37. Using software to support qualitative data analysis

Trena M. Paulus and Jessica N. Lester

INTRODUCTION

Qualitative data analysis software (QDAS) packages can be used by researchers to manage their entire study from start to finish, including the analysis of a wide variety of data useful for educational research. Packages can be used to analyze, for example, recordings of classroom interactions, observational fieldnotes, web pages, documents and records, social media conversations, images, videos, interviews, focus groups, Google Earth maps, and responses to open-ended survey questions. Notably, these software packages should not be confused with data analysis software such as SPSS, STATA, or SAS, which are often described as automatically analyzing the data. Rather, QDAS packages serve as platforms, or workbenches, in which the researcher can choose how to organize, store, and structure their unstructured or semi-structured qualitative data, and other documents related to a given study (see Elliott in this volume), in a systematic way that is aligned with their methodological approach.

This chapter will briefly describe the history of QDAS and the packages available as of early 2019, highlight the ways in which these platforms can be used to carry out a variety of analytic strategies, provide examples of how various components of the software can be used to do so, recommend best practices for describing the use of QDAS in research reports, and provide guidance for selecting a QDAS package.

HISTORY AND CURRENT STATUS OF QDAS

Prior to the development of QDAS, researchers relied upon manual methods for carrying out qualitative data analysis, physically cutting and pasting segments of transcripts, organizing data via paper-based folders, and using colored pencils/highlighters to take note of patterns, among other practices. Indeed, this remains true of some researchers today. With the rise of computing technology in the 1960s, some researchers began experimenting with using word processors and text retrievers to support their qualitative analysis (Wolski, 2018). In the early 1980s, the first two QDAS packages were developed: Ethnograph and Non-numerical Unstructured Data Indexing Searching and Theorising (NUD*IST).

Many of the earliest QDAS packages were not created with commercial intent. Rather, they were developed to meet the research needs of an individual (Wolski, 2018). Wolski (2018), who conducted 20 in-depth interviews with QDAS developers and propagators focused on its history, found that QDAS’ development was linked to an individual’s: (1) ‘previous experiences’ (e.g., computer programming experience); (2) ‘existing knowledge about computers’; and (3) ‘acquired knowledge’ (e.g., learning new programming languages) (p. 10). As Wolski noted, ‘knowing about qualitative research, computers and programming together provided...’
the formula for QDAS creation, whether this was done by one developer or in collaboration with a computer scientist’ (p. 10).

With an increase in desktop computing during the 1980s, QDAS gained more widespread use (Fielding, 2008). While according to Wolski the earliest propagation of QDAS packages occurred through word of mouth, the reach and use of QDAS quickly grew. By the early 1990s, the Computer Assisted Qualitative Data Analysis Software (CAQDAS) Network Project began. Located at the University of Surrey and funded by the UK-based Economic and Social Research Council (ESRC) from 1994 to 2011, the CAQDAS Network has served as a central site for discussion and activity related to the use of software for qualitative research. For instance, the very first CAQDAS conference was hosted at the University of Surrey and marked the emergence of an international network focused on the development and dissemination of QDAS packages. In the decades since, a strong network of developers, propagators, and committed users has emerged. Significantly, training events at major disciplinary conferences (e.g., the American Sociological Association and the American Educational Research Association) have become mainstays, with a special interest group focused on digital tools in qualitative research at the International Congress of Qualitative Inquiry marking the widespread interest in examining applications of QDAS to qualitative research practice.

Today, there are multiple QDAS packages available for researchers. As Wolski (2018) noted, there has been a 'rapid dissemination of knowledge and information about QDAS’ (p. 17), with many of the most commonly used packages offering similar features. As of 2019 when this chapter went to press, the three most robust and well-known packages were NVivo, MAXQDA, and ATLAS.ti. Relative newcomers, such as Dedoose, the first cloud-based package, and Quirkos, developed with ease of use in mind, have been introduced to meet the specific needs of researchers. Table 37.1 outlines some of the more well-known QDAS packages’ country of origin, date of initial commercial release, latest version, and supported platforms.

Even though QDAS packages have been available for over 30 years and continue to grow in popularity, the relationship between QDAS and the qualitative research community has remained an uneasy one (Davidson & di Gregorio, 2011; Paulus, Lester, & Britt, 2013). From paper and pen for recording observational notes to digital devices for recording interviews to QDAS to support analysis, qualitative researchers have long used tools to support the qualitative research process. It is QDAS, however, that has seemed to elicit the greatest concern from the scholarly community. Much of this concern has centered on the belief that the software produces the analysis for the qualitative researcher. Given qualitative researchers are generally described as the research instrument, the (mistaken) positioning of a software package as doing the analysis for the researcher grates against a core assumption of qualitative research.

