Evaluation Methods for Vulnerable Populations: The Case for the Deaf and Hard of Hearing

<p>Kathryn Cairns</p> <p>Centre for Program Evaluation University of Melbourne</p> <p>kathryn.cairns@unimelb.edu.au</p>		
<p>Patrick McLaren</p> <p>Centre for Program Evaluation University of Melbourne</p>	<p>Janet Clinton</p> <p>Centre for Program Evaluation University of Melbourne</p>	<p>Ruth Aston</p> <p>Centre for Program Evaluation University of Melbourne</p>

Keywords: evaluation ethics, participatory evaluation, collaborative evaluation, evaluation with Deaf stakeholders, evaluation with hard of hearing stakeholders, real-time captioning

Abstract

Conducting research and evaluation with vulnerable populations requires deliberate and mindful adherence to ethical standards and principles such as those outlined in the Belmont report, the Nuremberg Code and Declaration of Helsinki, the Australian Code for the Responsible Conduct of Research, the National Statement on Ethical Conduct in Human Research, as well as other professional standards of practice. When working with these populations, the principles of beneficence and non-maleficence, in particular, must be embedded from the project outset through to the dissemination of the findings. However, operationalising these principles can pose a challenge in practice. In this paper, we will present a case study of an evaluation of a real-time captioning program for Deaf/hard of hearing students, to illustrate the implementation of an inclusive and participatory evaluation methodological framework, in which adherence to ethical standards and principles was first and foremost.

Introduction

Research exploring literacy levels within the Deaf/hard of hearing population has revealed a gap in literacy attainment among this population, with Deaf/hard of hearing students typically demonstrating weaker literacy skills when compared to their hearing peers (King & Quigley, 1985). For example, a Victorian study found many Deaf/hard of hearing school leavers achieve a literacy level roughly equal to a Year 6 student (Walker & Rickards, 1992). Further, Brett (2010) has reflected on the challenges faced by Deaf/hard of hearing students due to the dominance of spoken language as the mode of instruction in academic settings, and noted that functional difficulty with spoken language excludes many students from participating in classroom learning activities. Research such as this highlights the need for assistive support services and technologies to help reduce the inequality in access to education and ensure that Deaf/hard of hearing students have an equal opportunity to reach their learning potential.

Literature Overview of Real-Time Captioning in the Classroom

Captioning is a form of assistive technology which has received attention in recent decades due to its potential to improve literacy and language skills amongst Deaf/hard of hearing populations. Research has demonstrated that classroom captioning can increase the comprehension of spoken material for Deaf/hard of hearing students (Boyd & Vader, 1972; Markham, 1989; Murphy-Berman & Jorgensen, 1980; Stinson, Stinson, Henderson & Miller, 1988). To illustrate this, in a study conducted by Boyd and Vader (1972), a group of Deaf/hard of hearing students were shown an educational television program with and without captions which were generated by a stenographer. Comprehension of the program's content was significantly higher after exposure to captioning (Boyd & Vader, 1972). Stinson et al. (1988), reported similar results among Deaf/hard of hearing students who were using captioning when compared with students using manual interpreting alone. The effects of captioning have also been observed with students for whom English is a second language. A study conducted by Markham (1989), examined the comprehension of video material by 76 university-level English as a Second Language (ESL) students and found greater comprehension with the captioned segment. Markham (1989) concluded that these findings support the proposition that learning capability is improved by simultaneous processing of different sensory modes.

In addition to comprehension, it has been suggested that captioning may generate an increased sense of inclusion and participation in the classroom for Deaf/hard of hearing students. Youdelman and Messerly (1996) explored the perceptions of students, teachers and note takers on the effectiveness of caption-like technology (computer-assisted note taking), and found that Deaf/hard of hearing students understood the lesson content better, could keep up with the pace of the lesson, and were better able to summarise the content of the lesson. The impact of this could lead to enhanced academic performance, but also to a more inclusive classroom environment with students able to participate more fully in the classroom as they can understand what is happening.

