparison study were granted by our university's human research ethics office. Participants provided consent for the initial study and for subsequent use of the data.

Background

To test the scribe method described herein, we scribed and transcribed interviews (n = 12), obtaining a data corpus from each of the methods. These interviews were semi-structured and ranged from 29.49 to 63.06 minutes (M = 42.43, SD = 11.10). We have previously reported on the results of the substantive thematic analysis of the scribed data (Eaton, Ohan, Stritzke, & Corrigan, 2016). In the current methodological investigation, we report on our method of training and using scribes featured in that study and compare the thematic outcomes of scribed data versus verbatim transcription.

Procedures for Training Scribes

Prior to scribing interviews, two students with clinical and interview experience enrolled in a postgraduate psychology program were trained to record and analyse scribe notes (one 3.5 hour session). The steps involved in training scribes are as follows.

Step 1: Prepare training materials. Training materials consisted of a pool of 10 audio-recorded segments extracted from de-identified interviews from a separate study for which a thematic analysis had already been completed (Eaton, 2013). Segments ranged from 16.3-22.4 minutes (M = 19.3, SD = 3.05). A master-set of scribed notes, codes, subthemes and themes were created for each segment using the procedure for using scribes described in the next section.

Step 2: Instruct scribes. The main task of a scribe (i.e., to document comprehensive notes for each interview) was clearly set out. Scribes were instructed to document: (1) topics raised by participants, (2) quotes (i.e., verbatim exemplars of participants' spoken words), (3) the interview question that led to participants' responses, (4) emotional detail (e.g., crying, laughter), (5) non-verbal detail (e.g., fidgeting), (6) contextual detail (e.g., aspects of the location of the interview), and (7) any other salient detail or detail likely to influence interpretation of meaning (e.g., participants' use of metaphor, sarcasm, emphasis, or voice inflection).

Step 3: Scribing of training segments. Listening to an interview segment, the scribes each generated their own scribed notations.

Step 4: Discrepancy check and rectification (of scribed notes). The scribed notes were compared, line-by-line, against the master-set. Discrepancies, such as omitting salient quotes or including filler words (e.g., "um," "ah") were resolved by negotiated agreement (Campbell, Quincy, Osserman, & Pedersen, 2013; Garrison, Cleveland-Innes, Koole, & Kappelman, 2006). This involved scribes discussing the information they had documented and their justifications for doing so in an effort to reconcile the discrepancies. We replicated this

process using subsequent interview segments until the discrepancy check revealed only minor disagreements between scribed notes and the master-set.

Step 5: Analytical process. Scribes were then trained to code and thematically analyse the scribed notes. Based on the social constructivist methods of Giorgi (2009), scribes began with pre-reading the scribed data to gain familiarity and a general sense of the messages conveyed within. Data were then segmented into meaning units (i.e., sentences and phrases) and descriptively coded (i.e., first pass coding). Codes were collapsed, amended, and deleted where necessary (i.e., second pass coding). Subthemes and themes were identified through a recursive and iterative process of code refinement, moving from individual codes to subthemes, to themes.

Step 6: Discrepancy check and rectification (of codes and themes). Codes and themes were compared against the master-set. Again, by negotiated agreement, each discrepancy was discussed and resolved. Consistency was deemed reached once all discrepancies were resolved.

Procedure for Using Scribes In-Person

On completion of the training, qualitative interviews commenced. These interviews were scribed in-person (i.e., within the interview) using the following steps.

Step 1: Informing and introducing. Participants were informed about the inclusion of the scribe in the participant information letter and again at interview scheduling. Participants were free to decline the scribe's presence without repercussion (an option not chosen by any participant). The scribe was introduced to participants at the commencement of the interview.

Step 2: Video record interview. Interviews were video recorded to obtain a first-hand record of the interview.

Step 3: Concurrent note-taking by scribe. The scribe sat outside of the immediate interview space to unobtrusively observe both the participant and interviewer. Scribed notes including quotes, contextual detail, and non-verbal detail were documented throughout the interview (as described in Step 2 in "Procedure for Training Scribes" above).

Step 4: Reflexive journaling (ongoing during analysis). Immediately post-interview, scribes made notes on their reflections regarding the interview process and content, and their opinions, thoughts, and feelings relative to the data they were collecting. Scribed notes were expanded on and ideas regarding emergent themes noted. Such reflexivity within qualitative research is essential to establish transparency, identify sources and evidence of potential bias, and enhance reliability of data collection, analysis, and findings (Ortlipp, 2008; Shaw, 2010).

Step 5: Review and revise scribe notes. Scribes reviewed and revised their notes using the video recordings as appropriate so as to accurately and comprehensively capture verbatim quotes and other relevant detail.

Step 6: Analytic process. Phases of preliminary and secondary coding were conducted, culminating with the derivation of subthemes and themes. The method for which has been described above in Step 5 of the "Procedure for Training Scribes." In addition, an audit trail was maintained and included reflexive notes and a record of theme evolution.

Transcribing

Two students, who had a minimum of a psychology honours degree were trained to transcribe verbatim from the interview video recordings using the training materials previously described (one 3.5 hour session). Verbatim transcripts were drafted using the Jefferson notation system (Atkinson & Heritage, 1999; Jefferson, 2004; Oliver, Serovich, & Mason, 2005). Transcribers were then trained (one 3.5 hour session) to code and thematically analyse transcriptions using the same method as the scribes. Consistent with scribes, transcribers maintained reflexive journals and a record of theme evolution.

The time taken to scribe, transcribe, and code each interview was recorded. Six of the 12 interviews were randomly selected to complete the methodological comparison.

Results

Amount of Data Produced by In-Person Scribes Versus Transcription

Because an integral purpose of using scribes is to obtain an easy-to-manage, compact dataset that facilitates later qualitative analysis, we first analysed the amount of information produced in scribed notes versus verbatim transcriptions. To do this, we examined the amount of text-based data (i.e., number of words) per method. Compared to transcription, scribing in-person produced 73% less data (9,254 versus 34,263 words).

Comparisons of Subthemes and Themes Derived from Scribed Versus Transcribed Data

There is a concern that the reduced dataset produced by scribes may result in a thinner thematic analysis and compromise the number of themes produced. Thus, the extent to which the reduction of data impacted (if at all) on the thematic analysis of these data was examined by comparing the number of subthemes and themes derived from in-person scribed notes versus transcriptions. Forty subthemes and nine themes emerged from each of the methods.

