Comparison of a non-spoken response mode and a spoken response mode in a test of phonological awareness

M Casey MA (AAC)
Educational Consultant, Centre for Augmentative and Alternative Communication, University of Pretoria.

KM Tonsing MA (AAC)
Junior lecturer, Centre for Augmentative and Alternative Communication, University of Pretoria.

E Alant PhD
Professor, Centre for Augmentative and Alternative Communication University of Pretoria.

ABSTRACT

Assessing students with moderate to severe disabilities, and especially those with little or no functional speech (LNFS), often necessitates accommodation of assessment material. These accommodations need to be equivalent to the original testing procedure. The aim of this study was to compare two different modes of response, one spoken and one using eye-gaze, to determine the influence of response mode on test results. A cross-over within-subject counterbalance design using multiple subjects was used[1]. Forty-eight typically developing Grade 1 children participated in the study. Participants were individually required to respond to 90 phonological awareness questions using either a spoken "Yes/No" response or a non-spoken "Yes/No" response by using eye-gaze. The response modes were controlled for order effect and analysed using ANOVA. No statistically significant difference was found between the two response modes.

Key words: assessment, eye-gaze transfer board (ETRAN), little or no functional speech (LNFS), phonological awareness, test accommodations

Introduction

Schooling reforms in South Africa aim to create equal learning opportunities for all learners[2] including those with barriers to learning (BtL). The term 'inclusion' is used to refer to a process of recognising and respecting differences, while supporting learners, educators as well as the education system to meet a range of learning needs. The focus is thus not only on learners that are to be made to 'fit in', but on system changes that will allow learners to benefit from least restrictive learning environments. The legislation reflects a move away from the medical model of disability towards a social and systems perspective.

Accommodating learners with severe physical disabilities and little or no functional speech (LNFS) in the classroom poses various challenges, one of these being how to assess such learners in a way that reflects their true function. It has been found that unless timed responses, fine motor responses and/or spoken responses are minimised in the assessment task, the risk of discrimination against this population cohort remains extremely high[3, 4, 5].

The sensitive and appropriate use of test accommodations can be a fitting course of action for increasing the relevant participation of children who had previously been excluded from the large-scale accountability assessments within the schooling system, and also to increase the possibility of valid and reliable test scores being obtained[4, 6].

Test accommodations, according to Thurlow, Elliott and Ysseldyke[4], refer to the changes in test materials or procedures that allow individuals with disabilities to be assessed in a way that ensures that their abilities rather than their disabilities are assessed. Test accommodations, in contrast to test modifications or alternate assessments, are mere changes to testing procedure, while still aiming to test the same content on the same standard. The accommodated test procedure should remain equivalent to the original test procedure and should not influence test results. Test accommodations and modifications are highly complex procedures. Although it is widely accepted and reported that teachers use test accommodations to make assessment material more accessible to the learner with moderate to severe disabilities, the data on the acceptability of making these accommodations remains limited[3, 7]. Phillips[8] emphasises that it is imperative to ensure that the scores obtained with and without accommodations are equivalent and that the integrity of the assessment remains intact. Phillips[8] suggests that when considering whether a proposed test accommodation is valid or invalid, the assessor should carefully consider the purpose of the test, the skills to be evaluated and what inferences the assessor wants to make from the test score.

While much research in respect of children diagnosed with learning disabilities is available, there is not much evidence to support the validity of test accommodations for learners with moderate to severe disabilities[9, 3], and in particular for those learners who have LNFS.

Of the research available, Wagner[5] reported that using a Yes/No response or binary communication has the potential for accurate communication of the inherent knowledge of the child. In his research, Wagner[5] aimed to measure whether assessment on the standardised form of the Peabody Picture Vocabulary Test–Revised (PPVT-R)[10] would render comparable scores when the test was administered in the binary communication mode. His results led him to the conclusion that the binary format was justified as a practical alternative to the standard PPVT–R administration.

Arvidson [7] conducted a study to establish whether the performance of a test of multiple-choice questions would render the same results if children used a scanning response mode as opposed to direct selection as a response mode on the Science and Technology Test from the 1999 Massachusetts Comprehensive Assessment System (MCAS). The findings indicated that neither unfair advantages nor unfair disadvantages occurred as a result of using the scanning mode of response as opposed to direct selection as a response mode.

