

Choosing a qualitative data analysis tool:

A comparison of NVivo and Leximancer

### **Abstract**

As the use and availability of qualitative software analysis tools increases so does ambiguity regarding the choice of the most appropriate software in sport management research. This paper uses NVivo and Leximancer to analyse the same set of data derived from interviews with sport management experts on high performance sport to showcase the differences in findings depending on the type of software used. The findings alert sport management and social science researchers to the importance of closely examining the choice of software for qualitative data analysis. Researchers may base their decision on aspects such as the type and size of data set, their competence and skills in data interpretation and the level of engagement with data analysis they plan on undertaking. This paper discusses the implications for researchers in the sport management field using these tools for data analysis and highlights the value that software can add to the research effort.

*Keywords:* Sport management research, qualitative methods, data analysis software, coaching policies, Leximancer, NVivo

## **Introduction**

The role of software in assisting researchers to analyse data is increasingly important. In the social sciences, the Statistical Package for the Social Sciences (SPSS) has been the dominant tool to analyse quantitative data for many years (Coakes & Steed, 2009). A much wider range of tools is available to assist qualitative data analysis. Additionally, over the past five to ten years the mention of qualitative software tools in published works has increased (Jones & Diment, 2010). In terms of qualitative management and business studies, NVivo has been the most used software package (Jones & Diment, 2010). However, the use of Leximancer software as an alternative package has been growing steadily as more researchers become familiar with its qualities (Cretchley, Gallois, Chenery, & Smith, 2010a).

Whilst Leximancer provides a form of automated analysis based on the properties of texts, the use of NVivo requires the manual handling of data at various points (Jones & Diment, 2010). While there are benefits with both packages, it is obvious that Leximancer has grown in popularity in research projects where large quantities of qualitative data are involved. Perhaps surprisingly, it is unclear which software is more useful in optimising research outcomes. Additionally, there has been little research to date about which software package is best suited to analysing what form of data and what type of sport management research (e.g., case study, phenomenology). There has also been limited effort made to determine which software produces the more trustworthy results. Welsh (2002) concluded that many social science researchers do not have the knowledge to make an informed assessment of the different software choices available to them, thus, he contested that most base their decision on nothing more than a colleagues' recommendation. It is also true that many researchers, often exposed early in their career to one software program which they learn how to use, become reluctant to switch away due to a lack of time, expertise, or money

(Lu & Shulman, 2008). Further complicating the situation, Jones and Diment (2010) found that almost all qualitative data analysis packages use specific file formats and that interoperability between different packages is either extremely limited or unavailable. Therefore, if a researcher chooses to change or try different software with which to explore a dataset, they are forced to start the data analysis process from the beginning.

These challenges make choosing the best form of software at the outset a critical consideration as it can shape the very premise on which the research is based. With this in mind, the purpose of this study is to use both NVivo and Leximancer to analyse the same set of data in order to compare and contrast the functions, coding, scope of the software and subsequent findings. The question driving this study is ‘How do data analyses and findings vary depending on the software used?’ Previous studies have used both NVivo and Leximancer to assist data analysis for different sets of data in the same study. For instance, Myers, Blackmore, Smith and Carter (2012) used Leximancer to analyse open ended questions of a survey. The researchers then used NVivo to compare the answers of the survey with a report. This study builds on previous knowledge and contributes to the understanding of qualitative software analysis tools because it draws conclusions based on comparisons obtained from the use of the same data set. This paper uses data from a study on high performance policies of a national sporting body, Australian Canoe, to critically compare the use of NVivo and Leximancer in assisting the analysis of qualitative data drawn from the sport management setting. Even though the data examined in this study focused particularly on sport management, the findings have a much broader applicability across leisure studies and all social sciences.

### **Literature review**

In an assessment of qualitative software, Jones and Diment (2010) divided computer-assisted qualitative data analysis (CAQDA) tools into two categories including tools that emphasise the manual handling of data (e.g., NVivo, Atlas.ti) and tools that provide automated analysis based on statistical properties of text (e.g., Leximancer). NVivo originated from the later development and enhancement of the first qualitative data analysis software, Non-numerical, Unstructured, Data: Indexing, Searching and Theorising, or NUD\*IST as it was more commonly known in 1981. NUD\*IST helped researchers by indexing components of textual documents and searching for words and phrases in the data. The analysis (linking of themes, contexts and categories) was actioned and done by the researcher and not undertaken automatically by the computer (Kellett, 1990). This software has undergone many transformations since its original release and in 2012 the latest version, NVivo 10, was released (QSR International, 2013). Over the years, other qualitative data analysis software have also been developed (e.g., Atlas.ti and Leximancer). Leximancer technology was created following seven years of research and development at The University of Queensland by Dr Andrew Smith at the university's psychology department (Leximancer, 2013). It is a software tool designed for analysing text data and uses statistics-based algorithms to automatically analyse text and visually displays the selected information in the form of concept maps, network clouds, and concept thesauruses (Smith & Humphreys, 2006). Leximancer has the capacity to search, add, remove, and merge terms, as well as extract semantic (meaning) and relational information.

Even though, Leximancer and NVivo are two discrete software packages that can be used on similar sets of empirical materials (e.g., interview transcripts, documents, open survey responses), they actually work in very different ways, which renders their comparison meaningfully. With purpose-built tools for classifying, sorting and arranging information,

NVivo helps a researcher manage and organise data and facilitates the analysis of data, identification of themes, glean insight and developing conclusions. Critically, NVivo requires the researcher to code the data and to develop themes or categories. Therefore, one can argue that the data analysis is principally subjective (an underlying philosophy of the constructivist paradigm) and allows the researcher to engage more meaningfully in the analysis process. Leximancer, on the other hand, produces results without the requirement for manual intervention. The software has been developed to identify a number of concepts and interrelationships without the need for researcher intervention. According to Hansson, Carey and Kjartansson (2010), the main limitation of using Leximancer as an analytic tool is that while the software produces mapping and relational data, the researcher's skill in interpretation, which is effectively suppressed by the manner in which Leximancer analyses the data, is the key to successful qualitative research. NVivo is not without its critics too. Cretchley, Rooney and Gallois (2010b) argued that because software like NVivo requires the researcher to derive the list of codes and rules for attaching these to the data, there is a lot of bias involved. Hence, this researcher-driven approach requires checks of reliability and validity. Other researchers (Tobin & Begley, 2004) have suggested that qualitative researchers should move away from the language of positivist concerns with validity and reliability and embrace trustworthiness and authenticity as measures of rigour.