We also examined the extent to which the subthemes and themes derived from in-person scribed data qualitatively differed to those derived from transcriptions. To do so, two independent raters, who were blind to the aims of the study, were presented with a series of subtheme pairs consisting of one subtheme (and brief description) from each method. Each scribe-derived subtheme was compared, one at a time, to each transcription-derived subtheme. Comparison pairs were counterbalanced to account for the two different methods (Gravois et al., 1992) and two distractor themes were included. The similarity of each comparison pair was rated on a 0-100 scale (0 = not at all similar; 100 = exactly the same). Value ranges for the similarity comparisons were based on the Koo and Li (2016) guidelines for the interpretation of Cohen's kappa or intraclass correlation inter-rater agreement. In adopting these fairly stringent guidelines, ratings of 90 (or above) indicated *excellent* similarity, 75-89 indicated *good* similarity, 50-74 indicated *moderate* similarity, 49 (or below) indicated *poor* similarity. This process was repeated at theme level.

The distractor subthemes/themes were rated as having 0% similarity with any other subtheme/theme and were removed from the comparison list. For each rater, the highest unique match for each subtheme was identified. This involved rank ordering the comparison scores, isolating the highest match, and removing it from subsequent comparisons to avoid overlapping. This process was continued until the highest match for each subtheme was identified. The process was again repeated at theme level.

Table 1 illustrates the number of subthemes and themes, and their similarity ratings. Between 80% (Rater A) and 82.5% (Rater B) of subthemes overlapped between the two

methods to a good to excellent degree. A similar good to excellent degree of overlap was found by each of the raters for 77.78% of themes.

Table 1. *Number of Subthemes and Themes, and Their Similarity Ratings, And Inter-Rater Reliability (i.e., Rater A and B)*

	>90 Excellent	75-89 Good	50-74 Moderate	<49 Poor	ICC	95%CI	*F*	(df,df)
Transcription to Scribe (*in-person*) subtheme level					.87*	[.86, .89]	7.88	(1599, 1599)
Rater A	31	1	5	3				
Rater B	31	2	1	6				
Transcription to Scribe (*in-person*) theme level					.89*	[.89, .90]	8.82	(80, 80)
Rater A	6	1	0	2				
Rater B	7	0	0	2				
Transcription to Scribe (*video*) subtheme level					.87*	[.86, .89]	7.87	(1599, 1599)
Rater A	30	2	5	3				
Rater B	30	4	2	4				
Transcription to Scribe (*video*) theme level					.95*	[.93, .97]	21.27	(80, 80)
Rater A	3	3	0	3				
Rater B	5	1	0	3				
Scribe (*in-person*) to Scribe (*video*) subtheme level					.90*	[.89, .91]	9.95	(1599, 1599)
Rater A	28	4	3	5				
Rater B	27	7	2	4				
Scribe (*in-person*) to Scribe (*video*) theme level					.94*	[.91, .96]	16.88	(89.89)
Rater A	5	1	1	2				
Rater B	5	1	1	2				

Note. Similarity value ranges based on the Koo and Li (2016) guidelines on inter-rater agreement; ICC = Interclass correlation; *$p < .001$

An inspection of the moderately (or below) overlapping themes revealed that these constituted similar subthemes, despite being structurally organised into themes differently. For example, as shown in Table 2, the scribe-derived theme "*privacy*" moderately overlapped with the transcription-derived theme "*shame*." But, inspection of the respective subthemes shows that the scribe-derived subthemes "*disclosure*" and "*none of your business*" within the theme "*privacy*" were very similar to the transcription-derived subthemes "*non-disclosure*" and "*within the family*" within the theme "*shame*." The two additional subthemes parcelled under the transcription-derived "*shame*" theme (i.e., "*embarrassment*" and "*social comparisons*"), which did not feature in the scribe-derived them "*privacy*" were still captured in the scribed data but were featured under a theme specifically about shame (not shown). Thus, at the subtheme level, raters showed an overall high level of similarity, indicating that the organisation of subthemes into themes was responsible for the slightly lower agreement at theme level.

In Table 3, we provide exemplars of subtheme pairs and illustrative data excerpt at each similarity level. As shown, a high similarity rating indicates a high level of congruence in the

subtheme title and description across the two methods. Additionally, the data excerpts illustrate the level of detail included in in-person scribed notes and in verbatim transcriptions. Scribed data clearly contains less of the paralinguistic nuances exhibited in the transcriptions, yet the scribed data capture with a high degree of overlap similar subthemes as the transcriptions. In Table 2, exemplars of theme pairs and their constituent subthemes at each similarity level are provided. As shown, a high similarity rating indicates a high level of congruence in the theme description and the theme's constituent subthemes across the two methods. In the interests of space, an exhaustive list of subthemes and themes is not provided[1].

Table 2. *Theme Exemplars (and Composite Subthemes) at Each Similarity Level for In-Person Scribe to Transcription Comparisons*

	>90 (Excellent)	75-89 (Good)	50-74 (Moderate)	<49 (Poor)
Transcription theme	**Theme**: Social stigma **Subthemes**: Blame, Bad-parent, False illness, Judgment, Exclusion, and Ignorance. **Explanation**: Parents are stereotyped as bad parents and at fault for their child's disorder. They are often excluded from social and family activities, and parent groups. Such judgments are based on an ignorance of the child's disorder, the difficulties experienced, and a belief that childhood mental illness does not exist; the child is naughty, and in need of better parenting.	**Theme**: Resilience **Subthemes**: Support, Optimism, Positive self-belief, Acceptance, Deflection, and Knowledge. **Explanation**: Others' support, and empathy promoted coping. Resilient parents were optimistic and positive about the future. Their self-belief as good parents was stronger than the stigma of being a bad parent. They had accepted their child and did not self-blame. The diagnosis helped; the more the parent knew about the child's problem, the stronger the resilience. This protected parents from bad-parent stigma.	**Theme**: Shame **Subthemes**: Embarrassment, Non-disclosure, Social comparisons, Within the Family. **Explanation**: Parents believed that their situation was shameful and that the child's behaviours were embarrassing. There was a desire to conceal information regarding the child's diagnosis and/or treatment as this was private family information and also to avoid feeling ashamed. Parents made downward comparisons between their own child and other children and their parent-self to other parents.	**Theme**: Child's stigma **Subthemes**: None **Explanation**: The child experiences his or her own stigma and parents are painfully aware of this. Parents witness their child being excluded, devalued and criticised. This causes them great concern for their child.

