



Accessible digital assessments of temporal, spatial, or movement concepts for profoundly motor impaired and non-verbal individuals: a pilot study

Mark Moseley, Lindsey Howat, Leigh McLoughlin, Sarah Gilling & Diane Lewis

To cite this article: Mark Moseley, Lindsey Howat, Leigh McLoughlin, Sarah Gilling & Diane Lewis (2021) Accessible digital assessments of temporal, spatial, or movement concepts for profoundly motor impaired and non-verbal individuals: a pilot study, *Disability and Rehabilitation: Assistive Technology*, 16:3, 350-360, DOI: [10.1080/17483107.2019.1683240](https://doi.org/10.1080/17483107.2019.1683240)

To link to this article: <https://doi.org/10.1080/17483107.2019.1683240>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 15 Nov 2019.



Submit your article to this journal [↗](#)



Article views: 1170



View related articles [↗](#)



View Crossmark data [↗](#)

Accessible digital assessments of temporal, spatial, or movement concepts for profoundly motor impaired and non-verbal individuals: a pilot study

Mark Moseley^{a,b} , Lindsey Howat^c, Leigh McLoughlin^d , Sarah Gilling^b and Diane Lewis^b

^aFaculty of Media and Communication, Bournemouth University, Poole, UK; ^bLivability Victoria Education Centre, Poole, UK; ^cDorset Healthcare University NHS Foundation Trust, Adult Community Speech and Language Therapy Service, Blandford Forum, UK; ^dFaculty of Science and Technology, Bournemouth University, Poole, UK

ABSTRACT

Purpose: Here we present a study of two new Assistive Technology (AT) accessible digital assessments which were developed to address the current paucity of (English) spoken language comprehension assessments accessible to individuals who are both non-verbal and have profound motor impairments. Such individuals may rely heavily upon AT for communication and control. However, many assessments require that responses are given either verbally, by physical pointing or manipulating physical objects. A further problem with many assessments is their reliance upon static images to represent language components involving temporal, spatial or movement concepts. These new assessments aim to address some of these issues.

Materials and methods: The assessments were used with 2 young people who are non-verbal and have profound motor impairments (GMFCS level IV/V) and who use eye gaze as their primary method of communication and access. One assessment uses static images and the other short video clips to represent concepts containing temporal, spatial or movement elements. The assessments were carried out with each participant, both before and after an intervention, as part of a larger study.

Results: The assessments were accessible using AT (eye gaze) for both participants, although assessment scores varied. The design of the assessments particularly suited one participant who scored near maximum, but they appeared less suitable for the other participant.

Conclusions: Making assessments AT accessible removes a barrier to assessing aspects of the spoken language comprehension abilities of some. Video may be a better medium for representing certain concepts within assessments compared with static images.

ARTICLE HISTORY

Received 5 March 2019
Accepted 17 October 2019

KEYWORDS

Spoken language comprehension; eye gaze; cerebral palsy; digital assessment; Assistive Technology; non-verbal; disability

► IMPLICATIONS FOR REHABILITATION

- The new assessments provided a deeper understanding of two members of a group who are traditionally difficult to assess, using two alternative physically accessible methods of assessing the spoken language comprehension of the target group;
- Accessible assessments are important for assessing complex individuals in order to identify knowledge limitations and set therapy (and education) goals;
- The alternative access features of communication software can provide a “wrapper” for providing accessibility features to assessments;
- Video clips may be a better means of representing certain concepts in assessments compared to their static equivalents;
- Ensuring that assessments are physically accessible is sufficient for the assessment of some individuals, but for some “cognitive” accessibility also needs to be considered.

Introduction

There are currently few suitable methods for assessing the comprehension of spoken language of those who have both anarthria (inability to produce clear, articulate speech) and profound motor impairments [1]. Hereafter these individuals will be referred to as the Target Group or TG.

It is important to assess children and young people (CYP) who have disabilities to develop an understanding of their existing knowledge and any intellectual impairments that they may have [2]. This understanding can provide a useful baseline, inform the direction and focus of therapy and education [3] and assist with

measuring progression. It can also help with identifying a person's suitability for Alternative and Augmentative Communication (AAC) [4]. Unsuitable or unsound assessment techniques may lead to inaccurate results and a misrepresentation of an individual's abilities which may lead to unrealistic [1] or reduced expectations [5].

At present, Speech and Language Therapists (SaLTs) may use a variety of published and standardized assessment batteries to assess the spoken language comprehension of the TG [6,7]. Nearly all such assessments are standardized using typically developing children [2]. This approach may mean that content and methods

of completion are not appropriate for the TG given their more limited life experiences and motor impairments.

The TG may be heavily reliant on Assistive Technology (AT) for communication and control but many assessments rigidly require that answers are given verbally, by physical pointing or even through the manipulation of physical objects.

Some assessment schedules contain a range of permitted adaptations, but these are usually minimal and do not accommodate the needs of those who are non-verbal and more motorically impaired [1]. To make the assessment materials and administration process suitable for use with the TG, modifications may be required which can then break the standardization and lead to invalidated results. Indeed, this may lead SaLTs to abandon conventional assessments completely and instead assess informally using observation or assessment schedules that they have developed themselves [7].

Adaptations to assessments may also alter the nature of what is being assessed, introduce assessor bias and increase cognitive loading [8]. Whilst potentially useful as initial screening tools, observation and bespoke assessment approaches lack standardization and so will have no evidence base to support their efficacy [7].

When assessing those who provide answers using eye-pointing, there is a risk of confusing “look to view” or “look to explore” with “interactive intention” [9]. This can lead to misinterpreted answers, so a more automated approach would be beneficial.

By their physical, often paper-based nature, many standard assessments are restricted to using static two-dimensional images to represent verbs. Symbolic or pictorial representations of certain verb concepts can be difficult to interpret, especially those involving more abstract elements. The artistic conventions used to represent movement in images e.g., “curved lines around joints” may not be understood by some [10]. Some verbs, such as “sleeping,” do not involve movement and so avoid this problem; others, such as “releasing” or “moving forwards,” may be better represented using moving images in the form of animations, for example Mayer Johnson’s PCS symbol animations [11], or video clips [10,12]. This is a key issue that this work is designed to address.