After comparing CAQDA packages for analysing comprehensive corpuses of qualitative data, Wickham and Woods (2005) argued that during the qualitative research process the researcher needs to interpret, transform and categorise disparity in the data, demonstrate crucial relationships and bring out theoretical conclusions. In addition, Miles and Huberman (1994) warned inexperienced researchers of the often surprising challenge included in transforming and analysing large corpuses of text.

Further to this, Cretchley et al. (2010b) made an important distinction between Leximancer and NVivo by suggesting that the former's grounded approach enabled the analyst to take an exploratory style, letting the list of concepts emerge automatically from the text. On that note, Bazeley (2013) argued that concepts do not 'emerge' from text. Rather, the researchers derive them as the process is not really automatic. Other qualitative content analysis software, such as NVivo, is useful when an 'a priori' model or set of factors exists. In other words, when there is an 'a priori' model with which to 'code' up the data, researchers can use this set of factors in NVivo. NVivo then allows the researcher to categorise and summarise the coded results easily. On the other hand, Leximancer is a useful tool when a researcher is exploring the textual data to attempt to uncover important factors (Davies, Presilla, Strathdee, & Thornicroft, 1994). In other words, it is highly useful when the researcher does not have an 'a priori' set of factors or a model by which to analyse the data.

#### *Software applications in sport management research*

Discussions on software features, strengths and weaknesses appear to dominate scholarly debate across various disciplinary areas of inquiry. For example, Leximancer has been used in the past for qualitative data analysis in academic research settings in business and the public sector, in social and cultural studies, and in research on education, leisure and tourism (Beamish, Bryer, & Davies, 2006; Fisher & Miller, 2008; Grimbeek, Bartlett, & Loke, 2004; Grimbeek, Bryer, Davies, & Bartlett, 2005; Rooney, 2005; Rooney, McKenna, & Keenan, 2006; Young & Denize, 2008). Similarly, NVivo has been popularly applied within the context of the social sciences, as well as fields such as forensics, tourism, hospitality, criminology and marketing (Bazeley, 2007; Veal, 2006). In studies that have used NVivo there is rarely any justification provided for the choice of software used; rather there is often a simple statement like 'NVivo was used to analyse the qualitative data'. Interestingly, studies that have used Leximancer have tended to justify their choice of

software and explain why they did not use NVivo (Davies, Green, Rosemann, Indulska, & Gallo, 2006). Therefore, it could be said that when NVivo is not suitable for data analysis, researchers feel compelled to justify an alternative software choice.

To identify the extent to which qualitative data-analysis software has been utilised in sport management related research, the authors performed a search for use of qualitative software tools in three mainstream sport management journals. These were the *European Sport Management Quarterly*, *Sport Management Review* and the *Journal of Sport Management* and included articles published through to the last issues of 2012. NVivo/NUD\*IST was reported as being used in 35 of the published articles and is the most popular form of software reported in qualitative studies published in the sport management journals, with Atlas.ti the next most popular with its use reported in 12 articles (Table 1).

Table 1 Overview of the use of qualitative software in sport management journals

| Journal        | European Sport<br>Management<br>Quarterly | Sport<br>Management<br>Review | Journal of<br>Sport<br>Management | Total |
|----------------|---|-------------------------------|-----------------------------------|-------|
| Period (years) | 2001-2012                                 | 1989-2012                     | 1987-2012                         |       |
| CAQDA          |   |                               |                                   |       |
| NVivo          | 9   | 14                            | 7                                 | 30    |
| Atlas.ti       | 4   | 3                             | 5                                 | 12    |
| NUD*IST        | 3   | 1                             | 1                                 | 5     |
| HyperRESEARCH  | 1   | 1                             | 1                                 | 3     |
| Alceste        | 1   | 1                             | 0                                 | 2     |
| Ucinet         | 1   | 0                             | 1                                 | 2     |
| MAXqda         | 0   | 2                             | 0                                 | 2     |
| Leximancer     | 0   | 0                             | 1                                 | 1     |



Even though Leximancer is popular in other academic fields, including tourism and event management (Scott & Smith, 2005; Swart, Linley & Hardenberg, 2012), only one sport management article in the three journals reviewed (Shilbury, 2012) was found to use Leximancer to assist with the data-analysis process. After analysing text of abstracts from sport management manuscripts and performing a content analysis of the reference list of the selected manuscripts, Shilbury (2012) suggested that Leximancer produced meaningful insights through concepts that were then clustered into higher-level themes, and depicted relationships between key concepts and themes.

Studies that have critically compared the use of different software programs are limited in the literature. For instance, Hansson et al. (2010) studied the potential of *Concordance*, *Atlas.ti* and *Leximancer* for analysing social scientists' use of *values* in journal articles. Their findings revealed that a combination of software tools yields useful information for further research on the contents of social science articles.

A growing body of sport management research has used software to facilitate qualitative data analysis as this form of inquiry has become more acceptable in recent years. However, sport management researchers appear to continue to favour software packages such as NVivo that are widely accepted by their peers rather than try or justify the use of less mainstream software packages such as Leximancer. Considering the investment in knowledge, time and money required for familiarisation and appropriate use of software, having the confidence to choose the right tool is important (Lu & Shulman, 2008). Therefore, it is essential to understand software differences, strengths and limitations for a much more informed decision and selection of software.

## **Method**

*Description of data*

The data used for the analyses in this study were the transcripts from 14 structured interviews with high performance managers, coaches and athletes in Australian Canoe, the national sporting body for sprint canoeing in Australia. The interview questions included an ‘a priori’ model; *the sport policy factors that lead to international sporting success* (SPLISS) (De Bosscher, De Knop, van Bottenburg, M., & Shibli 2006). The predetermined variables (i.e., policy factors) were classified using Houlihan’s (2009) three categories; Contextual, Processual, and Specific sport policies (see Table 2). The research question for the study was “What policies explain sprint canoe’s international elite athlete success?”