[1]A complete list of subthemes and themes can be obtained by contacting the first author.

Scribe (*in-person*) theme	**Theme**: Others' stigma **Subthemes**: Blame, Bad parent, Estrangement/ Exclusion, Stigma seed, Ignorance, Disorder not real. **Explanation**: Parents felt blamed for the disorder, are labelled as "bad" parents, and experience social exclusion. Stigma acts like a seed; once planted, it creates self-doubt about the parent-self. Others ignorant about the child's disorder; it does not exist, it is bad parenting.	**Theme**: Proof **Subthemes**: None **Explanation**: Parents needed to believe that they were "good parents" and that they were not to blame for their child's disorder. The proof that they were "good" parents was evident when the child's symptoms improved due to help and support provided by the parent.	**Theme**: Privacy **Subthemes**: Disclosure, None of your business. **Explanation**: Parents desired a right to privacy; no need to tell others about the child's disorder. This information was a private family matter. Privacy was also driven by a fear of stigma that may result when others are informed about the child's disorder.	**Theme**: Uncertainty/Worry **Subthemes**: None **Explanation**: Confusion and uncertainty about the child (and non-normative behaviour). Parents were aware that their child was different from other children and were concerned for the child's future due difficulties the child may face.

Note. Similarity value ranges based on the Koo and Li (2016) guidelines on inter-rater agreement.

Table 3. *Subtheme Examples at each Similarity Level for Scribe (in-person) to Transcription Comparisons*

	>90 (Excellent)	75-89 (Good)	50-74 (Moderate)	<49 (Poor)
Transcription subtheme	**Subtheme**: Blame **Explanation**: Parents feel others blame them for their child's disorder. **Data excerpt**: "Um, she never actually said 'I - blame – you,' um, her tone, um when she was talking to me about [son], um, intimated to me that I was doing something **wrong**. Um, not, not necessarily that I caused this, but I was [.hhh] making it worse. Um, and she did say, 'I don't know what you're going on about, because there's nothing **wrong** with him.'"	**Subtheme**: Support from others **Explanation**: Parents feel supported, understood, and cared for by others; not judged by others. **Data excerpt**: "And, um, my family's really supportive um. (hhh) So I mean they'll ask, mum 'specially, she'll ask about 'im [son] often and, me, like 'are you goin' alright?' and… they seem to get what goes on for 'im, which is, is, nice. They know I'm doing my best, it's nice when they say that."	**Subtheme**: Optimism **Explanation**: Parents' optimism for a good future for the child. **Data excerpt**: "but again I think, as he gets older, he does a bit better, keeps on going, he continues to get better; he'll do well. Y'know, he's smart and capable" [laughs].	**Subtheme**: False illness **Explanation**: Parents perceived that others believe that childhood mental illness does not exist, that the child is just naughty and in need of more discipline. **Data excerpt**: "↑ they don't believe in mental illness, like, um, they say there's nothing wrong with him [son]!; he's just naughty and needs a good **smack**"!
Scribe (*in-person*) subtheme	**Subtheme**: Blame **Explanation**: Parents feel others blame them for their child's disorder. **Data excerpt**: "She never actually said 'I blame you,' but her tone intimated to me that I was doing something **wrong**; not necessarily that I caused this, but that I was making it worse. She said, 'I don't know what you're going about, there's nothing **wrong** with him.'"	**Subtheme**: Acceptance **Explanation**: Parents feel that others understand the child and do not judge their parenting. **Data excerpt**: "And she [parent's own mother] said to them [neighbours], 'this is what behaviour is when it's not the child's fault, it's the ADHD; this is how it is, look all you like.' It's nice when someone defends you and tells others what it's actually like and that this is not about you being a bad parent."	**Subtheme**: Changes over time **Explanation**: The way parents feel about their child's mental illness, and their status as a parent of such a child changes over time to be more positive and optimistic. **Data excerpt**: How come it's different [seeing self as a good parent]? Partly time… understanding him [son]. We had some shaky times but we're on a much better pathway at the moment. More positive things are going on for both of us."	**Subtheme**: Estrangement **Explanation**: Friendships were compromised when a friend or family member stigmatised the parent and/or the child. This resulted in a termination of the friendship. This was felt as a regrettable, but often necessary, loss. **Data excerpt**: "I just don't talk to them anymore, they just have no idea what's going on for him [son]. It's a bit sad, because we were quite close."

Note. Similarity value ranges based on Koo and Li (2016) guidelines on inter-rater agreement.

Scribing from Video-Recordings of the Interview

In the event the scribe is unavailable, or the participant rejects their inclusion, defaulting to the electronic recording of the interview for analysis may be necessary. Thus, given that coding from in-person scribed data was found to result in less data, which, when analysed produced comparable subthemes (and to a slightly lesser degree, similar themes) as coding from transcription, we next explored if a second mode of scribing based on video-recordings of the interviews would produce a similar good-to-excellent overlap with transcription. To maintain independence, the video-scribe had not observed the interview in person. Scribing from video-recording (8,645 words) resulted in almost 75% less data than transcription (34,263words), and only 6.6% less data than in-person scribing. Thus, the amount of data produced across the two scribing modes was fairly consistent.

In examining the similarity between the subthemes and themes of the video-scribed data compared to those of the in-person scribed data, at least 80% of subthemes and almost 67% of themes overlapped between the two to a good to excellent extent (Table 1).

The number of, and similarity between, subthemes and then themes derived from video-scribed data versus transcription was then compared. Forty subthemes emerged from each of the methods. Ten themes emerged from the video-scribed data and nine from the transcriptions. As shown in Table 1, between 80% (Rater B) and 85% (Rater A) of subthemes overlapped between the two methods to a good to excellent degree. A lower good to excellent degree of overlap was found for almost 67% of themes. The figures for the subthemes are highly consistent with those obtained from the in-person scribe to transcription comparison (80%-82.5% versus 80%-85%) but were less so at theme level (67% versus 78%). Inspection of the extra theme identified in the video-scribed data revealed no comparative match greater than 20% with any transcription-derived theme. Further, this theme largely comprises subthemes rated as having poor similarity between the two methods.