Real-time captioning (RTC) is a rapidly developing technology that differs somewhat from other forms of captioning. Whereas typical captioning technology is retroactively added, often in a simplified form, in transcripts or as subtitles (see Daelemans, Hothker & Sang, 2004), RTC is generated in real time with a delay of only seconds, thus providing viewers with access to captioned content almost instantaneously. As research indicates that Deaf/hard of hearing students experience difficulties understanding content delivered in English, it is conceivable that RTC within a classroom context could potentially contribute to improved literacy by providing verbatim text to supplement spoken language. As a consequence RTC provides students with an opportunity to connect written and spoken language in a concrete way, and within an interactive context. While clearly an important area of research, the technology is in its infancy, and thus there has been no research to date examining how RTC affects the participation and performance of Deaf/hard of hearing students in educational settings.

Evaluation Case Study: RTC in Secondary Schools

To explore the value of RTC for participating students and educators, an evaluation was commissioned by a governmental body, to evaluate a pilot program implemented in 2011-2012. The pilot program used a captioning infrastructure developed by a private captioning provider, to deliver captions of teacher talk in real time for Deaf/hard of hearing students in the classroom. The captioning infrastructure produces RTC by transmitting a teacher's speech to a remote respeaker who translates the spoken content into verbatim text using speech recognition software. This text can then be returned to the student in the classroom via iPad or laptop within 7 seconds. The evaluation was focussed on investigating the impact of the pilot program using this technology for Deaf/hard of hearing students in years 10 - 12 across 8 metropolitan and regional facilities. The aim of the evaluation of this pilot program was to explore the impacts of RTC on:

- a. Comprehension and academic performance outcomes, such as language and literacy levels; and
- b. Participatory outcomes such as inclusion, engagement in the classroom environment and school attendance.

The challenge for the evaluator

As alluded to earlier, working with populations like the Deaf/hard of hearing requires that particular attention is paid to ensuring their participation is meaningful, and that the evaluation results in some benefit for the community. Disengaged and vulnerable populations tend to be one of the hardest to engage in evaluation, but paradoxically they are often the groups that stand to benefit the most from evaluation findings and recommendations. Thus the imperative for developing inclusive evaluation methods is clear for these groups. Additionally, evaluators and researchers are ethically obligated to ensure participants do not experience harm by virtue of their engagement with the

evaluation. Beattie (2001) speaks directly to this, noting the complexity of the interplay between ethics, deafness and education.

Several ethical challenges were present from the outset of this evaluation. These ranged from engaging and communicating with the multiple stakeholder groups within the pilot, developing appropriate evaluation procedures and protocols, and minimising participant burden and fatigue. To meet these challenges, we embedded ethical and professional standards of practice, which guided evaluation procedures, within the overarching evaluation framework and methodology. These standards are briefly described below, followed by an overview of the evaluation framework.

Ethical Standards and Professional Codes of Practice

The Belmont Report, Nuremberg Code and Declaration of Helsinki are considered to be the cornerstone of medical human research ethics. These principles are arguably equally applicable to evaluation practice within the education sector, particularly the principles of beneficence and non-maleficence. Therefore, these became a touchstone for ethical practice within the context of the evaluation. In addition, professional standards for ethical conduct of research and evaluation practice guided the development and implementation of the evaluation. These included: the *Australian Code for the Responsible Conduct of Research*; the *National Statement on Ethical Conduct in Human Research*; the *Guidelines for the Ethical Conduct of Evaluations*, developed by the Australasian Evaluation Society ([AES], 2013), which are designed to suit the cultural, social and institutional contexts of evaluation in Australia and New Zealand; the *Program Evaluation Standards* (2nd Edition, Sage 1994); and the *American Evaluation Association's Guiding Principles for Evaluators* (2004). Further, as the Deaf community, typically, consider themselves Culturally and Linguistically Diverse (CLD), it was necessary to adhere to guidelines for working with CLD populations as outlined in the *Standards for Educational Assessment and Psychological Testing* (American Evaluation Research Association [AERA], 1999) within the context of conducting psychological or educational assessments. The evaluation was also subject to review by the University of Melbourne Human Research Ethics Committee. The ways in which these standards and guidelines have been used to inform practice will be outlined in further detail below.