In this article, we describe two alternative assessment approaches which are designed for access by the TG to overcome the limitations of existing solutions. One of these is based on static images, the other on short video clips. These are accessible by eye gaze and designed to complement each other by representing concepts in both static and animated or moving formats to suit the concept.

A pilot study is also presented, which was used to evaluate the effectiveness of the new assessment methods. In this pilot, the assessment approaches were employed to assess two participants’ understanding of temporal, spatial and movement concepts both before and after an intervention (baseline and outcomes, respectively), described later.

Existing (digital) approaches for assessing the TG

The literature reveals a variety of alternative approaches to the assessment of groups who are difficult to assess using conventional methods. The approaches of interest in the current study are those which involve the use of digital technology.

Recently there has been a move by commercial assessment providers to digitize their standardized assessments, for example Pearson’s Q-Interactive [13]. However, most of these are literal translations of the physical versions and do not provide any

additional accessibility options, still requiring that answers be given by pointing, touch or verbal responses. Often designed for use with mobile technology such as tablets, their main technical focus is typically on automatic capture and analysis of the results. While the mobile platforms themselves may provide additional accessibility options, the assessment administration procedure may not permit their use.

Researchers have attempted to make standard forced-choice quadrant assessments accessible to a wider range of individuals. Friend and Keplinger [14] created an assessment based on touchscreen technology and standardized content for use with young infants. Warschausky et al. [15] converted and adapted the materials of several existing standardized assessments to a digital format, importing them into the communication software Boardmaker Speaking Dynamically Pro [16]. This provided a range of AT accessibility options, including support for linear switch scanning and a head mouse, which were used in the study. Adapting the assessments did not appear to affect the results of some assessments significantly when compared to the standard versions. This approach increases the range of accessibility options, but incurs the additional financial cost of the alternative access software (in this case Boardmaker Speaking Dynamically Pro). Also, the researchers noted that their approach raised legal issues concerning the copyright of the standard assessment materials.

Brain-Computer Interfaces (BCI) have been used as a method of identifying the understanding of spoken language in difficult to assess groups. Byrne et al. [17] presented participants with images and a matching or non-matching spoken word whilst measuring their brain activity. It appeared that the use of a BCI could be effective at detecting when a participant recognized a match or was conflicted by the image and a nonmatching word. This approach identified whether the participant understood the relationship between only one picture and one word. Choosing from multiple choices is more cognitively challenging. Huggins et al. [18] also used a BCI as a means of eliciting answers to a digitally adapted version of the Peabody Picture Vocabulary Test – 4th Edition (PPVT-IV) [19]. They compared the results of the unmodified and BCI-adapted versions of this test and found the results to be “within the expected variation of repeated test administration,” but stated that the adapted version took approximately three times longer to complete.

The use of BCIs has the benefits of requiring no motor or verbal responses from test subjects. However, BCIs may be unsuitable for some, including those who cannot tolerate wearing equipment on their heads, or those who have uncontrolled or involuntary movement. There is usually quite a significant attachment and detachment period too, which may test the patience of some. The use of BCI may, therefore, not be feasible in a clinical practice setting.

There are few assessments which are appropriate for use with eye tracking technology. Ahonniska-Assa et al. [20] used a digitally adapted form of PPVT-IV for assessing the receptive language of individuals who had Rett syndrome. The participants used “eye-tracking” technology and gave their answers by focusing on one of four forced-choice answer cells. This appeared to be a suitable access method for some, indicating greater proficiency than had been anticipated.

“CARLA” (Computer based Accessible Receptive Language Assessment) [21] is a commercially available assessment which works within the communication software MindExpress [22]. MindExpress supports a range of access methods which are “inherited” by CARLA. These methods include touch, switch scanning and eye gaze. It is not clear whether this assessment was

based on research or clinical experience and the assessment does not appear to have been standardized.

Geytenbeek et al. [6] created the Computer-Based Instrument for Low motor Language Testing (C-BiLLT), a tool for assessing the spoken language comprehension of groups who are difficult to assess, such as those who have severe cerebral palsy. This assessment provides a variety of access methods including eye gaze. The “sequencing of the linguistic complexity of items on the test was based on the Dutch version of the Comprehension Scale of the Reynell Developmental Language Scales (RDLS).” RDLS is a standardized assessment [23]. C-BiLLT is currently undergoing standardization trials and is being translated into other languages. At the time of the present study, there was no English language version available (the language of the participants in the present study).

Few studies have examined the use of video in the assessment of those who are unable to answer verbally (or motorically). Preferential looking is one approach that has been used to identify a person’s receptive language comprehension using video. Golinkoff et al.’s [10] Intermodal Preferential Looking Paradigm (IPLP) presented two different videos which were played simultaneously and accompanied by an auditory stimulus. Examinees’ gaze fixation was observed to identify which of the two videos were fixated upon most by participants and whether this preference matched with the auditory stimulus. The aim of the study was to identify whether the receptive language understanding of young preverbal children exceeded their expressive language, which was found to be the case.

Snyder et al. [12] used video in a stimulus preference assessment as an alternative to tangible objects or pictures. They considered video to be more suitable than static images for representing social interactions and activities. Assessment flexibility was key to their assessment approach, which enabled a broader range of individuals to be assessed.

Golinkoff et al. [24] reviewed the applications of their Intermodal Preferential Looking Paradigm (IPLP) spanning a period of 25 years. Their assessment paradigm typically presents only two answer cells. The authors described an inherent limitation with this approach as the “A not A” problem, i.e., the examinee does not know the answer to the question but knows the concept depicted in the incorrect cell, and that this does not match with the answer, and so using a process of elimination is able to deduce the correct answer.

Aims of the current research

The overall goal of this work was to provide assessment methods which are accessible by the TG. In particular, the focus was on the use of eye gaze to provide accessibility and to support the assessment of a set of temporal, spatial or movement concepts in an appropriate manner. To achieve this overall goal, two complementary assessment methods were developed and their effectiveness was evaluated in a pilot study.