Another accommodation that can be made to the testing procedure is to replace a spoken yes/no answer with a yes/no response indicated through eye-gaze. Eye-gaze communication is a skill that develops from early infanthood. More sophisticated communication means replacing eye-gaze as the typical infant matures. However, for children who fall along the disability continuum, gesture or natural speech may not become a viable option for effective communication. Therefore, eye-gaze frequently remains a viable avenue for communication exchange[11].

The aim of this study was thus to verify the equivalence between a yes/no response indicated by eye-gaze versus a yes/no response indicated by speech, in order to prove the former a valid test accommodation. A cross-over design between two groups was chosen, by which both

conditions for responding were administered to both groups. For this reason, a population of typically developing children was chosen, as both groups needed to respond in both ways (verbal and non-verbal).

Phonological awareness skills were chosen to form the test content. Deficits in these skills have been shown to have a negative effect on literacy acquisition, and certain aspects of phonological awareness have a causal relationship in the acquisition of mature reading skills[12]. The evaluation of these skills is paramount as children with LNFS are at a particular risk for deficient phonological awareness skills[13], a possible reason being that they cannot experiment with speech expressively as typically developing children do.

Method

Study Aim

The primary aim of this study was to investigate whether a non-spoken response mode (eye-gaze) would render equivalent results to a spoken response mode in a test of phonological awareness in typically developing English first language and English second language, Grade 1 children.

Design

A cross-over within-subject counterbalance design using multiple subjects was used[1], with participants comprising typically developing Grade 1, English first language and English second language children. The cross-over design enabled a within group comparison[14]. The non-spoken mode of response required the children to answer the questions using eye-gaze only (by looking at the top left hand corner of the eye-gaze transfer board (ETRAN) to indicate "Yes", or by looking at the bottom right hand corner of the ETRAN to indicate "No"). An ETRAN is a rectangular transparent sheet of Plexiglas with a central square cut out of the middle of the sheet. The answer for "Yes" indicated by a tick as well as the written word was placed on a green cardboard square and was stuck on the top left hand corner of the ETRAN, while the answer depicting "No" indicated by a cross and the written word was placed on red cardboard and was stuck on the bottom right hand corner. The spoken mode simply involved the children responding with a verbal "Yes" or "No". Participants were assigned to Groups 1 and 2. These two groups were matched for age, reading level and English first versus English second language. Within each group the test administration was furthermore counterbalanced for the following: a) item order within a task; b) task order across the three tasks; and c) presentation of spoken and non-spoken items. *Table 1* provides a layout of the design as for task order 1, 2, 3 (Task Order A). The same procedure was followed for Task Order B (3, 1, 2) and C (2, 3, 1). Group 1's participants were always required to answer the first five items in the spoken mode for the task administered first (be it Task 1, 2 or 3) regardless of item order. The tasks administered second and third required them to answer the first five items in the non-spoken mode, regardless of item order.

Participants

Because the population of children with LNFS is heterogeneous in nature [15, 16], a non-disabled population was chosen for greater homogeneity of the subjects. This would then also assist the researcher in comparing the two response modes. The participants were required to be enrolled in the regular curriculum of the Grade 1 year. The language of instruction had to be English. Receptive and expressive language had to be at a Grade 1 level. They were required to be six or seven years old at the time of assessment. They had to have normal hearing, vision and speech. Children diagnosed with ADHD had to be effectively medicated. Parents had to give consent for the child to participate in the study. The four Grade 1 classes of a privately funded school for typically developing children were selected as a subject pool. Eighty children complied with the selection criteria.

Forty-eight subjects from four different mixed abilities c1asses were asked to participate individually in the study. From the 80 children who complied with the selection criteria, 24 were randomly assigned to Group 1. The teacher grouped these 24 children into three groups according to reading ability: above average, average and below average. Twenty-four children (Group 2) were matched according to reading ability and English first language or English second language.

The mean age of Group 1 was 85.17 months (SD = 3.62) while the mean age of Group 2 was 84.17 months (SD =4.21). The median age for Group 1 was 85 months and the median age for Group 2 was 84 months. The Mann-Whitney Test indicated a p-value of 0.3150, thereby indicating no statistically significant difference between the groups at the 5% level of significance.

Materials

A Phonological Assessment Battery, that required only "Yes" or "No" responses, was developed for the purpose of this study. The subtests were drawn from the Phonological Awareness Literacy Screening (PALS) [17] and the Test of Phonological Awareness (TOP A) [18]. It included the following sub-sections:

- Task 1 - Rhyme recognition taken from the Phonological Awareness Literacy Screening[17]
- Task 2 - Initial sounds same taken from the Test of p~ Awareness[18]
- Task 3 - Ending sounds same taken from the Test of Phonological Awareness[18]

The assessment had a total of 90 items divided into three tasks.