Finally, the evaluation approach was conceptualised from a social justice perspective. Donna Mertens (2009), a researcher and program evaluator who has pioneered the philosophy of *transformative evaluation*, posits that evaluation should transform outcomes, and it should, where possible, result in improvements in the evaluand and the participants. In the case study presented in this article, the conduct of the evaluation was focussed on enhancing the RTC program for the benefit of the participating Deaf/hard of hearing students, with an emphasis on allowing the participants to have a voice through the evaluation.

An Inclusive Evaluation Framework

To meet the ethical challenges posed by the context for this evaluation, we employed

an evaluation framework which was participatory, collaborative, innovative, ethical, rigorous, and based on continuous feedback to stakeholders with a view to supporting program enhancement. The framework was underpinned by *The Framework for Program Evaluation of Public Health Initiatives*, developed by the Centers for Disease Control and Prevention (CDC&P, 1999). The model provides an overarching framework for the evaluation through the application of six steps ('stakeholder engagement'; 'program description'; 'focussing the evaluation design'; 'data collection'; 'justifying conclusions' and 'using and sharing lessons learned'). The framework is also clearly aligned to the aforementioned ethical and professional standards, to ensure the accuracy, reliability and validity of the evaluation process and findings. A mixed-methods research design was utilised to guide data collection and analysis. Such an approach combines qualitative and quantitative methods of inquiry, incorporating the strengths of both methods to better understand the research questions and strengthen the research design. In essence, a mixed-methods approach is; "... generative and open, seeking richer, deeper, better understanding of important facets of our infinitely complex social world". (Green, 2007, p. 20).

Finally, the evaluation framework was designed to be inclusive in nature. Briefly, an inclusive evaluation involves the systematic investigation of the merits or worth of a program to promote social change, with a focus on the involvement of all potential stakeholders of the project, with a particular focus on engaging those who have typically been under-represented (Mertens, 2009); in this instance, the Deaf/hard of hearing students in the pilot program. To support inclusive evaluation practice, evaluators must learn about the group that is under-represented in two ways: (1) by familiarising themselves with the literature, in this case research conducted with the Deaf/hard of hearing and the CLD within the context of education settings; and (2) by interacting with members of the community and wider stakeholders in a meaningful way (Mertens, 2009).

Methods and Principles

In the section that follows, we will outline the evaluation aims and methods and describe how we utilised the above evaluation framework to adhere to each of the principles of *respect for persons, justice, beneficence and non-maleficence* as outlined in the Belmont report (1979). In addition, reflections on the lessons learned through working with this population will be shared.

Summary of Evaluation Aims and Methods

The key aims of the evaluation were to explore the impacts of RTC on Deaf/hard of hearing students' comprehension, academic performance outcomes, inclusion, engagement and behaviour in the classroom environment. Further, we sought to examine unintended outcomes for key stakeholder groups, including teachers and parents. A brief overview of each method that was adopted within the evaluation is

given below, as a comprehensive discussion of evaluation methodology or findings is beyond the scope of this article. The evaluation reference group was involved in shaping and providing feedback on the development of each of the data collection instruments used within the evaluation. This section will then be followed with a commentary on how the evaluation methods and process adopted relate to each of these evaluative principles, with a particular focus on how this population was given the opportunity to have a voice in the evaluation. Further, we will discuss the ways in which the evaluation team promoted the use of data and subsequent conclusions and recommendations to inform the pilot program and future programs in the field of captioning in the classroom to benefit Deaf/hard of hearing students.

Literature Review: To ensure the rigor of an inclusive evaluation it is necessary to develop familiarity with the Deaf/hard of hearing population. As part of this process, the evaluation team conducted a review of the literature, focused on the impact of captioning on learning and participatory outcomes, particularly for the Deaf/hard of hearing. This literature review served as a foundation that guided the development of the broader evaluation methodology.

Online Survey with Students and Teachers: An online survey was delivered to participating teachers and students to gather background information, and experiences with and perceptions of the pilot program. The surveys also included a number of psychological and psychosocial constructs (for students) and questions targeted towards gathering the perceptions of RTC, transcripts and their implementation in the classroom.