The aims of the assessment techniques were:

Aim 1: To ensure that the new assessments were AT accessible for the TG, who usually use eye gaze as an alternative access method;

Aim 2: To represent concepts containing temporal, spatial or movement elements in a more suitable format;

Aim 3: To minimize the risk of assessor misinterpretation and bias.

The aim of the pilot study was to provide an indication of whether the participants’ knowledge of the concepts under investigation could be evaluated by the assessments.

Within the pilot study, the concepts that were measured concerned the participants’ knowledge of specific temporal, spatial or movement concepts, predominantly prepositions, verbs, adverbs and a small number of adjectives (colors). The necessity for the assessments described here resulted from a larger study, which concerned the effect of an eye gaze controlled robotic intervention on the comprehension of such concepts. Here, the design was for the assessments to be conducted both before and after the intervention, intended to improve the participants’ understanding of these concepts using an eye gaze controlled robotics system. The details of this intervention and its effectiveness are beyond the scope of this paper; the focus here is on the evaluation of the assessment methods.

Materials and methods

Two assessments were created: a static image-based assessment and a video-based assessment. These are complementary approaches, where the assessments collectively represented temporal, spatial, or movement concepts in the format most appropriate to the particular concept.

Both assessments are digital and computer-based, and designed to be run on the Microsoft Windows 7/10 Operating System. The static image-based assessment runs within the Grid 3 software [25].

The new assessments share some of the common features of existing standardized assessments. For example, both assessments were of the quadrant forced-choice variety, i.e., a 2×2 arrangement of cells with only one correct answer.

One of the main ways in which the new assessments differ from standardized assessments is in how answer cells are accessed i.e., using eye gaze technology. No written word labels were presented for the answer cells, as spoken language comprehension was the focus of investigation.

Both assessments were designed, implemented and tested by a team comprised of an Assistive Technologist (the first author) and five SaLTs. A subset of the SaLTs trialed and practiced the assessments in pairs – one adopting the role of SaLT and the other the “pupil” participant. This helped to refine the design of the assessments and to verify the administration process.

The static image-based assessment consisted of three parts: 1) an access check; 2) practice/familiarization; 3) the main assessment. The images used within the practice section of the static image-based assessment were selected from the libraries included with the Grid 3 software. The remainder of the images and all of the video clips were created by the first author. Many of the images and all the video clips feature a toy dog character. This character was deemed by the development team to be age appropriate for a wide range of users. The answer cell image designs were kept simple, often featuring a plain white or simple background and limited color palette, helping to establish clear figure-ground.

Ethical approval and consent

Ethical approval was granted by the Science, Technology and Health Research Ethics Panel of Bournemouth University. Informed consent was obtained from both the parents of the participants and from the participants themselves (using a specially adapted symbolized format suited to their communication needs).

The participants

The participants were recruited from students aged between 4 and 19 years attending Livability Victoria Education Centre at the time of the larger study.

Inclusion criteria

Candidates were eligible for inclusion if they were: at levels IV/V on the Gross Motor Function Classification System (GMFCS) [26]; level 5 on the Manual Ability Classification System (MACS) [27]; anarthric, with a clear discrepancy between their level of understanding and ability to speak, as assessed by SaLTs, with a minimum educational attainment of P Scale level 6 in Mathematics, English and Science [28].

Exclusion criteria

Candidates were not eligible for inclusion if they had: any known visual or perceptual impairments which would hinder the use of eye gaze technology; and any hearing impairments which would prevent them from hearing the assessment questions.

Two pupils elected to take part in the study: both were male; P1 had a diagnosis of athetoid cerebral palsy; P2 had a diagnosis of post-viral cerebral palsy; P1 was aged 16 years 7 months, and P2 was 19 years and 6 months at the time of the baseline assessments; neither had any verbal expressive language, but both were able to vocalize and had clear “Yes/No” responses; both were very experienced in using eye pointing for symbol-based communication (both low-tech using a communication book and a communication partner, and high-tech using eye gaze technology and dwell-select).



Figure 1. The hardware running the video-based assessment.

Equipment and set-up

The assessments were administered using a personal computer with a 22 inch touchscreen monitor, mounted on a height-adjustable mobile floor stand with an eye gaze camera attached to the lower frame (Figure 1). The PC ran Microsoft Windows 7. Sound was provided by stereo speakers.

All of the assessments were carried out at Livability Victoria Education Centre in a quiet, distraction-free room which was familiar to both participants. Only the participant, SaLT and first author were present in the room during each assessment. The SaLT would stand to the left of the participant, who would face the monitor. The static image-based assessment was always administered first and the video-based assessment second. All sessions were video recorded for the purposes of verification and analysis of the results.

The static image-based assessment

This assessment was built to work within the Grid 3 software [25]. An assessment question would be presented to the participant by displaying different images in four grid cells, only one of which would represent a correct answer. The typical format of the assessment screens or “panels” is shown in Figure 2. The actual question would then be “read out” using the speech synthesis feature of Grid 3. The questions were read out twice for each question and the examinee could select a cell to hear the question again if necessary. The concepts represented are given in Appendix Table A2.

The assessment is a “grid set” within Grid 3 and inherits all of the accessibility features of Grid 3 including, crucially, support for eye gaze and synthesized speech output. The examinee provided answers by “dwell-selecting” i.e., fixating their gaze upon a single cell for a brief time. The assessment would then automatically log their answer and move on to the next question. Automatic answer logging was achieved by linking a vbScript code file to each of the assessment cells using Grid 3’s Computer Control “start program” function. The logged answers were stored in a spreadsheet format file (see Appendix Table A1 for a sample).

The video-based assessment

At the time of the present study, no communication software was identified which provided support for four eye gaze accessible video cells on a single screen. The closest match to this behavior was found in the “video wall × 4” activity of the Look to Learn software [29]. However, this only provides a single page of videos and two pages were needed for this study. While the Look to Learn page can be edited and the videos changed, doing so during the assessment would interrupt the flow. For these reasons, the video-based assessment discussed here was created by the first author using the C# programming language [30]. This assessment enables eye gaze interaction.