Inter-Rater Reliability

To assess the extent to which the two independent raters were consistent in their assigning of similarity ratings, two-way mixed, average-measures intraclass correlation analyses (ICC; McGraw & Wong, 1996) were calculated for each block of comparisons. Overall, raters were highly consistent in their similarity ratings (ICC = .87 to .95; Table 1).

Economic Comparisons

To establish the economic viability of scribing, we compared the hours and costs involved in scribing versus transcribing. Based on our sample of 6 interviews (total interview time 234.66 minutes; $M = 39.11$, $SD = 6.40$), interviews that were transcribed took significantly longer to process into text format than when scribed in-person $t(5) = 7.96$, $p < .001$, $d = 3.25$, and when scribed from video $t(5) = 8.05$, $p < .001$, $d = 3.29$. Analysing transcribed data took significantly longer than in-person scribed data $t(5) = 4.85$, $p = .005$, $d = 1.98$, and video-scribed data $t(5) = 3.82$, $p = .012$, $d = 1.56$. Mean and standard deviation values are provided in Table 4.

A comparison of the total cost involved in collecting, processing, and analysing data for the six sample interviews is shown in Table 4. The comparison reveals that in-person scribing produced a net saving of AU$187.00 (US$108.00) per interview (of approximately 40 minutes each) using the Fair Work Commission's Market and Social Research Award (2010) minimum hourly rate for research assistants (i.e., $28.96 AUD) and minimum wage estimates

(i.e., $16.70 USD per hour; Bureau of Labor Statistics, 2015) as a basis for calculation. Scribing from video produced a comparative net saving (i.e., AU$191.17 [US$110.31] per interview).

Table 4. *Economic Comparison of Tasks Involved in Collecting, Processing, and Analysing Scribed Data Compared to Transcribed Data*

	Total Time (in minutes)		
Activity	*In-person* Scribe	*Video* Scribe	Transcription
Training	210.00	210.00	420.00
Interviewing	234.66	234.66	234.66
Creating text-based dataset	244.00 (*M* = 40.67, *SD* =	244.00 (*M* = 40.67, *SD* =	1655.00 (*M* = 275.83, *SD* =
Analysis	5.65)	6.56)	75.33)
Total	277.00 (*M* = 46.17, *SD* =	225.00 (*M* = 37.5, *SD* =	982.00 (*M* = 163.67, *SD* =
	23.07)	5.24)	79.39)
	965.66	**913.66**	**3219.66**
	Total Cost		
Currency	*In-person* Scribe	*Video* Scribe	Transcription
Australia (AUD)[1]	466.10	440.99	1588.75
United States of America (USD)[2]	268.70	254.30	916.18

Note. [1]Based on the Fair Work Commission's Market and Social Research Award (2010) minimum hourly rate for research assistances (i.e., $28.96 AUD); [2]Based on minimum wage estimates (i.e., $16.70 USD per hour Bureau of Labor Statistics, 2015).

Discussion

Despite its many benefits, qualitative research can be prohibitive given the considerable resources required to process and analyse transcribed data (Tessier, 2012). Further, transcription may not be possible in studies in which electronic recording of the interview is prohibited, precluded or unfeasible. As such, some have dispensed with transcription, instead using a scribe to take comprehensive notes during interviews and then analysing these data (e.g., Bex Lempert, 2016; Corrigan et al., 2015; Mowat, 2012). Whilst scribing offers an alternative to verbatim transcription and results in more manageable datasets, it is important to evaluate the impact of this alternative on the qualitative insights gleaned from subsequent analyses. In particular, compared to verbatim transcription, does using a scribe result in less of the interview content converted into a text-based dataset? If so, what is the extent of the data reduction compared to verbatim transcription, and does this reduction result in a loss to the richness of information available for thematic analysis of the interview data? The current findings indicate that the scribing method described herein results in a substantially smaller text-based dataset. Despite this, there was a high degree of overlap in the themes (highest at subtheme level) derived from the scribed data (both in-person and video-scribed) compared to those from the transcribed data. Both in-person and video-recording scribing were more time- and cost- effective than transcribing.

In describing a method for training and using scribes in qualitative research, we have outlined a process of video recording interviews, with concurrent note-taking by a scribe, followed by coding and thematic analysis. This process builds on those first created for the purposes of documenting and analysing field notes obtained during interviews (e.g., Halcomb & Davidson, 2006; McNall & Foster-Fishman, 2007). A fundamental issue with these earlier methods is that the interviewer is required to take notes. The interviewer, distracted by note-taking, is not free to focus on building and maintaining rapport, directing the interview, and adhering to protocols (Barker et al., 2016). The amount of detail captured is also compromised.

By shifting note-taking responsibility from the interviewer to the scribe, interview integrity is enhanced because the scribe can focus on capturing verbatim quotes and contextual detail. Such detail is important for the interpretation of meaning (Clausen, 2012).

This selective, yet judicious note-taking by scribes resulted in a reduction in the overall amount of data recorded. This addresses one of the key criticisms of transcription: that transcription results in voluminous and unwieldy datasets often containing extraneous detail that confuses the reader and impedes analysis (e.g., Evers, 2011; Kvale, 1983; Markle, West, & Rich, 2011; Tessier, 2012). There is a preference for lean datasets because they facilitate expedient analysis and reporting (Halcomb & Davidson, 2006; Neal et al., 2015). Our findings indicate that compared to transcription, both scribing in-person and scribing from video result in about a quarter of the volume of information being documented.

Despite less data recorded in scribed notes, that which remains still need to preserve the integrity of the analysis and the interpretations made from it (Bloom, 1993; Gravois et al., 1992). There is limited benefit to data reduction if the analytic strategies applied fail to yield findings comparable to that which would be obtained if the data were transcribed verbatim. In comparing the number of subthemes (and themes) derived from in-person scribed data to those obtained from the transcribed data, we found an equivalent number of subthemes (and themes) across the two methods. Although fine detail was not captured in the scribed notes, independent raters considered up to 82.5% of these subthemes to be highly similar between the two methods. Thus, the central messages being related remained, despite the data reduction.