Table 2 Factors contributing to elite success adopted from De Bossscher et al. (2006) and Houlihan (2009)

| Categories        | Policy Factors                         |
|-------------------|--|
| Contextual        | Culture                                |
|                   | Funding                                |
| <b>Processual</b> | Partnerships                           |
|                   | Administration                         |
|                   | Talent Identification                  |
|                   | Athlete Progress                       |
|                   | Resources                              |
|                   | <b>Coaching</b>                        |
|                   | Planning & Evaluation                  |
|                   | Lifestyle                              |
| Specific          | Competition opportunities              |
|                   | Facilities & Equipment                 |
|                   | Sport Science & Sport Medicine Support |
|                   | Research                               |

Before importing data into Leximancer and NVivo, paragraph styles (e.g., heading 1 for interview questions, heading 2 for participant answers) were applied to the interview transcripts using Microsoft Word to differentiate between questions and answers. This early formatting was deemed necessary and was preferred to formatting the text after it was entered into NVivo or Leximancer. A total of 52 pages of data formatted using Calibri font, size 11, and 1.15 line spacing was produced.

#### *Data analysis process using NVivo*

Formatted data was entered into NVivo 10 for analysis and a new standalone project was created. All interview transcripts were imported into the *Internals* folder and grouped under a folder named *Interviews*. As the interviews contained the same questions and were consistently structured in paragraph styles *Auto-Coding* was employed to organise and create a node for each question and organise data into nodes. Further, to avoid confusion between questions and answers, sport policy factors and sub-factors, and the structure of nodes and sub-nodes, *unique codes* were assigned to each question prior to analysis using two letters and one digit code. For example, PC1 was assigned for Processual (category), Coaching (sport policy factor), and Interview Question 1 (see Table 3).

As detailed earlier, in order to group the responses of each participant on the same question, data was first auto-coded. Then, open coding (i.e., *Code* or *Label* words and phrases found in the transcript or text) was used to identify codes for the phenomenon under investigation (Corbin & Strauss, 1990). Last, using axial coding (i.e., the creation of *Themes* or *Categories* by grouping codes or labels given to words and phrases), common codes were grouped into three major themes corresponding to three main categories of Contextual, Processual, and Specific (Corbin & Strauss, 1990). The example of ‘Coaching’ (one of the processual policy factors investigated, highlighted using bold in Table 2) is used in Table 3 to

illustrate the coding process from (a) the interview questions on coaching before and after coding to (b) the nodes and (c) the sub-nodes generated for each heading.

Table 3 Process from Interview Questions to Nodes and Sub-nodes on Coaching using NVivo

| Interview questions on Coaching                                      |   | Interview questions on Coaching Coded |   | Nodes in NVivo 10 | Sub-nodes |
|--|---|---------------------------------------|---|-------------------|-----------|
| How would you describe the quality of coaching staff in canoe/kayak? | PC1. Quality of coaching staff in canoe/kayak international performance | Quality of Coach_PC1                  | Culture differences<br>European influenced<br>Lack of experience<br>Less female coach<br>Not many quality coaches<br>Part time attitude<br>Post career challenges<br>World class  |                   |           |
| How do you believe coaching can be improved?                         | PC2. Improvement  | Improvement_PC2                       | Clear direction from head coach<br>Coach development<br>Coach training skills<br>Employ Australian coaches<br>Good planning and organised<br>More funding to coaches<br>Team work needed<br>Training in communication/culture |                   |           |
| How does coaching impact international performance?                  | PC3. Coach impact on success  | Coach impacts_PC3                     | Importance at early stage<br>Important impact   |                   |           |
| How well is ‘coach education and ‘training’ organised?               | PC4. Appropriate coach education and training in canoe/kayak            | Coach education_PC4                   | Higher education needed<br>Not appropriate<br>Opportunities open<br>Time limited for education  |                   |           |
| How well is ‘planning and development’ for coaches organised?        | PC5. Appropriate planning and development for coaches in canoe/kayak    | Planning and development_PC5          | No development and monitoring<br>No trust among coaches<br>Recycled plan  |                   |           |

*Data analysis process using Leximancer*

Using Leximancer 4.0 a new project was created in the *Leximancer Projects* folder. All the pre-formatted interview transcripts were loaded through *Load Data*. In order to conduct an exploratory analysis, *Project Run* was activated to run the project using default settings. Using this mode, Leximancer generated a list of 41 *Word-Like* concepts in order of declining occurrence and one *Name-Like* concept; Australian Institute of Sport (AIS) (i.e., a name concept that represents proper name, title or location). Some of the 41 *Word-Like* concepts were similar in their meaning or were used interchangeably (i.e., ‘athlete’ and ‘athletes’; ‘coach’, ‘coaches’ and ‘coaching’; ‘money’ and ‘funding’; ‘sport’ and ‘sports’). These concepts were merged or collapsed into one. Also, there were some concepts that were not relevant to the research question or of low semantic meaning (e.g., ‘things’, ‘look’, ‘better’, ‘doing’). These automatically generated concepts were removed from the text data. The final number of usable key concepts was 25 *Word-Like* and 1 *Name-Like* (a concept list is available in Table 5).

## **Results**

Using summarised findings this section illustrates the analysis and interpretation of the same dataset using NVivo and Leximancer software respectively.

*Summary of findings using NVivo*

All the data were coded using the ‘a priori’ categories of Contextual, Processual and Specific policy factors. In order to give a better view of the analysis at the deepest level (sub-nodes), the authors will discuss the findings using the core theme identified; “Coaching”. The core theme is central in relation to other identified themes and it allows the explanation of much of the data in the phenomena studied (Sotiriadou & Shilbury, 2010). The analysis of data using NVivo showed that the expectations of coaches, who had a European origin, from

athletes during training were very different to that of the athletes and much of this had to do with cultural sensitivities. The participants suggested that these coaches *'don't know what to expect or how to behave and they cannot handle it [coaching Australian athletes]'* (Respondent 14, a head coach). To put this into context, much of the cultural clash between athletes and coaches is the result of the existing talent identification process that sprint canoe uses. Specifically, Australian canoe athletes tend to transition to flat-water canoe/kayak from surf lifesaving. As Respondent 6, an athlete explained *'The culture is a surf culture, which I don't think is a bad thing..... a bit more relaxed, and not as regimented as some other sports. I think if we try and change this we will lose many good athletes'*. The seemingly laid back athlete approach to high performance versus the highly disciplined and authoritative approach engendered in the European coaching style generates coach-athlete conflict at a variety of levels but particularly in relation to values, expectations and beliefs.

At the same time though, the majority of participants expressed a concern that there were very few Australian quality coaches to draw from as they were lacking depth and experience. Participants argued that the sport is *'getting good results but we do not have quality [Australian] coaches as well as we do not have quality of coaching that are producing athletes'* (Respondent 9, a physiologist).