Computer Assisted Telephone Interviews (CATI): CATI were used to survey parents of students involved with the RTC pilot. The interview protocol was designed to gather information about the parent's linguistic background, their understanding and perceptions of RTC and its impact on their child, and their overall satisfaction with the pilot.

Interviews with Program Staff: Semi-structured interviews with program staff were conducted to elicit information about the effectiveness and efficiency of the development and implementation process for the RTC pilot from the perspective of various program stakeholders.

School Data Audit: An electronic data collection template, designed to elicit information from each participating student on their achievement, behaviour, and attitude over the course of their involvement with RTC was sent to each facility coordinator to complete for each student. The key evaluation contact at the school then populated this template with relevant documentary information for each of the students participating in the pilot.

Language and Literacy Assessments: These assessments were conducted in two waves. Initially, the Compass test (Australian Council for Educational Research [ACER], n.d.) was chosen to assess the language and literacy skills of the students. Compass is an online literacy and numeracy assessment (the numeracy component was not administered in this instance) specifically designed for adolescents and young adults who have had limited or disrupted exposure to formal education. The second wave

of testing, in response to feedback from the reference group and key stakeholders, utilised the Woodcock Reading Mastery Test, a standardised assessment of language literacy, to gather normed data on participating students.

Analysis of Transcripts: The evaluation team was provided with a selection of classroom transcripts (that is, the printed version of the captions that appear onscreen during the lesson), which were then analysed for evidence of change in teacher practice or interactions within the classroom over the course of the term since the introduction of captioning. A rubric-based coding scheme was developed as informed by the literature and program stakeholders and was used to analyse the transcripts. The rubric had four main components: structure of the lesson; clarity of communication; clarity of communication for captioning; and technology and presence of captioning in the classroom.

Given the richness and scope of the multiple data sources, the triangulation of data from all these sources enriched the overall evaluation findings, allowed for multiple stakeholders to have a voice throughout the evaluation, and ultimately led to a greater level of confidence in the accuracy of the conclusions generated. The following sections provide a discussion on how the evaluation team upheld ethical principles throughout the implementation of these data collection procedures. Challenges and lessons learned in this process are also included.

Commentary on Adherence to Ethical Principles

As mentioned earlier in the article, the ethical principles that inform the majority of research as outlined by the Belmont report are discussed in this section, within the context of the evaluation case study.

Respect for persons

The principle of respect acknowledges individuals' and groups' autonomy and their right to make choices, to hold views and to take actions based on their values and beliefs (Belmont Report, 1979). Within the context of this evaluation case study, this meant ensuring that all those with a stake in the program and its evaluation had an opportunity to engage with the evaluation, and to share their views on the program. Considerable effort was made to ensure the reference group for the evaluation was inclusive. It thus comprised of stakeholders from the evaluation commissioner (Executive director, program manager, and technical advisors); participating schools (principals, coordinators, teachers of the Deaf/hard of hearing); and parents of students involved in the program. The evaluation team members who sat on this group contributed expertise and experience in the fields of evaluation, education and psychology and Deaf education. In assembling this group of stakeholders, we aimed to ensure our approach was appropriate for the population of interest and that we were cognisant of any particular concerns or issues for this population. As a consequence, we hoped to facilitate a greater level of engagement in the evaluation process from all stakeholders

and evaluation participants.

At the initial stages of the evaluation, the reference group were involved in a program logic workshop where the key stakeholders came to a shared vision with regard to the desired outcomes for the program, ways in which implementation could support the achievement of these outcomes; and the appropriate measures for the evaluation with a view to inform the development and improvement for future roll-out of the program. Thereafter, regular reference group meetings were held approximately every two months, with meetings focused on the evaluation process and development of protocols and instruments, and the immediate dissemination of any emergent findings. Finally, the reference group was involved in a summative workshop in which we presented the draft evaluation findings and asked them to reflect on these within the context of their own experience. Based on this workshop, we revised the original program logic model, and also added some recommendations to the final evaluation report to ensure these were consistent with the lived experiences of those involved in the project.