At theme level, similarity was somewhat lower, with almost 78% of scribe-derived versus transcription-derived themes rated as similar between the two methods. We also found lower excellent-to-good similarity ratings at theme level when scribing from video-recordings of the interviews (i.e., 67%). Rather than a difference in the meaning of the interpreted data, the difference was in how the subthemes were organised into bigger/broader themes, which then drove the lower ratings at theme level. Variability in the configuration of subthemes in the formation of themes is not uncommon in qualitative research when more than one coder/analyst works with the data (Armstrong, Gosling, Weinman, & Marteau, 1994; Olszewski, Macey, & Lindstrom, 2007). To address this, current practice is to use a discrepancy management strategy (Campbell et al., 2013). Common methods require coders to independently segment and code interview data, then unite to discuss discrepancies in code names and meaning, and the construction of themes (e.g., Campbell et al., 2013; Garrison et al., 2006; Hill, Knox, Thomspon, Nutt-Williams, & Hess, 2005). By negotiating coding discrepancies, Garrison et al. (2006) increased agreement from 43% to 80% and Campbell et al. (2013) 54% to 96%. Although, discrepancy negotiation was undertaken during scribe training to help scribes understand the level of detail required, it was not used during method testing so as to evaluate the outcomes of scribing before any revisions were made to the subthemes/themes. It is possible that in adding a phase of discrepancy checking and rectification during the compilation of subthemes into themes, similarity at theme level may improve. Further, for consistency, researchers might opt to complete the thematic analysis themselves, rather than leave this to the scribe.

We also assessed the amount of data, as well as the number and content of subthemes and themes derived from video-scribed data. Again, despite less data being documented by the video-scribe compared to transcriptions, video-scribing also produced a high degree of similarity in subthemes (i.e., up to 85%). This finding has important implications for contingency planning in the event the scribe is unavailable or their inclusion rejected by the participant. Our results show that it is possible to default to the video recording without much compromise to the collection or analysis of the scribed data. Moreover, the video-recording also serves a valuable function in facilitating review and revision of notes, and the triangulation of themes (Asan & Montague, 2014; Garcez, Duarte, & Eisenberg, 2011).

However, there are situations in which the electronic recording of interviews is not practical, permissible, or agreeable to participants (e.g., Crichton & Childs, 2005). For example, prisons and other detention centres may prohibit the use of electronic recording equipment (Patenaude, 2004). Researchers must rely on memory and whatever limited notes they are able to jot down whilst attending to the maintenance of rapport, their own safety, and the conversation itself. Certain populations may find the recording equipment intrusive or even offensive (Ellis & Earley, 2006) and thus, prefer the scribe. In such cases, the inclusion of the scribe supports the capturing of important data, and as such, enables qualitative research to happen in contexts unsuitable for transcription. Further, researchers are increasingly accessing participants via interviews completed over the internet (e.g., via Skype; Lo Iacono, Symonds, & Brown, 2016) or video conferencing (Kazmer & Xie, 2008). Recording these exchanges for later transcription can present some technological challenges, which ultimately reduce the quality of the recorded output (Sullivan, 2012; Weller, 2015), and thus, impact on the quality of the transcription—if one can be produced at all. Scribing whilst the interview is live may be a valuable option for capturing comprehensive notes, potentially addressing some of these issues, as well as saving time and other resources.

Findings indicated a significant time saving in processing data into a text form (85% faster) and subsequent analysis (72% faster) using the in-person scribe method when compared to verbatim transcription. This saving in time, which accounts for scribe training, translates to a 70% cost saving, making the scribe method an economically better alternative to verbatim transcription. Using a scribe could save approximately AU$115 (US$70) per hour of qualitative interview analysis. This is not only beneficial from a funding perspective but may also increase the accessibility of qualitative research for student and/or community-level research where reviews of programs, interventions, or processes may have previously been prohibitive due to resources constraints. The development of more efficient methods to qualitative dataset creation, whilst still maintaining rigorous end-results—relative to the widely used and accepted method of verbatim transcription—can advance the reach of qualitative research (Neal et al., 2015; Tessier, 2012).

There are limitations to this study. We did not assess the influence of the scribe on the interview. The observer effect is well documented (Monahan & Fisher, 2010). It is suggested that by their mere presence, researchers change the dynamic of what is being studied (McDonald, 2005; Patton, 2002); this can be both positive (e.g., revealing of how individuals perceive themselves and want others to perceive them) and negative (e.g., self-censoring) (Monahan & Fisher, 2010). Some prefer to mitigate the observer effect (McDonald, 2005; Patton, 2002; West & Blom, 2017), whereas others see it as useful in revealing social interactions, relationships, or phenomena that might only become apparent because the researcher is present (Monahan & Fisher, 2010). Moreover, when scribes are used to document sensitive information, for example, during physician-patient interviews, the recording of data improves. This is because patients feel that the physician (interviewer) is more attentive (Pozdnyakova et al., 2018). Patients also report greater satisfaction regarding the interaction (Zallman et al., 2018). We attempted to limit any intrusive effects of the scribe by seeking participants' permission to include the scribe (permitting refusal without repercussion), explaining the scribe's purpose, and introducing the scribe. The scribe was unobtrusive, did not interact with the participant or interviewer, and sat outside the immediate interview space. Yet, the potential for a different exchange due to the scribe's presence, however unobtrusive, should be considered in future research.

We used relatively short segments for scribe training, and our test interviews were up to 63 minutes long. Thus, the data presented in the current study is the minimum for examining the analytic outcomes of the scribing method; but it may not be its maximum. In some forms of qualitative inquiry, such as ethnographic inquiry, interviewing can substantially exceed one

hour (Knoblauch, 2005; Smith & Osborn, 2008). As such, further testing is required to establish the utility of scribing over longer durations. Moreover, although we used existing interview data as training material, not all researchers will have access to such material. In these instances, pilot or test interviews could be used as an alternative (Baker, 1994; van Teijlingen & Hundley, 2001). To further build on this study, future research could consider broadening the scope of the testing of this method to other relevant domains of qualitative data collection (e.g., focus groups, observation studies).