We suggest that this belief is a mischaracterization of what is possible when using a QDAS package. Indeed, as Gibbs, Friese, and Mangabeira (2002) stated, QDAS is ‘just a tool for analysis, and good qualitative analysis still relies on good analytic work by a careful human researcher’ (p. 9). Another particularly common misconception about QDAS that has propelled arguments against its use has been the idea that it can only support one type of methodology, particularly grounded theory (Lonikila, 1995), or one analytic approach, specifically coding (Coffey, Holbrook, & Atkinson, 1996). Notably, the earliest QDAS packages were developed at the same time that grounded theory became popular, thereby perhaps explaining this misconception (Davidson & di Gregorio, 2011). And, indeed, qualitative coding is commonly associated with some of the earliest conceptions of grounded theory. Regardless, QDAS pack-
Table 37.1 Overview of QDAS packages (as of January 2019)

<table>
<thead>
<tr>
<th>Software</th>
<th>Country of origin and date of initial release</th>
<th>Supported platforms as of 2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLAS.ti</td>
<td>Germany, 1993 (Scientific Software Development GmbH)</td>
<td>Windows ATLAS for Mac ATLAS.ti Cloud Beta ATLAS.ti iPad ATLAS.ti Android</td>
</tr>
<tr>
<td>Dedoose</td>
<td>USA, 2009 (originally EthnoNotes) Cloud-based</td>
<td>Cloud-based</td>
</tr>
<tr>
<td>Elan</td>
<td>The Netherlands, 2002 (The Language Archive, Max Planck Institute)</td>
<td>Windows Mac Linux</td>
</tr>
<tr>
<td>f4 analyse</td>
<td>Germany, 2012</td>
<td>Windows Mac Linux iPad</td>
</tr>
<tr>
<td>HyperRESEARCH</td>
<td>USA, 1991 (ResearchWare)</td>
<td>Windows Mac</td>
</tr>
<tr>
<td>MAXQDA</td>
<td>Germany, 1989</td>
<td>Universal Windows/Mac MAXApp for iOS and Android</td>
</tr>
<tr>
<td>NVivo</td>
<td>Australia, 1981 (originally NUD*IST, now QSR International)</td>
<td>Windows Mac</td>
</tr>
<tr>
<td>QDA Miner</td>
<td>Canada, 2004</td>
<td>Windows Windows Windows</td>
</tr>
<tr>
<td>Quirkos</td>
<td>Scotland, 2013</td>
<td>Mac Linux Quirkos Cloud Android</td>
</tr>
<tr>
<td>Transana</td>
<td>USA, 2001</td>
<td>Windows Mac</td>
</tr>
</tbody>
</table>

ages can support whatever qualitative methodology and analytic approach a researcher might want to employ. Broadly, Jackson, Paulus, and Woolf (2018) argued that these entrenched misconceptions about the software often result from citation error rather than features inherent to the tools themselves, and warn that uncritically perpetuating these misconceptions discourages novice researchers from QDAS use.

To tackle a few of these misconceptions here, we highlight some things that QDAS packages do not do. First, as noted above, QDAS does not do analysis for a researcher. Just like Microsoft Word does not take control of or write papers on behalf of the author, QDAS does not take control of or do the analysis for the researcher. While packages like SPSS or SAS can be described as automatically computing, QDAS does not automatically analyze data or take control of any part of the analysis process. QDAS does not require quantification of data or use of a particular analytic approach. Despite claims by researchers that QDAS packages increase the rigor of analysis, simply using a package does not inherently make the analysis any better. Nor does it make the analysis process any easier or faster. That qualitative research is a time intensive process remains true even when using a QDAS package. As highlighted above, QDAS does not support only one methodological approach; rather, any approach can be enacted in QDAS, be it thematic analysis, grounded theory analysis, discourse analysis, or
ethnographic approaches, to name a few. In fact, there is not even one right way to use QDAS, as it can be leveraged to engage a variety of analytic practices (Silver & Woolf, 2015). Finally, QDAS does not inevitably impose distance between researchers and their data by placing a presumed barrier (i.e., computer) between them. Rather, as Jackson et al. (2018) argued, there are particular practices that should be engaged in – regardless of whether software is used or not – to ‘stay close’ to the data.

The advantage of QDAS packages (when used in an informed manner) is that they allow researchers to engage in the multiple analytic activities (that are central to qualitative research practice) in an efficient and effective way. We discuss some of these activities in detail next.