The principle of informed consent poses a challenge when working with CLD populations such as Deaf/hard of hearing groups, as traditional evaluation practices and procedures rely on written or oral communication and expressions of consent. In this case, the consent process needed to include the communication of information about the evaluation in a mode that all participants could understand. Thus, the evaluators needed to consider not only the preferred language and communication style of the participants, but also their literacy level. There was significant variation in the degree of deafness and preferred mode of communication across participants, with some students requiring everything to be translated into Auslan and others being able to comprehend English relatively easily by lip-reading, or hearing with the assistance of their cochlear implant. Further, there were considerable differences in the literacy level of participants which necessitated that protocols such as plain language statements and consent forms were written at a reading age that was accessible for all participants. In addition to ensuring that written communications were commensurate with the literacy level of the participants, information about the evaluation was also provided in person with an Auslan translator present when requested by the participants. To adhere with ethical guidelines regarding research with minors, the parents/guardians of the participants also received written information about the evaluation and were required to give their consent for their child to participate.

Justice

The principle of justice focuses on the equitable treatment of research participants, as well as ensuring an equitable distribution of benefits and burdens. As evaluators, we have a mandate to consider effective ways to balance the burdens of engaging in the evaluation process with the potential benefits. Considerable effort was made to collaborate with the reference group to create protocols that collected valuable information without impacting on participants in a negative way. In addition, data collection procedures were developed with a view to ensuring that data was collected

and analysed in a 'fair' way, such that the conclusions drawn from the data accurately reflected the outcomes and experiences of Deaf/hard of hearing participants involved in the program. Great care was taken to ensure that all participants, irrespective of their literacy level or communication preferences, were given equal opportunity to engage in a meaningful way with the evaluation.

Beneficence and non-maleficence

Fundamentally this evaluation was commissioned to determine the efficacy of RTC as an assistive technology to support the learning and achievement of Deaf/hard of hearing students. However, the notion of formative evaluation and continuous improvement was also central to the evaluation philosophy, which supports the principles of beneficence and non-maleficence. These principles assert that research and evaluation should seek to maximise benefits of the research project for the participants, whilst simultaneously minimising risks to the research subjects (Belmont Report, 1979). Deaf/hard of hearing populations are at greater risk of research-related harms, by virtue of their minority status, therefore it is critical that evaluators and researchers are mindful of this to ensure the evaluation design seeks to minimise harm and maximise benefit.

The main benefit of participation for the students in this evaluation was the opportunity to access a new technology which could support their engagement and inclusion in the classroom, and facilitate their learning more broadly. Further, through their participation in this evaluation, students also had the opportunity to contribute to the development and improvement of services for the broader Deaf/hard of hearing community. Evaluation participants trialled the effectiveness of a new technology designed to increase access to the spoken word, and ultimately to improve comprehension and learning, and, critically, provided their feedback on ways in which this technology and program could be improved; which information can subsequently be used to inform program development.

In relation to minimising potential harms resulting from the evaluation of the project, all data collected from students was anonymised. This data was not made available to teachers for review, despite some requests for this information, to preserve the authenticity of the consent process. Further, we triangulated several sources of achievement data (rather than just using one source of data) collected across time periods rather than relying solely on one observation point. The findings were also communicated with the necessary qualifying language to prevent misinterpretation or misuse of the results. Most importantly, the reference group was involved throughout to ensure that the evaluation was designed and implemented in a way that minimised harms to the participants.

To further illustrate adherence to these ethical principles, an example is provided below in the design and implementation of protocols to minimise participant fatigue in the language and literacy testing of participants.

Language and Literacy Testing Procedures

In the conceptualisation phase of the evaluation, it became evident that there is a dearth of appropriate assessment tools available to support accurate and meaningful assessment of the language and literacy skills of Deaf/hard of hearing students at a secondary school level. Given the reported delays in these skills for this population relative to their hearing peers, it was important to identify a tool that would fairly and accurately assess the student's literacy skills. After extensive research, and in consultation with the reference group, we chose to utilise an online assessment called Compass, which is produced by the Australian Council for Educational Research (ACER, n.d.).