In developing the assessment, the researcher endeavoured to adhere to the suggestions made by Blischak [19], which is to provide tasks which require the least adaptations when assessing persons who have LNFS. The words in each of the tasks were chosen to decrease the phonological similarity effect, as described by Conrad [20], as he found that when items presented together were phonologically similar to each other, the ability for accurate recall diminished. Each test item comprised three yes-no

Task order A: 1, 2, 3			
		Group 1	**Group 2**
Task 1	Item order 1-5, 6-10	Spoken response (items 1-5) Non-spoken response (items 6-10)	Non-spoken response (items 1-5) Spoken response (items 6-10)
	Item order 6-10, 1-5	Spoken response (items 6-10) Non-spoken response (items 1-5)	Non-spoken response (items 6-10) Spoken response (items 1-5)
Task 2	Item order 1-5, 6-10	Non-spoken response (items 1-5) Spoken response (items 6-10)	Spoken response (items 1-5) Non-spoken response (items 6-10)
	Item order 6-10, 1-5	Non-spoken response (items 6-10) Spoken response (items 1-5)	Spoken response (items 6-10) Non-spoken response (items 1-5)
Task 3	Item order 1-5, 6-10	Non-spoken response (items 1-5) Spoken response (items 6-10)	Spoken response (items 1-5) Non-spoken response (items 6-10)
	Item order 6-10, 1-5	Non-spoken response (items 6-10) Spoken response (items 1-5)	Spoken response (items 6-10) Non-spoken response (items 1-5)

Table 1: Visual Representation of the Research Design for the Task Order A (1, 2, 3)

	p-values					
Variable	**Task 1**		**Task 2**		**Task 3**	
	Item 1-5	**Items 6-10**	**Items 1-5**	**Items 6-10**	**Items 1-5**	**Items 6-10**
Task order	0.1117	0.6963	0.1676	0.1050	0.4110	0.3433
Response mode	0.9949	0.5073	0.1958	0.2022	0.2947	0.0756
Item order	0.3035	0.7239	0.2933	0.7965	0.7566	0.1874
Interaction between task order and mode	0.4311	0.3436	0.0774	0.5133	0.9422	0.3229
Interaction between task order and order of items	0.8495	0.2447	0.2224	0.5602	0.7757	0.3150
Interaction between mode and order of items	0.6573	0.8153	0.5780	0.5201	0.9399	0.1225
Interaction between task order, mode and order of items	0.3555	0.9524	0.2607	0.8933	0.7943	0.8312

Table II: ANOVA Results for Task Order, Item Order and Response Mode

Task	**Spoken**				**Non-spoken**				**p-value**
	Group 1		**Group 2**		**Group 1**		**Group 2**		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Task 1 Items 1-5	14.13	1.46	13.81	1.17	14.31	1.20	13.63	2.00	0.9949
Task 1 Items 6-10	14.44	1.31	13.38	2.13	14.38	1.19	14.19	1.17	0.5073
Task 2 Items 1-5	14.60	0.83	13.78	1.56	14.78	0.44	14.53	1.13	0.1985
Task 2 Items 6-10	14.22	1.30	14.60	0.91	14.80	0.41	14.78	0.67	0.2022
Task 3 Items 1-5	12.50	1.83	12.25	2.96	13.25	1.16	13.25	1.98	0.2949
Task 3 Items 6-10	12.63	1.41	12.38	2.06	13.50	1.59	13.00	2.14	0.0756

Table III: Performance in Spoken and Non-spoken Modes for Group 1 and Group 2, and P-value Comparing Spoken vs. Non-spoken Modes

questions; one correct answer ('yes') and two foils ('no'). Please refer to Casey[21] for the full assessment battery which was developed.

Test procedure and data analysis

The three tasks were individually administered to each participant. Each task was expected to take five to six minutes to complete, making the total time for each test an average of 15 to 20 minutes. This was confirmed during the administration of the pilot study. Identical instructions were given to each child. For all tasks, three practice trials were provided. Here corrective feedback was given to the child. The practice trials were repeated once if necessary. Thereafter the test proceeded. None of the questions was repeated once the practice trials were completed.