This is, in part, because of the lack of job security for coaches. Respondent 14, a head coach for instance, claimed that *'not many coaches think of this as a full-time job, secure job. It is like a hobby when they are only employed part-time'*. Mostly though, the problem with the lack of Australian coaches appears to reside within an inadequate coach education system and limited development pathway. One of the participants explained, as *'we do not have any education for the high performance coaches nor does Australian Canoe do any coach education for any level coaching'* and *'a lot of old players become coaches, so they assume you will make a good coach, which is not the truth because they need more training'*

(Respondent 11, a HP coordinator). This lack of experienced coaches and pathways to develop Australian coaches seems to affect talent identification and the progression of athletes to higher levels of performance.

Overall, through the use of NVivo, and by way of open, axial coding and the development of concept maps by the researchers some interconnections between concepts and overarching themes were identified. The findings suggested there was a relationship between talent identification, and surf lifesaving, as well as a link with European coaches and the clash of coach-athlete cultures. This analysis allowed the researchers to draw the conclusion that improvements were needed in the areas of *coach development* and *coach training skills* with respect to communication and culture.

#### *Summary of findings using Leximancer*

Following the concept editing outlined in the method section, the *Concept Map* (see Figure 1) contained 11 themes and 26 concepts.

Figure 1 Leximancer mapping of concepts with themes

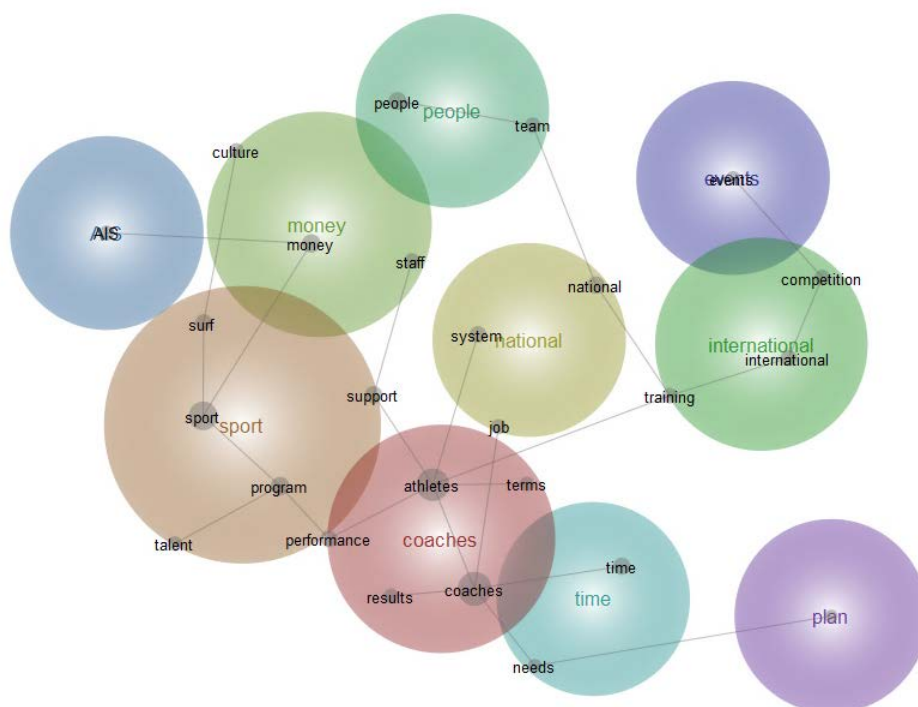




Figure 1 shows the dominant themes and associated concepts. The proximity of two concepts indicates how often they appear in similar conceptual contexts. The themes are the coloured circles around clusters of concepts. The lines show the most likely path between concepts. The connectivity score (see Table 4) reflects the degree to which the theme is connected to the other concepts in the map (Leximancer Manual, 2013).

The thematic summary shows that the themes that contribute to athletes' success as per the *heated map relevancies* and *connectivity score order* are 'coaches' (red), 'people', and 'sport', 'athletes', 'program', 'international' (green). The 11 themes and their connectivity are outlined in Table 4.

Table 4 List of the 11 themes in order of connectivity

| Theme         | Connectivity |
|---------------|--------------|
| Coaches       | 100%         |
| People        | 77%          |
| Sport         | 66%          |
| Athlete       | 66%          |
| Program       | 57%          |
| International | 27%          |
| Time          | 26%          |
| Surf          | 25%          |
| AIS           | 11%          |
| Events        | 8%           |
| Plan          | 6%           |

The 26 concepts within those themes, their count and relevance are displayed in Table 5. The Name-Like and Word-Like concepts are ranked based on the frequencies of occurrence in the transcripts. *Count* represents the number of times a concept appears in the entire corpus (2-sentence blocks), and *Relevance* represents the most frequent concept(s). For example, ‘coaches’ has a count of 174 and 100% relevance. Relevance shows proportionality of the concepts (representative) relative to each other.

Table 5 List of the 26 Name/Word–Like concepts

| <b>Related Name/ Word-Like Concepts</b> | <b>Count</b> | <b>Relevance</b> |
|---|--------------|------------------|
| AIS (Name-Like)                         | 38           | 22%              |
| Coaches                                 | 174          | 100%             |
| Athlete                                 | 139          | 80%              |
| Sport                                   | 138          | 79%              |
| Money                                   | 58           | 33%              |
| People                                  | 55           | 32%              |
| Time                                    | 50           | 29%              |
| Program                                 | 45           | 26%              |
| Performance                             | 43           | 25%              |
| Surf                                    | 37           | 21%              |
| Support                                 | 36           | 21%              |
| International                           | 34           | 20%              |
| Competition                             | 31           | 18%              |
| System                                  | 29           | 17%              |
| Team                                    | 28           | 16%              |
| Training                                | 28           | 16%              |
| Culture                                 | 28           | 16%              |
| National                                | 25           | 14%              |
| Job                                     | 24           | 14%              |
| Results                                 | 24           | 14%              |