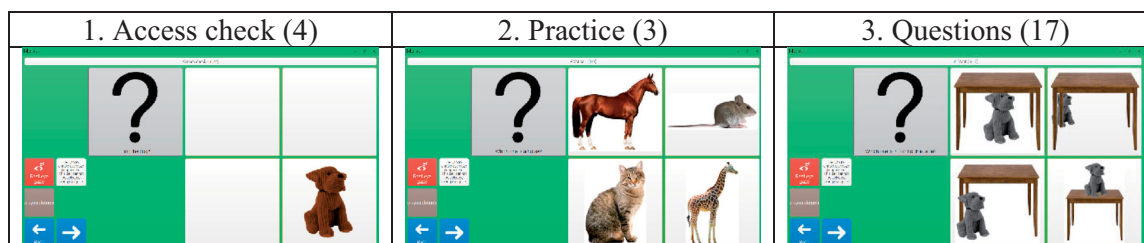


Figure 2. Static image-based assessment: Screen examples – the number in parentheses indicates the number of questions in each section.

The assessment contains just two onscreen “pages,” each with a grid of four video answer cells (Figure 3). Each video clip is between 2 and 4 s in duration, with no audio.

The concepts represented within the video cells are: Page 1: 1. Moving up; 2. Moving backwards; 3. Moving forwards; 4. Moving down; Page 2: 1. Moving left; 2. Gripping; 3. Releasing; 4. Moving right. The majority of the video clips are animations constructed by the first author. The “gripping” and “releasing” video clips are both live recordings created by the first author.

Once eye gaze control has been activated by the SaLT, each video cell animates when the examinee’s gaze falls within its boundaries, but pauses if gaze moves outside of the cell, resuming when gaze focus returns. Upon completion, the video clips pause (showing a black frame), “rewind” and then play from the beginning. Figure 4 shows the first and final frames of the “moving forwards” video clip.

Differences between the video and static image-based assessments

Key differences between the assessment design and administration procedures are summarized in Table 1. The main reasons for these were because Grid 3 was used for the static image assessment, but at the time of the study it did not support the use of four video cells and so could not be used for the video assessment. This mainly led to differences in how the questions were

“read out” (Grid 3 contains speech synthesis), how answers were given and how the responses were recorded or “logged.”

The pilot study

The pilot study consisted of a procedure for administering the assessments with the core aim of indicating whether the participants’ knowledge of the concepts under investigation could be evaluated by the assessments.

The SaLTs administered the assessments by following a predetermined assessment procedure. The participant’s SaLT first explained and then administered the assessment to the participant, helping them to work through and regulate the pace of completion, thereby reducing the likelihood of accidental selections [31]. This regulation was achieved in the static image-based assessment by providing the SaLT with touch-only activated cells that toggled whether eye gaze control was activated or deactivated and, therefore, whether selection of cells was possible. For the video assessment regulation was achieved by the SaLT who determined whether videos would play when focused upon and interpreted the participants’ responses.

The main procedure consisted of four stages, the first three of which involved the static image-based assessments (Figure 2) and the fourth, the video-based assessment:

1. Static Image Access Check: The access check ensures the participant can correctly activate all four of the answer

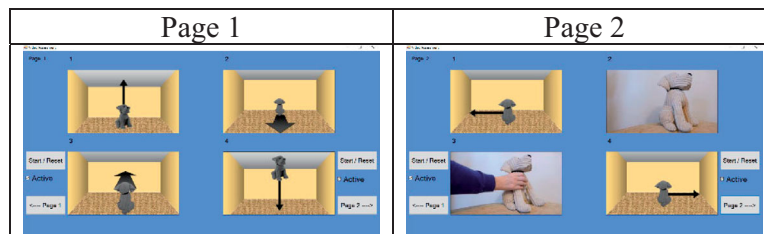


Figure 3. Video-based assessment: The two pages.

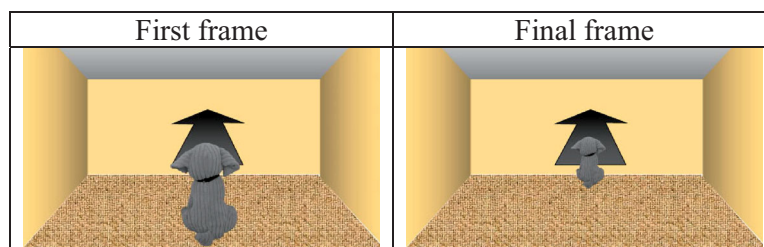


Figure 4. Video-based assessment: First and final frames of the “moving forwards” video clip (page 1, bottom-left cell).

Table 1. Differences in design and administration procedures of the static image and video-based assessments.

	Static image-based	Video-based
Method of question delivery	Synthesized speech – via the Grid 3. The participant can also activate a cell to listen to the question on demand	The SaLT reads the questions aloud
Answer selection method	Dwell-click (briefly fixating on an answer cell)	Continuous fixation. The SaLT determines which cell is being attended to based on the position of a visible cursor
Access check	4 questions, carried out through the digital assessment	No questions, the check was carried out by the SaLT (by visual observation)
Practice questions	3 questions, carried out through the digital assessment	None (assumed carry-over from previously completing the similar format static image-based assessment)
Recording of answers	Automatically by the digital assessment	Manually – by the SaLT
Answer cell type	Static image	Animation or video

cells in the assessment (Figure 2(1)). The system would present an image of the dog character in one of the cells, together with the question “where is the dog?” This would cycle through all four cells and the participant could only continue when all four questions were answered correctly. This addresses the assessment’s first aim, to determine whether the assessment approach is accessible by the participants.

2. Practice: A set of practice questions are presented to familiarize the participant with the assessment format (Figure 2(2)). The concepts here are nouns.
3. Main Assessment Questions: Following the same format as the practice questions, the main 17 assessment questions were presented to the participant and results logged automatically by the software. These questions related to prepositions and adjectives (colors); a complete listing is included in Appendix Table A2.
4. Video Assessment: After a brief explanation from the SaLT about what the assessment entailed, each participant completed a total of eight assessment questions. This time the SaLT read aloud questions of the format “Which one is...?” and recorded responses manually. These questions related to verbs and adverbs.

Table 2. P1 and P2: static image-based assessment results for baseline and outcome stages.