Each yes/no question was scored as either correct (one point allocated) or incorrect (zero points allocated). A participant's score for each of the three phonological awareness tasks was the total number of correct responses, with the maximum score for each test being 30. In total, each subject was asked 90 questions. As all tests provided the participant with three options, the correct option was placed randomly and occurred with approximately equal frequency in all positions.

It took five days to complete the study. A speech-language therapist assisted with the testing procedure. For the first two days both the researcher and the assistant scored each participant's responses. A comparison of the responses was undertaken to ensure that equivalence of scoring was achieved. Inter-rater reliability was 100%. Thereafter the researcher and the assistant administered the tests on their own.

The Mann-Whitney Test was used to obtain a *p*-value of the means obtained in comparison of the ages of Group 1 and Group 2. The one-way analysis of variance (ANOVA) was used to determine any statistically significant differences between the population means. This study required that two or more sample means were compared on one independent variable. In using the ANOVA procedure, the researcher was able to test the different variables both individually and in combination between both groups, thereby making more accurate probability statements than would be possible if using a series of separate *t*-tests[22].

Results and Discussion

Association between Task Order, Item Order, Response Mode and Scores

Table II gives the ANOVA results to show possible association between task order, item order, response mode and the results, as well as the interaction of these variables with each other. It is clear from the table that none of the p-values is lower than 0.05, indicating that there is no significant association between any of the variables and the scores. Thus item order and task order were effectively controlled for and the scores of the participants did not differ significantly between the response modes.

Spoken versus Non-Spoken Response Modes

Figure 1 depicts the average number of correct responses in each mode for the whole test battery. The Mean, Standard Deviation and p-values for the spoken and non-spoken modes are given in *Table III*.

From *Figure 1* as well as *Table III* it is clear that, as expected, there are no statistically significant differences between spoken and non-spoken

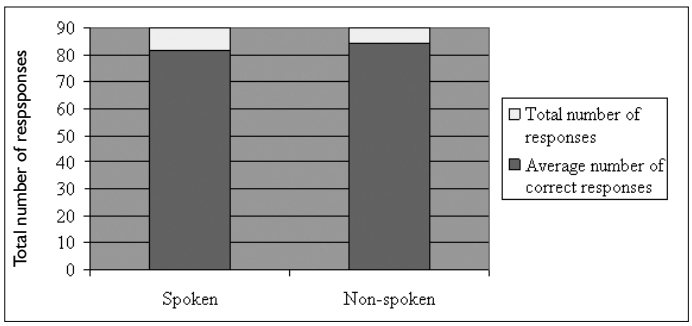

Figure 1:Average number of correct responses for the spoken versus non-spoken mode across all three tasks

modes on any of the tasks at the 5% level of significance.

The fact that no statistically significant difference was found between the two response modes indicates that the cognitive and physical demands of the response mode did not appear to impact either positively or negatively on the performance of these school children - they performed comparably under both conditions. This finding might reflect that eye-gaze was in fact a natural response mode for these children [23] as it precedes the verbal response mode in typically developing children. It is therefore comparable in terms of ease of production in relation to the spoken mode of response. Across responses, the non-spoken mode rendered slightly higher means for number of correct responses. The novelty aspect of the response mode, which encouraged greater attention to detail, as well as sustained levels of attention and concentration, might explain this phenomenon.

Conclusion and Recommendations for Further Studies

This is an initial study to investigate test accommodations which could possibly be used for children who have LNFS. The need for evidence-based practice to determine the equivalence or non-equivalence of assessment accommodations is necessary as the barriers to assessment are being addressed[4]. The findings suggest that, in typically developing children, those who receive accommodations in their assessment will be neither advantaged nor disadvantaged in their response via eye-gaze.

This is one of few studies that investigate accommodations of the response mode in the assessment of children. Other than the Arvidson[7] and Wagner[5] studies, no other studies have investigated the validity of accommodations in response modes. As with the Wagner [5] and Arvidson[7] studies, this study found that accommodations in the response mode neither advantaged nor disadvantaged the cohorts participating in the study.

The findings are, however, limited to a group of typically developing six and seven year old children. Studies done on children younger or older may not produce the same results. Should the same age children be assessed on more complex aspects of phonological awareness, those results may also differ from the results of this study.