Analytic Activities Supported by QDAS

QDAS packages provide a suite of various management and analysis tools that can support researchers in a variety of research-related activities. Paulus and Lester (in press) argued that these packages can be used to support the entirety of the research process – from reviewing the literature to generating and analyzing data to writing up and representing findings, to name only a few activities. Notably, however, as Woods, Paulus, Atkins, and Macklin (2016) found in their content analysis of 763 empirical articles reporting use of NVivo or ATLAS.ti, the majority of researchers use the packages only for data management and analysis, with far fewer using the packages for other parts of the research process. Thus, there indeed remains a gap between how QDAS packages can be used and how researchers are currently leveraging them.

The methodological literature base includes several useful frameworks for helping qualitative researchers think about their use of QDAS when engaged in analytic activities. Christina Silver and Ann Lewins (2014), for instance, described five main categories of analytic activities that QDAS packages can support, including: integrating data sources and analytic approaches; exploring the content and structure of the data; organizing materials and ideas; reflecting on data, interpretations, processes, and results; and retrieving, reviewing, and rethinking ideas about the data. Five-level Qualitative Data Analysis (Silver & Woolf, 2015) is a framework that helps guide researchers through the process of translating their analytic strategy into the tactics of the software in order to harness it in a powerful manner. Woolf and Silver have offered detailed guidance on how to engage this perspective for users of ATLAS.ti, MAXQDA, and NVivo (see https://www.fivelevelqda.com).

More specifically, this method aims to ‘resolve the contrast between the emergent nature of analytic strategies and the more cut-and-dried nature of the software tactics in a conscious and effective manner’ (Silver & Woolf, 2015, p. 537). The five levels are designed to support researchers in (1) making the focus and purpose of the study explicit (knowing what it is they want to do); (2) articulating their analytic plans (whether software is to be used or not); (3) translating analytic tasks to software tools in a cyclical way (matching the tools to the analysis); (4) engaging individual software operations (using the tools in a straightforward way); and (5) ultimately using software operations in a customized way (harnessing the full power of the software in innovative ways). For instance, a qualitative researcher collecting walk-along interviews focused on people’s meaning-making in connection to space and place, may desire to make links between their interview data and geo-data that illustrate the places that were navigated and referenced during the interview. This is possible in a QDAS package, such as ATLAS.ti 8, wherein you can produce geo-documents that can be used to triangulate your
interview data with locations through Google Maps. Informed by the Five-level Qualitative Data Analysis method, a qualitative researcher establishes a clear sense of what they intend to do methodologically and analytically before ever touching the software; that is, they should know what they want to do with a given QDAS package prior to engaging with it. As such, the qualitative researcher leverages the QDAS features that are most useful to their analytic work in the most powerful way.

We next highlight some of the analytic activities that can be supported using QDAS tools, but will not necessarily be used by all researchers in all studies. Keep in mind that different QDAS packages can support these activities to a greater and lesser extent, and potential users should review the features available as part of the selection process. Some packages offer unique features (e.g., Transana supports the creation of transcripts (Jefferson, 2004) that use specialized symbols which can represent both what and how things are said; Dedoose supports multiple collaborators as it is cloud-based, etc.).

Organizing and Managing Data

At the most basic level, QDAS packages can be used as a ‘textual laboratory’ to organize data and other project-related documents all in one place (Konopasek, 2008). This ‘laboratory’ can be where the entirety of the project ‘lives’ for the duration of the study. This includes not only all types of data sources (e.g., images, text files, audio and/or video recordings, etc.), but also relevant project files (e.g., ethics board approval forms, researcher journals, team meeting notes, data collection instruments, etc.), and even the reviewed literature sources. Data and documents can be stored and organized in ways that are relevant to the unique project focus. This allows for greater visibility of the entire research process, which is particularly useful when generating an audit trail of the decision-making processes. It also supports collaboration in ways that manual analysis cannot. It is far easier to share, review, and merge copies of analyzed software files than it is to share stacks of highlighted data segments in an attempt to reach consensus or engage in collaborative analysis. And, if a research team is interested in engaging in ‘real-time’ collaborative analysis, some packages support this activity (i.e., cloud-based Dedoose, NVivo for Teams, Quirkos for Cloud, and ATLAS.ti Cloud).

Reviewing the Literature

Second, features within QDAS can be leveraged to complete an entire literature review within the software. Given literature reviews are quite similar to a thematic analysis of qualitative data (in this case the ‘data’ are published research studies or other literature sources), QDAS can be harnessed to support a paperless process from start to finish (Lubke, Britt, Paulus, & Atkins, 2017; O’Neill, Booth, and Lamb, 2018; Pope, 2016). The major packages now support direct import of bibliographic data from reference management software packages, such as EndNote, Mendeley, or Zotero, making the process even more seamless.