Compass was selected for several reasons. First, it is delivered online, a space with which Deaf/hard of hearing students are potentially as conversant as their hearing peers. Second, the test contains engaging young adult stimulus material and lines of questioning for senior school students. Third, the test is not designed to be diagnostic in itself, and should be used in conjunction with other observation points, which was compatible with the mixed methodology employed in the overall program evaluation. Finally being a virtual, 'off-the-shelf' product, the instrument was cost-effective (ACER n.d.). The test was also confirmed as being suitable by a Deaf/hard of hearing education academic.

In a later stage of the evaluation, the evaluation commissioners requested that the students were assessed using a standardised measure of literacy. The Woodcock Reading Mastery Test (WRMT) was selected for this purpose, due to its robust psychometric properties, particularly test-retest reliability, as well as its ability to assess a variety of linguistic skills. As outlined by the *Standards for Education and Psychological Testing* (AERA, 1999), test administration should follow the standardised procedure for administration outlined by the test unless a situation arises, such as working with CLD populations, which dictates that an exception should be made, and any exceptions made must be documented (Standard 5.1 and 5.2) (American Educational Research Association [AERA], 1999).

It is important to acknowledge the challenges of using norm-referenced tools such as the WRMT testing in CLD populations. Issues such as content bias, linguistic bias and disproportionate presentation in normative samples need to be addressed when using such tools to inform evaluative judgments (Laing & Kamhi, 2003). For instance, content bias occurs when a test assumes shared experiences, concepts or vocabulary between examinees. As the WRMT was normed using a United States sample, it was examined for possible content bias. While there were a small number of items in the *Reading Comprehension* selection that would be considered unfamiliar for Australian students, overall, the test was determined to be appropriate for use.

In addition to reviewing the test for content bias, the evaluation team consulted with Auslan interpreters to identify potential linguistic bias that may be present for native Auslan signers. As some linguistic issues were identified in the conduct of the test, such as the need to respond to questions with one word, which may be difficult for Auslan signers who may use more than one word to describe the meaning of a concept, it was

necessary to create a protocol to ensure equal fairness of testing for Auslan signers, English speakers and bilingual students.

The evaluation needed to ensure that all examinees understood the test instructions, and were able to respond fairly. This language requirement presented the evaluators with an obstacle, as team members suitably qualified to administer the WRMT were not proficient in Auslan. It was essential to enlist the support of Auslan interpreters, who were provided for all students for whom oral communication was not their primary mode of communication. However, this scenario presented additional complications in relation to the testing environment which, as specified by the standards, should only include the administrator and participant. To minimise the impacts of the presence of the interpreter, all interpreters were made aware of the testing protocols and given clear instruction regarding their involvement.

Finally, standard procedures for the WRMT were amended to take the reading level of a year 6 student as a starting point. This was informed by research indicating the reading levels of Australian Deaf/hard of hearing school leavers are approximately at this level, by amending the starting point, the potential for floor effects was minimised (Walker & Richards, 1992).

Both tests, when amended and administered as described above, were found to be suitably engaging and sensitive in assessing the language and literacy skills of the students in this study. However, there remain obvious limitations in the use of such instruments in the absence of meaningful norms for the population of interest. The identification of appropriate assessment tools for this population remains an important area for the research agenda in Deaf/hard of hearing education.

Conclusions

This paper has described the ways in which an inclusive evaluation methodology, underpinned by ethical and professional standards and a philosophy of transformative evaluation, was utilised to facilitate engagement with a CLD population, namely Deaf/hard of hearing secondary students.

Engaging the Deaf/hard of hearing population effectively in evaluation and research is an important step in improving services and outcomes for this population, who are often marginalised and poorly served in the community and society as a whole (Beattie, 2001). It is through the continued development of new methodologies that are inclusive and underpinned by ethical and professional standards, that members of this population will be enabled and empowered to have a voice in the ongoing development of services, programs and policies of which they are beneficiaries. Through this process, policy makers and service providers will be better informed as to how best to meet the needs of this population (Mertens, 2009).