	P1		P2	
	Baseline	Outcome measures	Baseline	Outcome measures
Practice	3/3	3/3	2/3	3/3
Assessment	15/17	15/17	4/17	7/17
Totals	18/20	18/20	6/20	10/20

Results

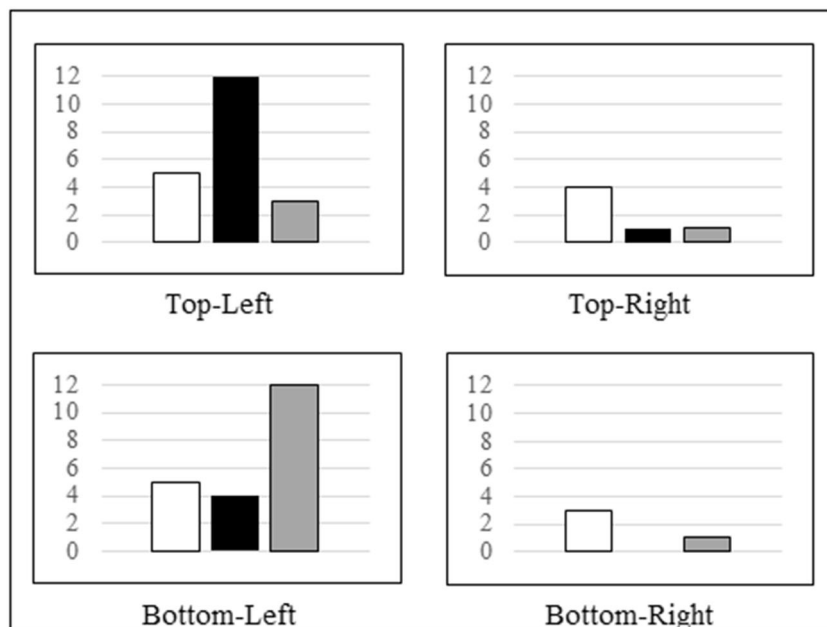
In this section the results of the assessments are presented for the purpose of evaluating their suitability and viability. The pilot study covered baseline and outcomes measurements, mainly related to the larger study. This demonstrates the use of the assessments for an experimental setup.

Static image-based assessment results

During the static image-based assessment, each participant completed a total of 24 questions: 4 access check questions, 3 practice questions and 17 assessment questions. Both participants completed all of the questions. An overall summary of the results of the static image-based assessments for both P1 and P2, for the baseline and outcome measures assessments, are displayed in Table 2. This shows that P1 scored a near maximum 15 correct answers out of a possible 17 questions in both the baseline and outcome measures stages of the static image-based assessment. Conversely, P2’s assessment scores were low across both baseline and outcome measures stages, with below chance scores at baseline and a small improvement at outcome.

Figure 5 shows the spread of P2’s answers in the static image-based assessment for both the baseline and outcomes stages. P2 appeared to favor answers positioned on the left side of the screen.

There were some changes in P2’s scores between the baseline and outcome measures but it is difficult to attribute these to progression: Figure 5 shows that P2’s answers were heavily skewed towards those cells positioned on the left side of the answer grid, with 16 out of 17 baseline and 15 out of 17 outcome measure answers being on the left. There did not appear to be any evidence of this behavior for P1.



Legend/Key

□	Number of actual correct answers in this position
■	Number of P2's answers given in this position for the baseline
■	Number of P2's answers given in this position for the outcome measures

Figure 5. P2: static image-based assessment results – answers displayed and grouped by cell position (baseline and outcome stages).

Video-based assessment results

For the video-based assessment, each participant completed a total of eight assessment questions spread over two onscreen pages. There was no explicit access check and no practice questions. Both participants completed all eight questions. An overall summary of the results of the video-based assessments for both P1 and P2, for the baseline and outcome measures assessments, is displayed in Table 3. This shows that P1 achieved high scores in the video-based assessments with a score of 6 out of 8 in both the baseline and outcome measures. P1 answered the same two questions incorrectly for both baseline and outcome measures stages. These questions involved the concepts of “Moving Forwards” and “Moving Backwards.” P1 gave the opposite answer for each concept depicted i.e., “Moving Forwards” for the question “Moving Backwards” and vice versa. It is worth noting that during the development and trials stages one of the SaLTs involved also interpreted these questions in this manner. P2’s score in the baseline video-based assessment was 4 out of 8 and this decreased to 3 out of 8 for the outcome measures.

Figure 6 shows the spread of P2’s answers in the video-based assessment for both the baseline and outcomes stages. It shows that similar to the static image-based assessment, P2 appeared to exhibit bias towards answers on the left side of the grid of 4. Six out of 8 of his answers were left-sided in both the baseline and outcome measures.

Table 3. P1 and P2: video-based assessment results (baseline and outcome stages).

P1		P2	
Baseline	Outcome measures	Baseline	Outcome measures
6/8	6/8	4/8	3/8

Discussion

The two new assessments were trialed in a pilot study with two individuals from the TG. The extent to which the assessment techniques and pilot study achieved the aims set are now discussed.

Assessment techniques: aims revisited

Aim 1

To ensure that the new assessments were AT accessible for the TG who usually use eye gaze as an alternative access method:

This aim was achieved, with both participants able to access and complete the assessments using eye gaze as a means of selecting or indicating their answers.

In common with other studies [15,21,22], communication software was used successfully as a “wrapper” for the static image-based assessment to provide a range of accessibility features. This also avoided the issue of misinterpreting gaze direction, as highlighted by Sargent et al. [9]. The limitations of existing communication software, identified earlier, precluded the adoption of a “wrapper” approach for the video-based assessment. Nevertheless, the bespoke approach used did deliver basic accessibility in the form of eye gaze interaction.

Aim 2

To represent concepts containing temporal, spatial or movement elements in a more suitable format:

Moving images were used to depict concepts which contain temporal, spatial or movement aspects, a format which maps more closely to the real world. The examinee was able to watch each video clip by fixating on the relevant cell.