In summary, researchers are commonly faced with decisions regarding which features of the content and structure of the conversation to transcribe or otherwise process into a text-based form (Bailey, 2008; Cook, 1990; Lapadat & Lindsay, 1999). The level of detail required is determined by the aims of the research and the type of analysis being done (Bailey, 2008; Halcomb & Davidson, 2006; Tashakkori & Teddlie, 1998). The decision should be made on the basis of what is most useful, effective, and efficient (Kvale, 1996; Tilley, 2003). Should the need arise to scribe the interviews instead of transcribe, the findings of our study provide preliminary evidence that scribing is a promising alternative to transcription. However, there are certain research paradigms that may be more appropriate for scribing; in particular, studies that seek to identify common ideas or meanings (e.g., those using an interpretive or descriptive phenomenology framework) and for those employing thematic or content analysis (such as the study on which the current study is based; i.e., Eaton et al., 2016). Such forms of inquiry do not necessarily require verbatim transcription (Halcomb & Davidson, 2006; Smith & Osborn, 2008). Scribing might also be a robust addition to ethnographic or anthropological studies, providing more detailed data than just the field notes that are commonly relied on in these studies (e.g., in REAM studies; Beebe, 2014; McNall & Foster-Fishman, 2007). Given the reductions in the amount of data documented and the qualitative differences this entails, scribing is potentially less suited to studies requiring detailed transcription of the specific linguistic elements of the dialogue (e.g., conversation, discourse, or narrative analysis). The "ums," "ahs" and tangential discussions are important pieces of information for such studies as it is the function and form of language itself that is of interested (e.g., Edwards & Lampert, 1993; Psathas & Anderson, 1990; Schiffrin, 1994; Smith, Jarman, & Osborn, 1999).

If not requiring such a level of detail, scribing has the potential to make accessible qualitative research projects previously unfeasible due to labour- and resources- intensive verbatim transcription. Further, using an in-person scribe to collect text-based data has utility in research environments in which verbatim transcription from electronic recording is prohibited. As such, scribing shows promise in the field of qualitative research.

References

Armstrong, D., Gosling, A., Weinman, J., & Marteau, T. (1994). The place of inter-rater reliability in qualitative research: An empirical study. *Sociology, 31*(3), 597-606. doi:10.1177/0038038597031003015

Asan, O., & Montague, E. (2014). Using video-based observation research methods in primary care health encounters to evaluate complex interactions. *Informatics in Primary Care, 21*(4), 161-170. doi:10.14236/jhi.v21i4.72

Atkinson, J. M., & Heritage, J. (1999). Jefferson's transcript notation. In A. Jaworski & N. Coupland (Eds.), *The discourse reader* (pp. 158-166). London, UK: Routledge.

Averill, J. B. (2002). Matrix analysis as a complementary analytic strategy in qualitative inquiry. *Qualitative Health Research, 12*(6), 855-866. doi:10.1177/104973230201200611

Bailey, J. (2008). First steps in qualitative data analysis: Transcribing. *Family Practice, 25*(2), 127-131. doi:10.1093/fampra/cmn003

Baker, T. L. (1994). *Doing social research* (2nd ed.). New York, NY: McGraw Hill Inc.

Barker, C., Pistrang, N., & Elliott, E. (2016). *Research methods in clinical psychology: An introduction for students and practitioners* (3rd ed.). Chichester, UK: John Wiley and Sons, Ltd.

Beebe, J. (2014). *Rapid qualitative inquiry: A field guide to team-based assessment* (2nd ed.). London, UK: Rowman and Littlefield.

Bentley, M. E., Pelto, G. H., Straus, W. L., Schumann, D. A., Adegbola, C., de la Pena, E., … Huffman, S. L. (1988). Rapid ethnographic assessment: Applications in a diarrhea management program. *Social Science and Medicine, 27*(1), 107–116. doi:10.1016/0277-9536(88)90168-2

Bex Lempert, L. (2016). *Women doing life: Gender, punishment, and the struggle for identity.* New York, NY: New York University Press.

Bloom, L. (1993). Transcription and coding for child language research: The parts are more than the whole. In J. A. Edwards & D. Lampert (Eds.), *Talking data: Transcribing and coding in discourse research* (pp. 149-166). Hillsdale, NJ: Lawrence Erlbaum.

Britten, N. (1995). Qualitative interviews in medical research. *British Medical Journal, 311*(6999), 251-253. doi:10.1136/bmj.311.6999.251

Bureau of Labor Statistics. (2015). *Occupational employment and wages: Social science research assistants* (19-406). Retrieved from http://www.bls.gov/oes/current/oes194061.htm

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semi-structured interviews: Problems of unitization and inter-coder reliability and agreement. *Sociological Methods and Research, 42*(3), 294-320. doi:10.1177/0049124113500475

Clausen, A. S. (2012). The individually focused interview: Methodological quality without transcription of audio recordings. *The Qualitative Report, 17*(19), 1-17. Retrieved from http://nsuworks.nova.edu/tqr/vol17/iss19/1

Cook, G. (1990). Transcribing infinity: Problems of context representation. *Journal of Pragmatics, 14*(1), 1-24. doi:10.1016/0378-2166(90)90061-H

Corrigan, P. W., Pickett, S., Kraus, D., Burks, R., & Schmidt, A. (2015). Community-based participatory research examining health care needs of African Americans who are homeless with mental illness. *Journal of Health Care for the Poor and Underserved, 26*(1), 119-133. doi:10.1353/hpu.2015.0018

Crichton, S., & Childs, E. (2005). Clipping and coding audio files: A research method to enable participant voice. *International Journal of Qualitative Methods, 4*(3), 40-49. Retrieved http://ejournals.library.ualberta.ca/index.php/IJQM/issue/view/365

Eaton, K. (2013). *The stigmatisation of the provision of alcohol and other drug services* (Unpublished thesis). School of Psychology, Edith Cowan University, Western Australia, Australia.

Eaton, K., Ohan, J. L., Stritzke, W. G. K., & Corrigan, P. W. (2016). Failing to meet the good parent ideal: Self-stigma in parents of children with mental health disorders. *Journal of Child and Family Studies, 25*(10), 3109-3123. doi:10.1007/ s10826-016-0459-9

Edwards, J. A., & Lampert, M. D. (Eds.) (1993), *Talking data: Transcription and coding in discourse research*. Hillside, NJ: Lawrence Erlbaum.