The evaluation was designed to better inform program developers, academics and educators about the challenges Deaf/hard of hearing students face in education, and

the potential for an assistive technology, namely RTC, to help students navigate and better manage these challenges. In addition to illustrating their challenges, participants also shared what worked well for them and enabled them to engage in education. Overall, the evaluation process sought to ensure that the participants' engagement in the evaluation served to improve the RTC program, accurately capture the impact of program, and generate information about the challenges and enablers in education for Deaf/hard of hearing students in an Australian context. To do so, the evaluators adopted Merten's (2009) view of evaluation, ensuring that the evaluation itself led to better outcomes for vulnerable and marginalised populations which in this case were Deaf/hard of hearing students in selected secondary schools.

As a consequence of this evaluation, the evaluation commissioners have a strong foundation on which to base further research into the use of RTC. They also have evidence from which to develop an improved model for the provision of assistive technology services in education to maximise access to the curriculum for students who are Deaf/hard of hearing.

References

- Australian Council for Educational Research (ACER). (n.d). *Compass Literacy and Numeracy Test*. Retrieved from <http://www.acer.edu.au/tests/compass/compass-overview1>.
- American Evaluation Association. (2004). *Guiding Principles for Evaluators*. Retrieved from <http://www.eval.org/p/cm/ld/fid=51>.
- American Educational Research Association (AERA), American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for Educational Assessment and Psychological Testing*. American Educational Research Association: Washington: AERA.
- Australasian Evaluation Society (AES), (2013). *Guidelines for the Ethical Conduct of Evaluations*. Retrieved from http://www.aes.asn.au/images/stories/files/membership/AES_Guidelines_web.pdf.
- Beattie, R.G. (2001). *Ethics in deaf education: The first six years*. San Diego: Academic Press.
- Brett, M. (2010). Challenges in Managing Disability in Higher Education, Illustrated by Support Strategies for Deaf and Hard of Hearing Students. *The Open Rehabilitation Journal*, 3, 4-8.
- Boyd, J., & Vader, E. (1972). Captioned Television for the Deaf, *American Annals of the Deaf*, 117, 34-37.
- Centres for Disease Control and Prevention. (1999, September 17). Framework for Program Evaluation in Public Health. *Morbidity and Mortality Weekly Report*,

48(RR11), 1-40.

Daelemans, W., Höthker, A. & Sang, E.T.K. (2004). Automatic sentence simplification for subtitling in Dutch and English, Retrieved from <http://www.clips.ua.ac.be/~walter/papers/2004/dhs04.pdf>

Green, J. (2007). *Mixed Methods in Social Inquiry*. San Francisco: John Wiley & Sons.

King, C.M., & Quigley, S.P. (1985), *Reading and Deafness*. San Diego: College-Hill Press.

Laing, S.P., & Kamhi, A. (2003), Alternative Assessment of Language and Literacy in Culturally and Linguistically Diverse Populations, *Language, Speech & Hearing Services in Schools*, 34, (1), 38-41.

Markham, P. (1989). The Effects of Captioned Videotapes on the Listening Comprehension of Beginning, Intermediate and Advanced ESL Students. *Educational Technology*, 29(10), 38-41.

Murphy-Berman, V., & Jorgensen, J. (1980). Evaluation of a Multilevel Approach to Captioning Television for Hearing-Impaired Children. *American Annals of the Deaf*, 125, 1072-1081.

Mertens, D. (2009). *Transformative research and evaluation*. New York: Guilford Press.

National Commission for the Protection of Human Subjects of Biomedical and Behavioural Research. (1979), *The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research*. Retrieved from <http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html>

Sanders, J. (1994). *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs*. The Joint Committee on Standards for Educational Evaluation. Thousand Oaks, CA: SAGE Publications.

Stinson, M., Stinson, R., Henderson, J., & Miller, L. (1988). Perceptions of Hearing-Impaired College Students Toward Real-Time Speech to Print: RTGD and Other Educational Support Services. *The Volta Review*, 90, 336-338.

Walker, L., & Rickards, L. (1992). Reading Comprehension Levels of Profoundly, Prelingually Deaf Students in Victoria. *The Australian Teacher of the Deaf*, 32, 32-47.

Youdelman, K., & Messerly, C. (1996). Computer-Assisted Notetaking for Mainstreamed Hearing-Impaired Students. *The Volta Review*, 98(4), 191