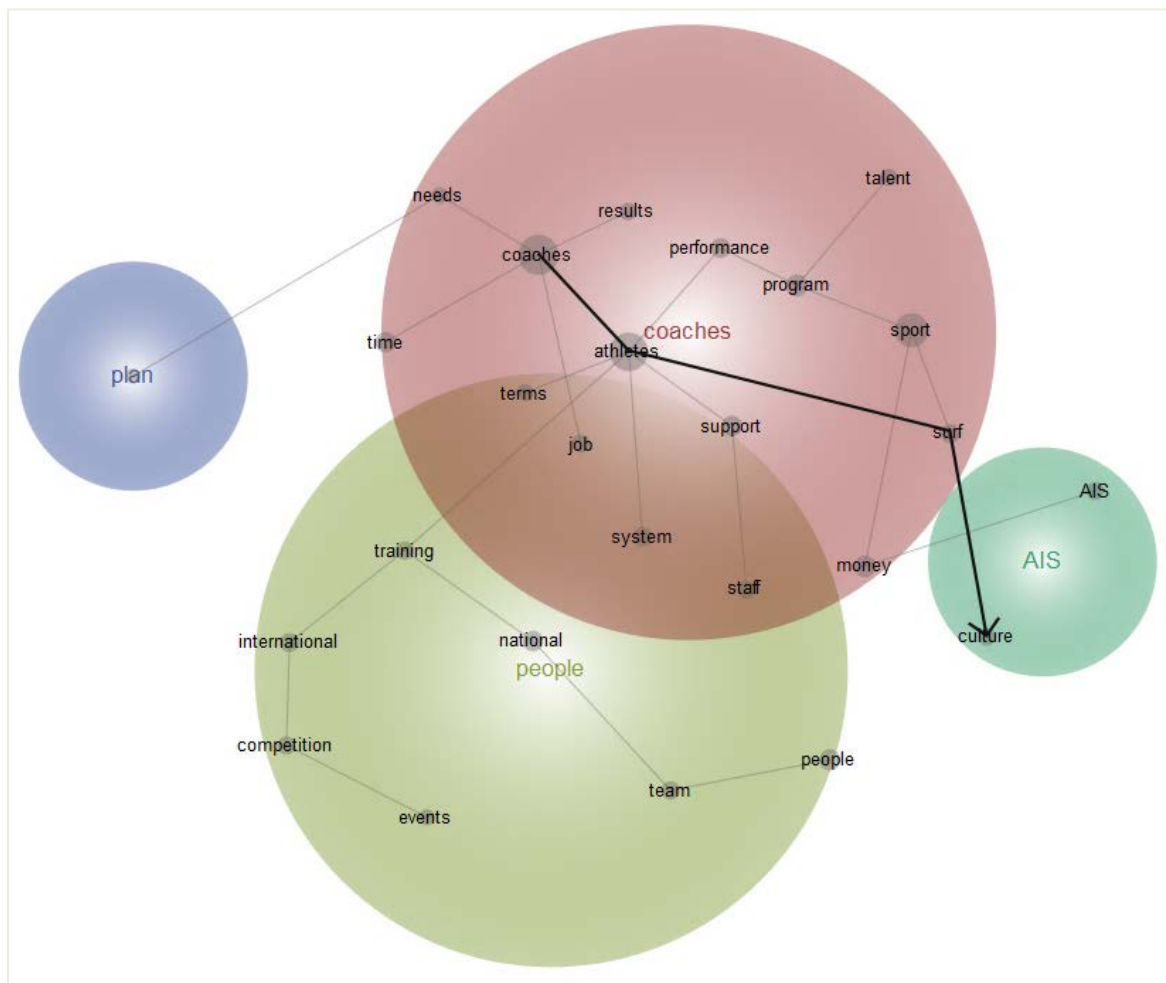
|        |    |     |
|--------|----|-----|
| Talent | 24 | 14% |
| Needs  | 23 | 13% |
| Staff  | 22 | 13% |
| Events | 22 | 13% |
| Plan   | 21 | 12% |
| Terms  | 19 | 11% |

In the concept map, the concepts appear as black labels. The relationships of the groups of concepts to one another are illustrated on the concept map with 95% visibility, 33% theme size and 96 degree of rotation. The concepts are clustered according to weight and relationship to create a concept cluster map. The concepts are contextually clustered on the map. In other words, concepts that appear together frequently in the text or in similar situations are close together on the map. The map is produced in colours with concepts sharing a theme in the same colour as their cluster group circle and cluster label (Leximancer, 2013). For example, the cluster of the conceptually related concepts of ‘system’, ‘national’, ‘team’, ‘training’ and ‘people’ is grouped in the theme ‘People’. The themes are heat-mapped to indicate importance according to the colour wheel. This means that the hot colours (red, orange) denote the most important themes, and cool colours (blue, green), denote the least important themes

Based on the count, relevance and connectivity, ‘coaches’ was the most important theme but why coaches were so central was not quite evident yet. What follows is some of the key findings on the theme ‘coaches’. Further examination in *high detail level* showed that the theme ‘coaches’ is in strong association with other identified concepts such as job (50%), system (34%), support (28%), terms (26%), athlete (26%), talent (25%), and results (25%) (see Figure 2). An exploration of the concept ‘job’ using a query, showed all the comments participants made on coach quality (e.g., *‘Some coaches have been fast tracked from scholarship coaches to higher roles way too quickly and are not experienced enough for their*

*job [an athlete]’), as well as the nature of the job being part-time, lacking prospect and not being a secure occupation (e.g., ‘Unlike tennis you can’t really get private lessons or make a career out of being a part-time coach. It does not feel like a secure job’, [a coach]).*

Figure 2 Relationship among ‘coaches’ with other concepts



In order to further explore the indirect relationships between concepts within the theme ‘coach’, the authors employed a *pathway* (showing in darker and thicker line in Figure 2) between the concept ‘coaches’ and the concept ‘culture’. This allowed navigating the most likely path in conceptual space from a start concept (i.e., ‘coach’) to an end concept (i.e., ‘culture’). The pathway showed that there is an interrelationship between two concepts. For instance, the following quote states: ‘We have a culture of individuals and that stems from

*our background in surf. If you look at our history 90% of our medals have been from individual athletes who have come from surf*" (a coach). Figure 2 indicates that other concepts are bypassed in order to move from 'coaches' to 'culture'. The pathway illustrated the connections emerging from the transcripts. That pathway is Coaches>Athletes>Surf>Culture. Preliminary exploration of the probability of each leg of the path and excerpts of text linking the two concepts involved in each leg indicated that 'coaches' has a connection with 'athletes' who are from and influenced by the surf culture. A program manager, for example, commented: *'We get surf athletes who are picked up and that elite coach takes them forward'*.