Generating Data

QDAS packages can also help with generating data. Social media data, such as Twitter feeds, and survey data, can be imported into the major packages. In addition, mobile apps are provided by some packages, while others have integrated cloud-based tools such as Evernote,
which can be used for fieldwork. Mobile devices and/or Evernote can be used to record notes, take photos or videos, or record geographical locations in the field. These data points then can be exported and uploaded to the full project file via Dropbox or iTunes file shares, and work can then be continued on the laptop or desktop station. QDAS packages can be used to maintain a research journal, and, if appropriate, the journal entries can be treated as data to be analyzed alongside the rest of the dataset. For example, a new data document within the package can be generated and used to record assumptions, biases, and unfolding understandings. In this way, all of the main activities of the study are maintained in one place, adding portability, visibility, and organization to the process.

Transcribing Audio-Visual Data

Next, most QDAS packages support transcription of audio and video recordings within the software itself. Keyboard shortcuts, foot pedals, and in some instances, specialized transcription symbols (e.g., Transana’s Jeffersonian symbols) are supported. The researcher still needs to actually do the transcription, but it can all take place in the same software package. Even more importantly, many of the packages provide the ability to synchronize the transcript with the original media files (audio and/or video recordings). This is useful so that when reading the transcript during analysis, a simple click on a particular part of the transcript will result in the original recording being played. In this way, the researcher can both read and listen to or view how something was said rather than relying on the typed transcript alone. As such, it is possible to stay closer to the data sources than in manual analysis. Notably, many of the QDAS packages provide tools to analyze audio or video recordings directly without first transcribing them. This is particularly useful for researchers working with large, interactional datasets in which they do not intend to transcribe the dataset in its entirety.

Analyzing and Interpreting Sources of Data

Of course, QDAS also supports the analysis of data, and provides multiple tools with which to do so. Audio and video files, as well as image-based data, can be analyzed directly in Transana, as well as in other packages (Estrada & Koolen, 2018). Many packages also have some automated analysis tools, including text search features, word frequency counts, word clouds, and auto-coding features, which can provide a bird’s-eye view of the data content, before moving into closer analysis. Memo tools support annotation of data while reading, viewing, or listening to it. Memos are a central analytic activity within many qualitative research designs (Miles, Huberman, & Saldaña, 2014) and can typically be customized for each study. For example, a researcher may want to produce a ‘free memo’ – one that is not linked to their data – that broadly describes their current analytic activities. Other studies may require that analytic or reflexive memos be directly connected to a particular segment of the dataset. Coding tools, too, allow a researcher to label segments of data – be it text-based data, video/audio recordings, or even images/photographs – in analytically meaningful ways. Once all related segments are labeled with codes, they can be retrieved together and reviewed. This is particularly useful as a researcher moves from low-level to high-level inferences. Once all the data are coded, the dataset can be queried for patterns of codes among various groups to help answer research questions that may require such comparisons. Visualization tools are also available in QDAS packages, and provide a way to see various relationships between
data sources, coded data, memos, and even the previously reviewed literature. For instance, a researcher might generate a network that examines the relationships across segments of interview transcripts, photographs collected while engaged in fieldwork, and observational fieldnotes. These sorts of visual representations can be useful when moving into higher-level interpretations and engaging in deeper exploration of a given dataset.

Writing and Reporting Research Findings

Finally, the entire research report can be written up within the same QDAS package. This is particularly useful as it allows production of the written product in the same location where all of the data, interpretations, and previous literature live – keeping everything close. For researchers wanting to move outside of QDAS for the writing stage, all analytic work can be exported into text files or spreadsheets, and can even be considered within other software packages that provide different tools (see Lester, 2015, for an example of this).

ACCOUNTING FOR THE USE OF QDAS IN THE RESEARCH REPORT

Given the flexibility of QDAS, as described in this chapter, it should be clear that there is not one way to use these tools. Thus, researchers need to be quite specific in their methods section about how, exactly, QDAS was used. However, this is not yet common practice. In a review of all peer-reviewed journal articles published from 1994 to 2013 that reported use of ATLAS.ti and QSR NVivo, Woods et al. (2016) noted that few researchers included details about how they used the software other than mentioning that they did so.