Subjectively, comments made by one of the SaLTs who helped to develop and administer the assessments appeared to support the opinions of other authors [12] about the value of “moving”

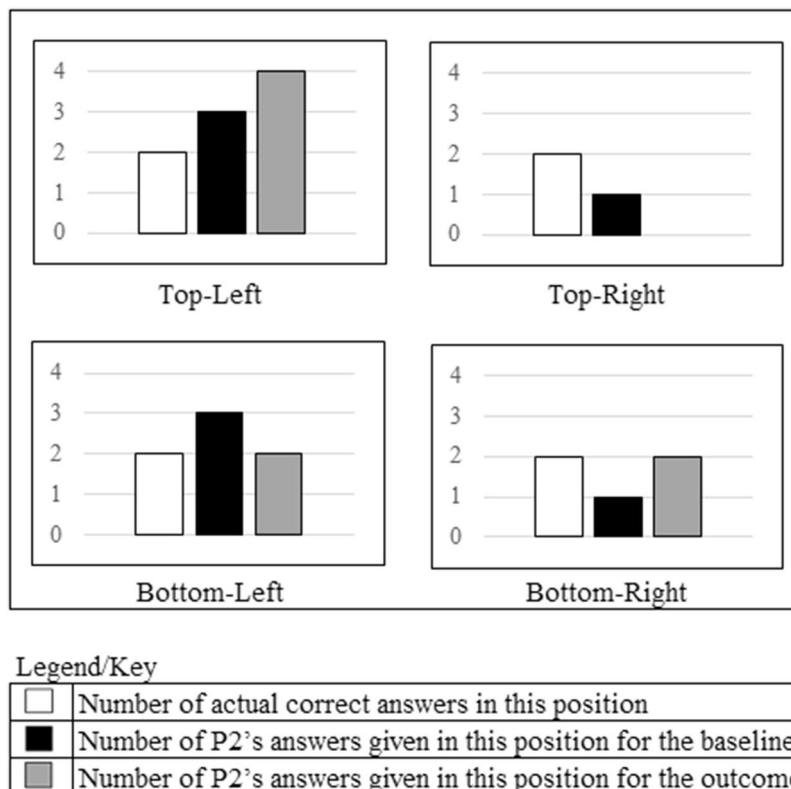


Figure 6. P2: video-based assessment results – answers displayed and grouped by cell position (baseline and outcome stages).

images in assessment. The SaLT reported: “The video assessment is particularly impressive. I’ve never seen this used before and gave a much better representation of the concepts.” The SaLTs who administered the assessments considered the videos to be more engaging to CYP and a better representation of the temporal, spatial or movement properties of real-world concepts such as “gripping” or “moving left.” Such concepts are difficult to represent using static images and usually involve the use of conventions which need to be understood such as arrows to indicate direction, or curved lines around moving parts to indicate motion [32].

Aim 3

To minimize the risk of assessor misinterpretation and bias:

The static image-based assessment gave the examinee more autonomy during the assessment by allowing them to select their answers directly using eye gaze. The amount of input required from the SaLT during the administration process was minimized, and the participants’ responses were logged automatically by the software, thereby reducing the risk of misinterpretation of answers or the introduction of bias. This design does bring an increased risk of accidental selections, as the answers are not confirmed by the SaLT, as may be the case with assisted scanning approaches.

The video-based assessment was less effective in this regard. There was a greater reliance on the SaLT during the assessment procedure i.e., asking the questions and interpreting and manually logging the answers, increasing the risk of misinterpreted answers or assessor bias.

Pilot study: aim revisited

The aim of the pilot study was to provide an indication of whether the participants’ knowledge of the concepts under investigation could be evaluated by the assessments:

This appeared to be the case for P1 with high scores throughout the assessments. This was less clear for P2 whose scores were relatively low throughout the assessments.

The results appear to indicate that P1 was able to reveal his knowledge of the concepts being assessed to a high level. In fact the scores were so high in the baseline that it was not possible to measure progression between the baseline and outcome measures. The assessments appeared to be less suitable for uncovering the knowledge of P2.

P2 predominantly chose answer cells positioned on the left in both the static image-based and video-based assessments, over both baseline and outcome measures stages. P2 did not appear to exhibit any signs of anxiety or distress and appeared to be relaxed during all of the assessments. This bias towards answers positioned on the left may be indicative of an inability to inhibit, side-preference or perseveration behaviors [10,33,34], perhaps exacerbated through difficulty understanding the concepts being tested. P2’s SaLT commented that, in her opinion, he had exhibited perseveration behaviors during assessments undertaken by her on other occasions. Another possible explanation for P2’s answering behavior may be that he had limited experience of initiating or making choices and so found the choice-making aspect of the assessments difficult.

It is interesting to note that during the practice stage of the assessment P2 scored 2 out of 3 in the baseline and 3 out of 3 in the outcome. P2 also achieved proportionately higher scores for assessment questions relating to colors, achieving 2 out of 4 correct answers in the baseline and 3 out of 4 correct answers in the outcomes. These results may indicate that P2 had a better understanding of nouns and noun attributes (i.e., colors) compared to the other concepts being tested. Golinkoff et al. [24] state that

“nouns are easy; verbs are hard” for young children and Hsu and Bishop’s [32] study comparing language-impaired children with age-matched and grammar-matched controls discuss the difficulties children can have with understanding spatial prepositions. P2 may have experienced similar difficulties.

Additional information of value emerged from the use of the assessments in the pilot study. P1 scored more highly than had been anticipated, surprising the staff who works with him and challenging their perceptions. This has also been observed by other authors where accessibility options appropriate to an individual’s needs have been employed [20]. P2’s scores reinforced that other assessment techniques and approaches are needed to verify his knowledge and understanding. Furthermore, P2’s difficulties raise the issue that the new assessments may be accessible using AT but perhaps not “cognitively” accessible. Further work needs to be undertaken to understand perseveration, side preference and other behaviors that may affect assessment, and how assessments might be designed to allow for such behaviors.

Conclusion and future work

The overall goal of this work was to provide assessment methods which are accessible by the TG with a focus on the use of eye gaze to provide accessibility and to support the assessment of a set of temporal, spatial, or movement concepts. This was motivated by the needs of a larger study and a lack of suitable existing assessment techniques. Two complementary assessment methods were presented, using static images and video clips to represent the concepts in an appropriate format. A pilot study was conducted with two participants to evaluate the effectiveness and suitability of the assessment techniques.