### **Discussion**

This study used NVivo and Leximancer to analyse the same set of qualitative data in the domain of sport management. The findings demonstrate that CAQDA tools facilitate coding and analyses and provide varied perspectives on a single dataset through the use of alternative types of software. This study allowed comparisons on several levels, including (a) comparison of the results, and their usefulness in answering the research question and in drawing meaningful conclusions and practical implications, and (b) comparisons of the analyses and processes that took place whilst using different software packages. These comparisons are discussed in this section. The implications of the findings in this study and the value that CAQDA tools add to the analysis are then presented.

In relation to answering the research question for the study that created the dataset, both NVivo and Leximancer resulted in a list of key sport policies that are perceived by respondents to contribute to athlete success. In both instances, a particular theme stood out with the emphasis on the role of coaches and coach development. Leximancer offered a greater list of potentially useful key-words (e.g., time) that were not identified in the NVivo

results that forced the researchers to look for their meaning. Some of these emerging concepts were atypical or erroneous.

Even though both tools provided results that served to assist the researchers in answering the research question posed, and although there were some common outcomes, the findings identified, and the level of their interpretation and practicality varied depending on the tool used. For instance, NVivo allowed for an elaborate 'story telling' on what sport policies influenced international success. This analysis offered a much more critical view of the results which raised the discussion and interpretation of findings to a higher level conceptually. Hence, the study findings reiterate the value of NVivo to studies using a grounded theory approach (see Bringer, Johnston, & Brackenridge, 2006). Leximancer, on the other hand, offered a visually attractive display of the key themes and concepts, their importance and proximity. Yet, developing a storyline from those diagrams was tedious and perplexing. Leximancer highlighted semantic relationships and proximity between words and was the only software that indicated concepts such as 'time' as a high frequency key-word related to sport policies.

NVivo allowed the researchers to draw conclusions and suggest that when making decisions such as recruiting a coach, sport managers should take into account the sports culture, athlete background and other policies in order to sustain their competitive advantage and avoid potential cultural conflicts (e.g., coach-athlete conflict). Subsequently, NVivo enabled the researchers to draw similarities and contrasts with the SPLISS model that was tested in the study and comment on the theoretical contribution of the study to sport policies. It was evident that, in contrast to the point argued by De Bosscher et al. (2006), strong beliefs about policymakers' inability to influence culture, a sport's culture, and in this instance the coach-athlete clash, can be influenced or avoided by policymakers or sport managers directly. Hence, recruiting coaches from within an Australian pool of talent or having to make

considerable changes to the coaching system itself could be avoided. The drawing of such conclusions or comments was not possible through a review of the first run of Leximancer results. Importantly, further tests conducted to identify a link between coaches, athletes and culture were only revealed in Leximancer because the researchers were prompted by prior NVivo findings. Even though NVivo appeared to be more useful in drawing meaningful results and implications, this could be because NVivo is recommended in studies using a priori variables (e.g., McDermott & Keating, 2012).

A comparison of the analyses and processes involved with the different software showed that both NVivo and Leximancer were purposeful in that each increased the overall level of organisation of the project and the ability to sort, retrieve and search the data (refer to Table 6 for a comparison of Leximancer and NVivo strengths and weaknesses). This finding is consistent with previous evaluations of CAQDA tools (Richards & Richards, 1987). Our analysis also confirmed previous findings that NVivo emphasises manual handling of data whereas Leximancer offers an automated analysis based on statistical properties (Jones & Diment, 2010). As a result, NVivo was more labour intensive and required the researchers to engage with the data and in the analysis process. The researchers found this process rewarding as it enabled them to ‘get closer’ to the data and understand in greater depth the issue under investigation, which is paramount in qualitative research. This proximity to the data and the importance of ascertaining what the data is telling the analyst is also supported by previous work (Silverman, 2011). Leximancer, on the other hand, allowed the researchers to explore the data in a timely manner with the list of concepts emerging automatically from the text without the need for researcher input (Coffey & Atkinson, 1996; Richards & Richards, 1994). However, it is argued that use of such a software package tended to alienate the researchers from the data itself to such an extent that in many instances, the researchers had to read the passages to make sense of the words that Leximancer decided to pick as

concepts or themes to verify their meaning. Whilst the use of NVivo is, to a certain extent, similar to manual coding, Leximancer was however different in the sense that the automated data analysis visually plotted groups of concepts and their relationship to each other.

Table 6 Strengths and Weaknesses of Leximancer and NVivo

|                   | NVivo   | Leximancer   |
|-------------------|---|--|
| <b>Strengths</b>  | <ul style="list-style-type: none"> <li>▪ Suitable for researcher driven study</li> <li>▪ Applicable for various types of data (text, photo, video, audio)</li> <li>▪ Efficient for small sample size</li> <li>▪ Effective for (semi) structured interview</li> <li>▪ Suitable for interpretive approach</li> <li>▪ Similar process with manual handling of data</li> <li>▪ Widely accepted for publication in three mainstream Sport Management Journals</li> </ul> | <ul style="list-style-type: none"> <li>▪ Suitable for exploratory and predictive study</li> <li>▪ Effective for large volumes of data</li> <li>▪ Quick identification of concepts</li> <li>▪ Objective data analysis as researcher bias coder subjectivity is removed</li> <li>▪ Reliable as there is minimal manual intervention from the researcher</li> <li>▪ Suitable for a positivistic approach</li> </ul> |
| <b>Weaknesses</b> | <ul style="list-style-type: none"> <li>▪ Time consuming to identify concepts for unstructured interviews</li> <li>▪ Data analysis is subjective as researcher bias may occur in manual coding (validity)</li> <li>▪ Reliability is questionable as there is intervention from the researcher</li> <li>▪ Data can be analysed without being specific about the actual embedded methods of analysis (Jones &amp; Diment, 2010)*</li> </ul>                            | <ul style="list-style-type: none"> <li>▪ Text analysis only</li> <li>▪ Not able to capture the style or implied tone of voice</li> <li>▪ Findings may contain unexpected or unexplained concepts and relationships</li> <li>▪ Limited used for publication in three mainstream Sport Management Journals</li> </ul>  |

*\*Note: Jones and Diment (2010) reported on the use or misuse of methodology in business and management related articles that used CAQDA tools and were published between 2005 and 2009. Of those articles 70% used NVivo/Nudist, 20% used Atlas.ti, 2% used Leximancer and 6% used other CAQDA software.*



These findings lend themselves to various implications for sport management researchers and other social scientists. The ease of using Leximancer may draw the attention of inexperienced researchers who would potentially run the danger of oversimplifying the data analysis process and drawing incorrect results and conclusions. Yet, Leximancer is considered a more objective analytic tool as the researcher bias coder reliability and subjectivity are removed (Isakhan, 2005). A key consideration when using NVivo is that the analysis can be subject to researcher bias occurring in the manual coding process. Hence, it could be argued that due to researcher intervention, the findings from NVivo are questionable unless trustworthiness and credibility of coding is offered. Some researchers have suggested that the searching tools in NVivo allow a level of interrogation of data hence an improved rigour of analysis (Welsh, 2002). By and large, validity and reliability in qualitative research has been debated over many years (Kirk & Miller, 1986) with some researchers suggesting that such terms are inappropriate and instead words such as 'trustworthiness' and 'quality of data' are preferred in the context of qualitative research (Miles & Huberman, 1994).