Such lack of detail may perpetuate persistent misconceptions – that QDAS can automatically analyze the data, for example, or that using QDAS inherently improves a study’s rigor (Jackson et al., 2018). To avoid this, Paulus, Woods, Atkins, and Macklin (2017) advised researchers to include the following details when reporting their use of QDAS in a study. First, given that the features available in a QDAS package change with each new version, it is important to mention which version of the software was used. Second, so as not to give the impression that the software, rather than the researcher, is doing the analysis, writers should strive to use active voice (‘the research team created quotations and assigned codes to the data’) rather than passive voice (‘ATLAS.ti was used to analyze the data’) when describing their use of the software. Third, because not all qualitative researchers are familiar with all QDAS packages, the researcher should provide a brief description of what the software is, what it was used for, why it was selected, and which features or tools were used and how. If possible, the researcher should include software outputs (e.g., code lists and definitions or visual representations) as part of the data display and findings in order to retain the connection between the use of the software and the final researcher interpretations. Finally, if increased rigor through use of the software is being claimed, the researcher must substantiate this claim with specific details of how the tool was used in such a way that the quality was improved.
SELECTING A QDAS PACKAGE

Deciding which QDAS package to use can be an overwhelming process. In this section, we provide some recommendations for selecting a package. First, if there is institutional access to and support for learning and using a particular package, it is probably best to start with the already available package. Often this will be ATLAS.ti, NVivo, or MAXQDA. In the absence of institutional support, it can be helpful to, at minimum, find out which software colleagues are using and then choose the same one, so that someone is nearby to ask for help.

If there is no institutional access to the software and no colleagues using QDAS, it may be worth starting with a free trial version. Most software companies provide free introductory webinars, face-to-face training, or online training, and there are also an abundance of video tutorials available on sites such as YouTube. We recommend attending training whenever possible, as the QDAS packages are complex and can be difficult to learn on one’s own.

While most of the QDAS packages are cross-platform, not all were developed as natively Mac or Linux programs. It can be useful to check to be sure that the software offers the needed features for the platform that will be used (e.g., check first to be sure that video analysis is available in the Mac version, or Windows version, etc.). This information should be available on the website or in the user manual.

Similarly, if there is a particular feature that is needed for analysis, check to be sure that the software package provides it. For example:

- Can XML files from citation management software be imported in order to analyze annotated PDFs for literature review purposes?
- Is there a mobile app that can be used for data collection in the field?
- Can the needed data type be easily imported and analyzed (e.g., video files, social media data, survey data, etc.)?
- Can audio/video files be imported, transcribed, and synchronized?
- Can team members view and analyze the data synchronously?
- Will team members need to work across different QDAS packages?

Not all QDAS packages as yet support the import of social media data, video analysis, internal transcription features, the use of a mobile app, or working via the cloud to collaborate in real-time with team members – so if these features are needed, a careful review of the choices must be completed before a particular package is selected. The CAQDAS network offers reviews of many packages that can assist in the decision-making process: [https://www.surrey.ac.uk/computer-assisted-qualitative-data-analysis/support/choosing](https://www.surrey.ac.uk/computer-assisted-qualitative-data-analysis/support/choosing)

Historically, it has been impossible to work across platforms; that is, analysis done in ATLAS.ti could not be imported or viewed within NVivo, or vice versa. In 2019, the QDA-XML exchange standard that enables interoperability between different software programs (Evers, 2018) was launched. If not all team members have access to the same QDAS package, it is important to choose a package that is participating in this exchange standard (for a list of participating companies, see [http://www.qdasoftware.org/](http://www.qdasoftware.org/)). Finally, it can be immensely helpful to read about others’ experiences of using QDAS to engage in specific projects, such as Oswald’s (2017) experience with MAXQDA, Paulus and Lester’s (2016) experience with ATLAS.ti, and LeBlanc’s (2017) experience with NVivo.
CONCLUSION

In this chapter, we offered a general overview of QDAS packages, illustrating how researchers might leverage such packages to support the entirety of the research process. We briefly overviewed the history of QDAS and highlighted how such packages came to be, along with common misconceptions about their capabilities. We provided a general overview of the analytic activities that a QDAS package can support, including: (1) organizing and managing data; (2) generating data; (3) transcribing audio-visual data; (4) analyzing and interpreting sources of data; and (5) writing and reporting research findings. Further, we described best practices related to reporting uses of QDAS when writing up a research report. To conclude, we offered guidance for how a researcher might go about selecting a QDAS package.

NOTES

1. CAQDAS, Computer Assisted Qualitative Data Analysis Software, is another common term used to describe software supporting qualitative analysis.

2. NUD*IST was the earliest version of what today is called NVivo.

REFERENCES