Ellis, J. B., & Earley, M. A. (2006). Reciprocity and constructions of informed consent: Researching with Indigenous populations. *International Journal of Qualitative Methodology, 5*(4), 1-9. Retrieved from http://www.ualberta.ca/~iiqm/backissues/5_4/pdf/ellis.pdf

Evers, J. (2011). From the past into the future: How technological developments change our ways of data collection, transcription and analysis. *Forum: Qualitative Social Research, 12*(1), Art. 38. Retrieved from http://nbn-resolving.de/urn:nbn:de:0114-

fqs1101381
Fair Work Commission (2010). *Market and Social Research Award 2010* (Cwlth) s.4 (Australia). Retrieved from: https://www.fwc.gov.au/documents/documents/modern_awards/pdf/MA000030.pdf
Garcez, A., Duarte, R., & Eisenberg, Z. (2011). Production and analysis of video recordings in qualitative research. *Educação e Pesquisa, 37*(2), 249-261. doi:10.1590/S1517-97022011000200003
Garrison, D. R., Cleveland-Innes, M., Koole, M., & Kappelman, J. (2006). Revisiting methodological issues in transcript analysis: Negotiated coding and reliability. *Internet and Higher Education, 9*(1), 1-8. doi:10.1016/j.jheduc.2005.11.001
Giorgi, A. (2009). *The descriptive phenomenological method in psychology*. Pittsburgh, PA: Duquesne University Press.
Gravois, T., Rosenfield, S., & Greenberg, B. (1992). Establishing reliability for coding implementation concerns of school-based teams from audiotapes. *Evaluation Review, 16*(5), 562-569. doi:10.1177/ 0193841X9201600507
Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.
Halcomb, E., & Davidson, P. M. (2006). Is verbatim transcription of interview data always necessary? *Applied Nursing Research, 19*(1), 38-42. doi:10.1016/j.apnr.2005.06.001
Hamo, M., Blum-Kulka, S., & Hacohen, G. (2004). From observation to transcription and back: Theory, practice, and interpretation in the analysis of children's naturally occurring discourse. *Research on Language and Social Interaction, 37*(1), 71-92. doi: 10.1207/s15327973rlsi3701_3
Hill, C. E., Knox, S., Thomspon, B. J., Nutt-Williams, E., & Hess, S. A. (2005). Consensual qualitative research: An update. *Journal of Counselling Psychology, 52*(2), 196-205. doi:10.1037/0022-0167.52.2.196
Jarnow, J. (2017, April 26). Transcribing audio sucks – So make the machines do it [Blog post]. Retrieved from https://www.wired.com/2017/04/trint-multi-voice-transcription
Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13-31). Amsterdam, Netherlands: John Benjamins Publishing.
Johnson, B. E. (2011). The speed and accuracy of voice recognition software-assisted transcription versus the listen-and-type method: A research note. *Qualitative Research, 11*(1), 91-97. doi:10.1177/1468794110385966
Kazmer, M. M., & Xie, B. (2008). Qualitative interviewing in internet studies: Playing with the media, playing with the method. *Information, Communication and Society, 11*(2), 257-278. doi:10.1080/13691180801946333
Kieren, D. K., & Munro, B. (1985). *The observational recording dilemma* (Report No. 143). Edmonton, AB: University of Alberta, Department of Family Studies.
Knoblauch, H. (2005). Focused ethnography. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 6*(3), Art. 44. Retrieved from http://nbn-resolving.de/urn:nbn:de:0114-fqs0503440
Koch, T. (2006). Establishing rigour in qualitative research: The decision trail. *Journal of Advanced Nursing, 53*(1), 91-100. doi:10.1111/j.1365-2648.2006.03681.x
Koo, T. K., & Li, M.Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155-163. doi:10.1016/j.jcm.2016.02.012
Kvale, S. (1983). The qualitative research interview. *Journal of Phenomenological Psychology, 14*(1), 171-196. doi:10.1163/156916283X00090
Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing.* Thousand Oaks, CA: Sage.

Lapadat, J. C., & Lindsay, A. C. (1999). Transcription in research and practice: From standardization of technique to interpretive positionings. *Qualitative Inquiry, 5*(1), 64-86. doi:10.1177/107780049900500104

Lincoln, Y. S., & Guba, E. G. (2003). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The landscape of qualitative research: Theories and issues* (2nd ed., pp. 253-291). London, UK: Sage.

Lo Iacono, V., Symonds, P., & Brown, D. (2016). Skype as a tool for qualitative research interviews. *Sociological Research Online, 21*(2), 1-15. doi:10.5153/sro.3952

MacLean, L. M., Meyer, M., & Estable, A. (2004). Improving accuracy of transcripts in qualitative research. *Qualitative Health Research, 14*(1), 113-123. doi:10.1177/1049732303259804

Manderson, L., & Aaby, P. (1992). An epidemic in the field? Rapid assessment procedures and health research. *Social Science and Medicine, 35*(7), 839-850. doi:10.1016/0277-9536(92)90098-B

Markle, T. D., West, R. E., & Rich, P. J. (2011). Beyond transcription: Technology, change, and refinement of method. *Forum Qualitative Social Research, 12*(3), Art. 21. Retrieved from http://nbn-resolving.de/urn:nbn:de:0114-fqs1103216

McDonald, S. (2005). Studying actions in context: A qualitative shadowing method for organizational research. *Qualitative Research, 5*(4), 455-473. doi:10.1177/1468794105056923

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30-46. doi:10.1037/1082-989X.1.1.30

McNall, M. A., & Foster-Fishman, P. G. (2007). Methods of rapid evaluation, assessment, and appraisal. *American Journal of Evaluation, 28*(2), 151–168. doi:10.1177/1098214007300895

Monahan, T., & Fisher, J. A. (2010). Benefits of "observer effects": Lessons from the field. *Qualitative Research, 10*(3), 357-376. doi:10.1177/1468794110362874

Mowat, I. J. (2012). 'It wasnae (only) me': The development of self-regulation and self-responsibility in pupils experiencing social and emotional behavioural difficulties. In J. Visser, H. Daniels, & T. Cole (Eds.), *Transforming troubled lives: Strategies and interventions for children* (Vol. 2, pp. 251-271). Bingley, UK: Emerald Group Publishing Ltd.