Commonly though, researchers use software such as NVivo to analyse qualitative data as their mantra, and not only do they not illustrate their measures for reliability and validity but they also fail to give detail on their analysis process. This is considered something of a failing though for as Jones and Diment (2010) contended, when analytical tools are used there needs to be a justification of 'why' they are used and of 'how' the tool was implemented. Merely stating that NVivo or Leximancer was used in the data analysis process is not considered sufficient.

Consequently, it can be argued that researchers in the sport management and the leisure studies field more broadly, would benefit from incorporating the use of both NVivo and Leximancer in the analytic process as the combination would allow a comparison and integration of results from discrete means of analyses (Bazeley, 2013). It is contended that the

use of alternative types of software yields useful information and provides different perspectives on a single dataset. However, the possibility of using dual CAQDA tools may be somewhat restricted in many settings due to investment, time and knowledge requirements associated with each.

### **Conclusion**

The comparison of the study results, the usefulness of these results in answering the research question and in drawing conclusions and the comparison of the processes involved whilst using different software helped draw certain conclusions. First, the data analysis process can be enhanced in various ways by the use of either NVivo and/or Leximancer. It is apparent that both of these software packages can speed up the analysis process markedly, making it easier for researchers to experiment with different codes, test different hypotheses about relationships, and draw diagrams of emerging theories. Second, it is clear that *contextual*, *specific* and *philosophical* considerations often drive the choice of software used. Contextual considerations include the researcher's time, knowledge, money and other resources that serve to influence the decision as to which software package to use. The specific requirements and design of a study should also be key in determining choice. When selecting a CAQDA tool, researchers should give due consideration to the differences resulting from their expertise and philosophical considerations such as their values and beliefs concerning research in general. To illustrate, in this study the researchers felt that NVivo brought them 'closer' to the data, offered a focus on meanings and a better connectivity within the data to show how one concept may influence another. This evaluation however needs to be contextualised on the basis of the researchers' *ontological* (i.e., the assumptions about how the world is made up) and *epistemological* (i.e., the beliefs about how one might discover knowledge about the world) standings (Burrell & Morgan, 1979, Bryman, 2012; Creswell, 2013; Jennings, 2010).

To conclude, a well-designed research study using appropriate qualitative software to assist in the analysis of datasets is a pathway to increasing the rigor and flexibility of the research. The authors contend that the choice of software package also has the potential to distance the researcher from data interpretations. This is particularly the case for novice researchers. Hence, for the researcher to stay clear of this risk it would be best to choose a combination of automatic and manual text analysis. Subsequently, a key message for social science researchers is that perhaps the integration of the results from separate analyses is more telling than their separation. Regardless of software choice, social science researchers need to critically explore the types of reports, visuals, summaries, or types of output they wish to produce and decide what software has the capacity to produce these outcomes as early as possible in the life of the research project planning process. Overall, this paper encourages the utilisation of a more informed decision on choosing and clearly discussing the application of software packages in analysing qualitative research and enables researchers to be specific about the actual embedded methods of analysis.

The limited set of data (14 transcripts) analysed with Leximancer and NVivo might present a limitation in this paper. Analysing a much larger amount of transcripts using NVivo would have been very time consuming whilst more data would have made no difference in the time required to analyse them using Leximancer. It is likely that NVivo allowed the researchers to categorise and summarise the coded results best because there were ‘a-priori’ concepts. Hence, it is uncertain whether the results would have favoured NVivo if there were no ‘a priori’ concepts in the study. Should this study be replicated on data where there are no ‘a priori’ concepts it may be the case that Leximancer would be more useful. Another possible limitation of this study is that it has used the basic functions of NVivo and Leximancer when both software tools offer a variety of more advanced functions than the ones used in this study. For example, both software tools have functions to compare different

groups of respondents with each other. These functions were not tested as this study only included a group of respondents from a single sport and from one organisation. Furthermore, NVivo has functions that run automatically such as *word clouds*, *tree maps* and *word trees* that were not explored in this study as the aim was to focus on the manual data analysis of this software. Further research with more data that uses additional functions of NVivo and Leximancer would add even more depth to researchers' understanding of the software strengths and weaknesses and enable an even more informed choice of software when analysing qualitative data. An outstanding area of research relates to technological advances and the continuing emergence of new or modified software. Such advances pose epistemological and ontological challenges on the impact that software innovation may have on research rigour and the way scientific and empirical research is conducted and reported.

### References

- Bazeley, P. (2007). *Qualitative data analysis with NVivo*. London: Sage.
- Bazeley, P. (2013). *Qualitative data analysis: Practical strategies*. London: Sage.
- Beamish, W., Bryer, F.K., & Davies, M.D. (2006). Teacher reflections on co-teaching a unit of work. *International Journal of Whole Schooling*, 2(2), 3-19.
- Bringer, J. D., Johnston, L. H., & Brackenridge, C. H. (2006). Using computer-assisted qualitative data analysis software to develop a grounded theory project. *Field Methods*, 18(3), 245-266.
- Bryman, A. (2012). *Social research methods*. Oxford: Oxford University Press.
- Burrell, G., & Morgan, G. (1979). *Sociological paradigms and organisational analysis: Elements of the sociology of corporate life*. Heinemann: London.
- Coakes, S. J., & Steed, L. (2009). *SPSS: Analysis without anguish using SPSS Version 14.0 for Windows*. John Wiley & Sons, Inc.

Coffey, A., & Atkinson, P. (1996). *Making Sense of Qualitative Data*. California: Sage.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, 13(1), 3-21.

Creswell, J. W. (2013). *Qualitative inquiry & research design: choosing among five approaches*. Los Angeles: SAGE Publications.