In this study, a binary response mode was used. However, binary communication is limited, as there is a 50/50 chance of providing the tester with the correct response. Using binary communication to test phonological awareness results in the test only assessing the skill at a receptive level and not at an expressive level. By not being able to generate words or pseudo words, the whole range of phonological awareness (rhyme oddity, rhyme production, oral production of phonemes, and articulating individual segments in words, phoneme deletion and synthesis) cannot be assessed[19]. Further recommendations for research would therefore include:

- A replication of this study should target participants with severe speech disorders and severe physical disabilities who might benefit from eye-gaze as a response mode in the testing context. This study could then be extended to compare the efficiency of the use of eye-gaze with other direct or indirect selection strategies that would not require the presence of another individual in the testing situation.
- Studies should investigate the expanded use of an eye-gaze response mode to test responses that go beyond binary responses as used in this study.

References

1. DePoy E, Gitlin LN. Introduction to research. Multiple strategies for health and human services. St. Louis: Mosby Year Book Inc, 1994.
2. White Paper 6: Building an Inclusive Education and Training System. National Department of Education: Cape Town, South Africa, 2001.
3. Lloyd LL, Arvidson H, Bennett DE, McLaughlin TR. Accommodations taxonomy for improved assessment of AAC users. Unpublished conference proceedings, ISAAC Mini-seminar, Odense, Denmark, 2002, August.
4. Thurlow ML, Elliott JL, Ysseldyke JE. Testing students with disabilities. Practical strategies for complying with district and state requirements. 2nd ed. Thousand Oaks: Corwin Press, 2003.
5. Wagner PA. Adaptations for administering the Peabody to individuals with severe communication and motor dysfunctions. Mental Retardation 1994; 32:107-112.
6. Elliott SN, McKivitt BC, Kettler R J. Testing accommodations research and decision-making: the case of "good" scores being highly valued but difficult to achieve for all students. Measurement and Evaluation in Counselling and Development 2002; 35:153-166.
7. Arvidson H. Comparative study of two modes of response: direct selection and scanning. Unpublished doctoral dissertation. Purdue University: West Lafayette, IN, 2000.
8. Phillips SE. High-stakes testing accommodations: validity versus disability rights. Applied Measurements in Education 1994; 7:93-120.
9. Grise P, Beatie S, Algozzine B. Assessment of minimum competency tests in Fourth Grade learning disabled students: test modifications make a difference. Journal of Educational Research 1982; 76:35-40.
10. Dunn LM, Dunn LM. Peabody Picture Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service, 1981.
11. Goossens' CA, Crain SS. Overview of non-electronic eye-gaze communication techniques. Augmentative and Alternative Communication 1987; 3:77-89.
12. Wagner RK, Torgesen JK. The nature of phonological processes and its role in the acquisition of reading skills. Psychological Bulletin 1987; 101:192-212.
13. Vandervelden MC, Siegel LS. Phonological processing and literacy in AAC students with motor speech impairments. Augmentative and Alternative Communication 1999; 15:191-211.
14. Jones B, Kenward MG. Design and analysis of cross-over trials. London: Chapman and Hall, 1989.
15. Higginbotham DJ, Bedrosian JL. Subject selection in AAC research: Decision points. Augmentative and Alternative Communication 1995; 11:11-13.
16. Light J. Interaction involving individuals using augmentative and alternative communication systems: state of the art and future directions. Augmentative and Alternative Communication 1988; 4:66-82.
17. Invernizzi M, Meier JD, Swank L, Juel C. Phonological Awareness Literacy Screening (PALS). Charlottesville, VA: University of Virginia Press, 1999-2000.
18. Torgesen JK, Bryant BR. Test of phonological awareness. Austin, TX: Pro-Ed, 1994.
19. Blischak DM. Phonological awareness: implications for individuals with little or no functional speech. Augmentative and Alternative Communication 1994; 10:245-254.
20. Conrad R. Acoustic confusion in immediate memory. British Journal of Psychology 1964; 55:75-84.
21. Casey M. A comparison of a non-spoken response mode and a spoken response mode in a test of phonological awareness. Unpublished master's thesis. University of Pretoria: South Africa, 2004.
22. McMillan JH, Schumacher S. Research in education. A conceptual introduction. 5th ed. New York: Longman, 2001.
23. Goossens' CA. Aided communication intervention before assessment: A case study of a child with cerebral palsy. Augmentative and Alternative Communication 1989; 5:14-26.

❐

Maureen Casey: makcee@mobileemail.vodafonesa.co.za

Kerstin Tonsing: kerstin.tonsing@up.ac.za

Erna Alant: erna.alant@up.ac.za