Neal, J. W., Neal, Z. P., van Dyke, E., & Kornbluh, M. (2015). Expediting the analysis of qualitative data in evaluation: A procedure for the rapid identification of themes from audio recordings (RITA). *American Journal of Evaluation, 36*(1), 118-132. doi:10.1177/1098214014536601

Oliver, D. G., Serovich, J. M., & Mason, T. L. (2005). Constraints and opportunities with interview transcription: Towards reflection in qualitative research. *Social Forces, 84*(2), 1273-1289. doi:10.1353/sof.2006.0023

Olszewski, B., Macey, D., & Lindstrom, L. (2007). The practical work of <coding>: An ethnomethodological inquiry. *Human Studies, 29*(3), 363-380. doi:10.1007/s10746-006-9029-2

Ortlipp, M. (2008). Keeping and using reflective journals in the qualitative research process. *The Qualitative Report, 13*(4), 695-705. Retrieved from http://nsuworks.nova.edu/tqr/vol13/iss4/8

Patenaude, A. (2004). No promises, but I'm willing to listen and tell what I hear: Conducting qualitative research among prison inmates and staff. *The Prison Journal, 84*(4), S69-S91. doi:10.1177/0032885504269898

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Perrier, M-J., & Kirkby, J. (2013). Taming the 'Dragon': Using voice recognition software for transcription in disability research within sport and exercise psychology. *Qualitative Research in Sport, Exercise, and Health, 5*(1), 103-108. doi:10.1080/2159676X.2012.712996

Phillippi, J., & Lauderdale, J. (2017). A guide to field notes for qualitative research: Context and conversation. *Qualitative Health Research, 28*(3), 381-388. doi:10.1177/1049732317697102

Pozdnyakova, A., Laiteerapong, N., Volerman, A., Feld, L. D., Wan, W., Burnet, D. L., & Lee, W. W. (2018). Impact of medical scribes on physician and patient satisfaction in primary care. *Journal of General Internal Medicine, 33*(7), 1109-1115. doi:10.1007/s11606-018-4434-6

Psathas, G., & Anderson, T. (1990). The 'practices' of transcription in conversation analysis. *Semiotica, 78*(1-2), 75-99. doi:10.1515/semi.1990.78.1-2.75

Schiffrin, D. (1994). *Approaches to discourse*. Oxford, UK: Basil Blackwell.

Seaman, C. B. (1999). Qualitative methods in empirical studies of software engineering. *IEEE Transactions on Software Engineering, 25*(4), 557-572. doi:10.1109/32.799955

Shaw, R. (2010). Embedding reflexivity within experiential qualitative psychology. *Qualitative Research in Psychology, 7*(3), 233-243. doi:10.1080/14780880802699092

Smith, J. A., Jarman, M., & Osborne, M. (1999). Doing interpretive phenomenological analysis. In M. Murray, & K. Chamberlain (Eds.), *Qualitative health psychology: Theories and methods* (pp. 218-240). London, UK: Sage.

Smith, J. A., & Osborn, M. (2008). Interpretive phenomenological analysis. In J. A. Smith (Ed.), *Qualitative psychology* (pp. 53-80). London, UK: Sage Publications.

Sturges, J. E., & Hanrahan, K. J. (2004). Comparing telephone and face-to-face qualitative interviewing: A research note. *Qualitative Research, 4*(1), 107-118. doi:10.1177/1468794104041110

Sullivan, J. R. (2012). Skype: An appropriate method of data collection for qualitative interviews? *The Hilltop Review, 6*(1), 53-60. Retrieved from: http://scholarworks.wmich.edu/hilltopreview/vol6/iss1/10

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology.* Thousand Oaks, CA: Sage.

Tessier, S. (2012). From field notes, to transcripts, to tape recordings: Evolution or combination? *International Journal of Qualitative Methods, 11*(4), 446-460. doi:10.1177/160940691201100410

Tilley, S. A. (2003). Challenging research practices: Turning a critical lens on the work of transcription. *Qualitative Inquiry, 9*(5), 750-733. doi:10.1177/1077800403255296

van Teijlingen, E. R., & Hundley, V. (2001). The importance of pilot studies. *Social Research Update, 35*(Winter), 1-4. Retrieved from http://sru.soc.surrey.ac.uk/SRU35.pdf

Weller, S. (2015). *The potentials and pitfalls of using Skype for qualitative (longitudinal) interviews: National Centre for Research Methods working paper* (Report No. 3757). Retrieved from the National Centre for Research Methods: http://eprints.ncrm.ac.uk/3757/1/Susie%20Weller.pdf

West, B. T., & Blom, A. G. (2017). Exploring interviewer effects: A research synthesis. *Journal of Survey Statistics and Methodology, 5*(2), 175-211. doi:10.1093/jssam/smw024

Yilmaz, K. (2013). Comparison of quantitative and qualitative research traditions: Epistemological, theoretical and methodological differences. *European Journal of Education, 48*(2), 311-325. doi:10.1111/ejed.12014

Zallman, L., Finnegan, K., Roll, D., Todaro, M., Oneiz, R., & Sayah, A. (2018). Impact of medical scribes in primary care on productivity, face-to-face time, and patient comfort. *Journal of the American Board of Family Medicine, 31*(4), 612-619.

doi:10.3122/jabfm.2018.04.170325

Author Note

Kim Eaton is a clinical psychologist registrar, and PhD researcher at the University of Western Australia (UWA) interested in the stigma and self-stigma associated with mental illness. Correspondence regarding this article can be addressed directly to: kim.eaton@research.uwa.edu.au.

Werner G. K. Stritzke is a senior lecturer in clinical psychology and the director of the clinical psychology training programs at UWA. Correspondence regarding this article can also be addressed directly to: werner.stritzke@uwa.edu.au.

Jeneva L. Ohan is a senior lecturer in child and adolescent psychology at UWA. Correspondence regarding this article can also be addressed directly to: jeneva.ohan@uwa.edu.au.

Funding: This research was funded by the Australian Government Research Training Program Scholarship, and The University of Western Australia School of Psychological Science.

Conflict of interest: The independent raters performed the similarity comparisons in a paid capacity.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the University of Western Australia Human Research Ethics Office.

Informed consent: Participants provided informed consent.

Article Citation

Eaton, K., Stritzke, W. G. K., & Ohan, J. L. (2019). Using scribes in qualitative research as an alternative to transcription. *The Qualitative Report*, *24*(3), 586-605. Retrieved from https://nsuworks.nova.edu/tqr/vol24/iss3/12