Cretchley, J., Gallois, C., Chenery, H., & Smith, A. (2010a). Conversations between carers and people with schizophrenia: a qualitative analysis using Leximancer. *Qualitative Health Research*, 20(12), 1611-1628.

Cretchley, J., Rooney, D., & Gallois, C. (2010b). Mapping a 40-year history with Leximancer: themes and concepts in the Journal of Cross-Cultural Psychology. *Journal of Cross-Cultural Psychology*, 41(3), 318-328.

Davies, I., Green, P., Rosemann, M., Indulska, M., & Gallo, S. (2006). How do practitioners use conceptual modeling in practice? *Data & Knowledge Engineering*, 58(3), 358-380.

Davies, S., Presilla, B., Strathdee, G., & Thornicroft, G. (1994). Community beds: The future for mental health care? *Social Psychiatry and Psychiatric Epidemiology*, 29, 241-243.

De Bosscher, V., De Knop, P., van Bottenburg, M., & Shibli, S. (2006). A conceptual framework for analysing sports policy factors leading to international sporting success. *European Sport Management Quarterly*, 6(2), 185-215.

Fisher, R., & Miller, D. (2008). Responding to student expectations: a partnership approach to course evaluation. *Assessment & Evaluation in Higher Education*, 33(2), 191-202.

Grimbeek, P., Bartlett, B., & Loke, K.K. (2004). Using Leximancer to identify themes and patterns in the talk of three high-distinctions students, in Bartlett, B., Bryer, F. and Roebuck, D. (Eds), *Education: Weaving Research into Practice*, Vol. 2, Griffith University, Nathan, pp. 122-8.

- Grimbeek, P., Bryer, F., Davies, M., & Bartlett, B. (2005). Themes and patterns from three Annual Meetings of the International Conference on Cognition, Language, and Special Education Research, in Bartlett, B., Bryer, F. and Roebuck, D. (Eds), *Stimulating the "Action" As Participants in Participatory Research*, Vol. 2, Griffith University, Nathan, pp. 101-14.
- Hansson, T., Carey, G., & Kjartansson, R. (2010). A multiple software approach to understanding values. *Journal of Beliefs & Values*, 31(3), 283-298.
- Houlihan, B. (2009). Mechanisms of international influence on domestic elite sport policy. *International Journal of Sport Policy*, 1(1), 51-69.
- Isakhan, B. (2005). *From despotism to democracy: Reporting Iraq's January 2005 election in the Australian and Middle Eastern print media*. Retrieved July 17, 2013, from <http://live-wirez.gu.edu.au/jea.papers/Isakhan%20.doc>
- Jennings, G. R. (2010). *Tourism research*. Milton, Qld: John Wiley & Sons
- Jones, M., & Diment, K. (2010). *The CAQDA Paradox: A divergence between research method and analytical tool*. Paper presented at the International workshop on Computer-Aided Qualitative Research Asia.
- Kellett, P. (1999). Organisational leadership: Lessons from professional coaches. *Sport Management Review*, 2(2), 150-171.
- Kirk, J. & Miller, M.L. (1986). *Reliability and Validity in Qualitative Research*. London: Sage.
- Leximancer. (2013). Leximancer: From words to meaning. Retrieved July 17, 2013, from <http://www.leximancer.com/cms/>
- Lu, C. J., & Shulman, S. W. (2008). Rigor and flexibility in computer-based qualitative research: Introducing the Coding Analysis Toolkit. *International Journal of Multiple Research*

*Approaches*, 2(1), 105-117.

McDermott, A. M., & Keating, M. A. (2012). Making Service Improvement Happen The Importance of Social Context. *The Journal of Applied Behavioral Science*, 48(1), 62-92.

Miles, M., & Huberman, A. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd Ed.). Thousand Oaks, CA: Sage.

Myers, S. A., Blackmore, M. J., Smith, T. F., & Carter, R. W. (2012). Climate change and stewardship: strategies to build community resilience in the Capricorn Coast. *Australasian Journal of Environmental Management*, 1-18.

QSR International (2013). *What's new*. Retrieved July 17, 2013, from <http://www.qsrinternational.com>

Richards, L., & Richards, T.J. (1987). Qualitative data analysis: Can computers do it? *Australian and New Zealand Journal of Sociology*, 23, 23-35.

Richards, L. & Richards, T. (1994). From filing cabinet to computer. In Alan Bryman, & Robert G. Burgess (Eds.), *Analysing Qualitative Data* (pp.146-172). London: Routledge.

Rooney, D. (2005). Knowledge, economy, technology and society: the politics of discourse. *Telematics and Informatics*, 22(4), 405-22.

Rooney, D., McKenna, B., & Keenan, T. (2006). Copyright and cultural production: a knowledge and wisdom theory perspective on education policy. *Policy Futures in Education*, 4(4), 380-95.

Scott, N., & Smith, A. E. (2005). Use of automated content analysis techniques for event image assessment. *Tourism Recreation Research*, 30(2), 87-91.

Shilbury, D. (2012). Competition: The heart and soul of sport management. *Journal of Sport Management*, 26(1), 1-10.

Silverman, D. (2011). *Interpreting qualitative data*. London: Sage.

Smith, A., & Humphreys, M. (2006). Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods*, 38, 262-279.

Sotiriadou, K., & Shilbury, D. (2010). Using Grounded Theory in sport management research. *International Journal of Sport Management and Marketing*, 8(3), 181-202.

Swart, K., Linley, M., & Hardenberg, E. (2012). A media analysis of the 2010 FIFA World Cup: A case study of selected international media. *African Journal for Physical, Health Education, Recreation and Dance*, 18, 131-141.

Tobin, G. A., & Begley, C. M. (2004). Methodological rigour within a qualitative framework. *Journal of advanced nursing*, 48(4), 388-396.

Veal, A. J. (2006). *Research methods for leisure and tourism: A practical guide*. Pearson Education

Welsh, E. (2002). Dealing with Data: Using NVivo in the Qualitative Data Analysis Process [12 paragraphs]. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 3(2), Art. 26. Retrieved 17 July 2013 from <http://nbn-resolving.de/urn:nbn:de:0114-fqs0202260>.

Wickham, M., & Woods, M. (2005). Reflecting on the strategic use of CAQDAS to manage and report on the qualitative research process. *The Qualitative Report*, 10(4), 687-702.

Young, L., & Denize, S. (2008). Competing interests: the challenge to collaboration in the public sector. *International Journal of Sociology and Social Policy*, 28(1/2), 46-58.