The results indicate that the new assessment techniques were accessible to both participants. They were effective in revealing knowledge in one participant but were inconclusive for the other. In this second case, a side preference or perseveration behavior was observed, further demonstrating some of the challenges in assessing the TG.

The pilot study and the assessment techniques described here go some way to addressing the limitations of conventional spoken language comprehension assessments for individuals with profound motor impairments. However, further developments in this field are required.

Lack of comparison

Ideally, the new assessment techniques should be compared with existing solutions to determine how effective they are in revealing knowledge and understanding. The main motivation for this work was the current lack of suitable assessment techniques for the TG and, therefore, in this context, the work provides a baseline for future comparisons. Future work could be undertaken with individuals similar to the TG, who have verbal skills, or with typically developing children and young people. Comparisons could then be made to help verify the effectiveness of the new techniques.

Hearing impairments

In their current form, both assessments are not suitable for those who have hearing impairments.

Video-based assessment features

A comparison of the two assessments identified a number of possible improvements for the video-based assessment. Ideally, both

assessments should have been more aligned in terms of design, accessibility, and administration. Most of the differences were due to the static image-based assessment being contained within Grid 3 (and so inheriting Grid 3's accessibility features), whereas the video-based assessment incorporated new software, developed by the first author. Recommended modifications to the video-based assessment are as follows:

1. Incorporation of an explicit access check: Analysis of video recordings revealed that P2 had not fixated upon and watched all of the video cells prior to the first assessment question, possibly affecting his answer. An access check would ensure this cannot happen.
2. Incorporation of practice questions: The static image-based assessment contained a series of practice questions to familiarize the examinee with the format and procedure. The video-based assessment did not contain practice questions and it should not have been assumed that the participants knew what was expected of them.
3. The inclusion of speech synthesis: Questions should be "read out" by the assessment to maintain consistency with the approach used by the static image-based assessment.
4. Include the ability to dwell-select cells: The participant should be able to dwell-select answer cells for themselves. Instead, for the video-based assessment, answers were given by the examinee fixating upon an answer cell and the SaLT interpreting and confirming the cell. This potentially introduces additional cognitive loading for the examinee and increases the risk of assessor misinterpretation or bias. This could be mitigated against by a change in the design i.e., when the examinee is ready to answer, the software enters a "dwell-select" mode;
5. Automatic logging of the answers: The video-based assessment requires the administering SaLT to interpret and note examinee answers manually.

Further recommendations

Rather than "reinventing the wheel," it seems logical, as suggested by Warschausky et al. [15], to digitally adapt standardized assessments and to embed these within communication software. Developing bespoke assessment material, as was done in this study, means that there is no standardization of the data gathered and to qualify as a "valid" assessment, it would need to go through a protracted process of validation and trials (as has been done with Geytenbeek et al.'s [4] C-BiLLT). As a starting point, assessment providers could develop additional administration procedures that have greater flexibility for those who use alternative access methods, as has been done by Geytenbeek et al. [4] and Guerette et al. [3].

The results of the pilot study and the new AT accessible assessments described here indicate an advance in our ability to assess some members of the TG. The authors suggest that standardized assessment providers, communication software manufacturers and researchers work together to ameliorate the current situation and provide more appropriate assessments for this group.

Acknowledgements

The research described here was carried out at Livability Victoria Education Centre with the assistance of both pupils and staff. The

authors thank everyone involved, with special thanks to Nicola Mearing, Sian Awford and Ellen Johnson.

Disclosure statement

Grid 3 communication software was provided by Smartbox Assistive Technology Limited. The MindExpress communication software and CARLA software assessment packages were both provided by Techcess Communications Limited.

Funding

This work was supported by the Engineering and Physical Sciences Research Council [EP/L016540/1].

ORCID

Mark Moseley  <http://orcid.org/0000-0001-7292-0218>

Leigh McLoughlin  <http://orcid.org/0000-0002-7566-3568>

References

- [1] Geytenbeek J, Harlaar L, Stam M, et al. Utility of language comprehension tests for unintelligible or non-speaking children with cerebral palsy: a systematic review. *Dev Med Child Neurol.* 2010;52(12):e267–e277.
- [2] Yin Foo R, Guppy M, Johnston LM. Intelligence assessments for children with cerebral palsy: a systematic review. *Dev Med Child Neurol.* 2013;55(10):911–918.
- [3] Guerette P, Tefft D, Furumasa J, et al. Development of a cognitive assessment battery for young children with physical impairments. *Infant-toddler intervention. Transdisciplinary J.* 1999;9(2):169–184.
- [4] Geytenbeek JJ, Mokkink LB, Knol DL, et al. Reliability and validity of the C-BiLLT: A new instrument to assess comprehension of spoken language in young children with cerebral palsy and complex communication needs. *Augmentative Alternative Commun.* 2014;30(3):252–266.
- [5] Encarnação P, Alvarez L, Rios A, et al. Using virtual robot-mediated play activities to assess cognitive skills. *Disabil Rehabil Assist Technol.* 2014;9(3):231–241.
- [6] Geytenbeek JJM, Heim MMJ, Vermeulen RJ, et al. Assessing comprehension of spoken language in nonspeaking children with cerebral palsy: Application of a newly developed computer-based instrument. *Augmentative Alternative Commun.* 2010;26(2):97–107.
- [7] Watson RM, Pennington L. Assessment and management of the communication difficulties of children with cerebral palsy: A UK survey of SLT practice. *Int J Lang Commun Disord.* 2015;50(2):241–259.
- [8] Pennington L. Symposium: Special needs: Cerebral palsy and communication [Article]. *Paediatrics Child Health* 2008; 18(9):405–409.
- [9] Sargent J, Clarke M, Price K, et al. Use of eye-pointing by children with cerebral palsy: What are we looking at? *Int J Lang Commun Disord.* 2013;48(5):477–485.
- [10] Golinkoff RM, Hirsh-Pasek K, Cauley KM, et al. The eyes have it: lexical and syntactic comprehension in a new paradigm. *J Child Lang.* 1987;14(1):23–45.
- [11] Tobii Dynavox LLC. PCS animations. 2018 [cited 2018 Dec 12]. Available from: <https://goboardmaker.com/collections/all/products/pcs-animations-bundle>

- [12] Snyder K, Higbee TS, Dayton E. Preliminary investigation of a video-based stimulus preference assessment. *J Appl Behav Anal.* 2012; 45(2):413–418.
- [13] Pearson Education Ltd. Q-interactive: Pearson. [cited 2018 Dec 13]. Available from: <https://www.pearsonclinical.co.uk/q-interactive/q-interactive.aspx>
- [14] Friend M, Keplinger M. An infant-based assessment of early lexicon acquisition. *Behav Res Methods Instrum Comput.* 2003; 0135(2):302–309.
- [15] Warschausky S, Van Tubbergen M, Asbell S, et al. Modified test administration using assistive technology: Preliminary psychometric findings. *Assessment* 2012;19(4):472–479.
- [16] Tobii Dynavox LLC. Boardmaker with speaking dynamically pro v.6. 2018 [cited 2018 Dec 13]. Available from: <https://goboardmaker.com/collections/boardmaker-software/products/boardmaker-with-speaking-dynamically-pro-v-6>
- [17] Byrne JM, Dywan CA, Connolly JF. An innovative method to assess the receptive vocabulary of children with cerebral palsy using event-related brain potentials. *J Clin Exp Neuropsychol.* 1995;17(1):9–19.
- [18] Huggins JE, Alcaide-Aguirre RE, Aref AW, et al. Brain-computer interface administration of the Peabody Picture Vocabulary Test-IV. 7th International IEEE/EMBS Conference on Neural Engineering (NER); 2015 Apr 22–24; Montpellier, France. *IEEE Xplore*; 2015. p. 29–32.
- [19] Dunn LM, Dunn DM, Pearson A. PPVT-4: Peabody picture vocabulary test. Minneapolis, MN: Pearson Assessments; 2007.
- [20] Ahonniska-Assa J, Polack O, Saraf E, et al. Assessing cognitive functioning in females with Rett syndrome by eye-tracking methodology. *Eur J Paediatr Neurol.* 2018;22(1): 39–45.
- [21] Techcess Communications Ltd., National Institute of Health Research, Devices for Dignity, et al. CARLA 2018 [cited 2018 13 Dec]. Available from: <https://www.techcess.co.uk/carla1/>
- [22] Techcess Communications Ltd. Mind express. 2018 [cited 2018 Dec 13]. Available from: <https://www.techcess.co.uk/mind-express/>
- [23] Edwards S, Letts C, Sinka I. The New Reynell developmental language scales. London, UK: GL Assessment Limited; 2011.
- [24] Golinkoff RM, Ma W, Song L, et al. Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned? *Perspect Psychol Sci.* 2013;8(3):316–339.
- [25] Smartbox Assistive Technology Limited. Smartbox. <https://thinksmartbox.com/2016> [cited 2016 Dec 05]. Available from: <https://thinksmartbox.com/>
- [26] Palisano R, Rosenbaum P, Bartlett D, et al. Gross motor function classification system - expanded and revised (GMFCS - E & R) - Level V. 2007. Available from: https://canchild.ca/system/tenon/assets/attachments/000/001/399/original/GMFCS_English_Illustrations.pdf
- [27] Eliasson A-C, Krumlind-Sundholm L, Rösblad B, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: Scale development and evidence of validity and reliability. *Dev Med Child Neurol.* 2006;48(7): 549–554.
- [28] Department for Education. Performance - P scale - Attainment targets for pupils with special educational needs. UK: Department for Education; 2017 [cited 2018 Dec 14]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/617033/Performance_-_P_Scale_-_attainment_targets_for_pupils_with_special_educational_needs_June_2017.pdf
- [29] Smartbox Assistive Technology Limited. Look to learn. <https://thinksmartbox.com/product/look-to-learn/>. 2018 [cited 2018 Dec 13]. Available from: <https://thinksmartbox.com/product/look-to-learn/>
- [30] Tools for windows apps and games [Internet]. Microsoft; [cited 2019 Oct 1]. Available from: <http://visualstudio.microsoft.com/vs/features/windows-apps-games/>
- [31] Jacob RJK. What you look at is what you get: eye movement-based interaction techniques. *Conference on Human Factors in Computing Systems Proceedings*; 1990.
- [32] Hsu HJ, Bishop D. Training understanding of reversible sentences: A study comparing language-impaired children with age-matched and grammar-matched controls. *PeerJ.* 2014;2:e656.
- [33] Berger SE. Demands on finite cognitive capacity cause infants' perseverative errors. *Infancy* 2004;5(2):217–238.
- [34] Piaget J. The construction of reality in the child. New York: Basic Books; 1954.

Appendix

Table A1. A sample of a static image-based assessment log file.

Date	Time	Page	Cell/correct answer	Given answer	Correct/incorrect
27/6/2017	14:16:47	Right	Which one is right?		
27/6/2017	14:16:54	Right	Which one is right?		
27/6/2017	14:17:01	Right	Right (Top-Left)	Left (Bottom-Right)	Incorrect
27/6/2017	14:17:07	Behind	Which one is behind the table?		
27/6/2017	14:17:20	Behind	Which one is behind the table?		
27/6/2017	14:17:27	Behind	Behind (Top-Right)	Behind (Top-Right)	Correct
27/6/2017	14:18:09	Lower	Which one is lower?		
27/6/2017	14:18:18	Lower	Which one is lower?		
27/6/2017	14:18:24	Lower	Lower (Top-Right)	Lower (Top-Right)	Correct

Table A2. The words that were tested within the static image-based assessment.

	Concept	Type
Practice		
1	Horse	Noun
2	Boat	Noun
3	Banana	Noun
Assessment		
1	Red	Adjective
2	Above	Preposition
3	Green	Adjective
4	Left	Preposition
5	In front	Preposition
6	Middle	Preposition
7	Blue	Adjective
8	Right	Preposition
9	Behind	Preposition
10	Lower	Preposition
11	Yellow	Adjective
12	On	Preposition
13	Higher	Preposition
14	Under	Preposition
15	In	Preposition
16	Far apart	Preposition
17	Near	Preposition