Author Name: Jane Strickland
Contact Email: jane.strickland@unimelb.edu.au

ASSESSING AND UNDERSTANDING EARLY NUMERACY FOR STUDENTS WITH ADDITIONAL LEARNING NEEDS

Jane Strickland, Kerry Woods, and Masa Pavlovic
Assessment Research Centre, Melbourne Graduate School of Education

Abstract

Almost 5% of Australian students fail to reach minimum standards for numeracy as assessed on the national assessment and monitoring program, and this is of particular concern in light of evidence that poor numeracy contributes to lifelong problems such as difficulty getting and maintaining employment. Learning progressions, particularly in mathematics, have been identified as effective ways to support teachers' understanding of skill development and connect assessment and pedagogy. The aim of this study was to design an assessment of early numeracy skills for students with additional learning needs that could be mapped to a learning progression. Trialled with 2597 students from 65 schools, a set of items based on teacher observation and judgement measured skills that provide a foundation for participation in numeracy learning across the school years. Analysis of the data, using a partial credit approach to item response modelling, found that the items could describe numeracy learning across a broad range of proficiency, although gaps were identified for students at the highest and lowest extremes of numeracy understanding. Analysis of differential item functioning (DIF) suggested that the items measured numeracy skills and understanding consistently for students with different sub-types of additional learning needs.

Introduction

Data gathered from recent national tests of numeracy showed that between 4% and 5% of Australian students were either working below the accepted minimum level of numeracy ability or exempt from testing due to the nature or severity of their additional learning needs (Australian Curriculum, Assessment and Reporting Authority, 2015). Students typically enter school with a wide range of numeracy skills and understanding (Fuson, Clements & Sarama, 2015; Powell & Fuchs, 2012), and factors such as early life experiences or the presence of a general or specific disability may impede or complicate their opportunities to build a firm foundation for numeracy learning. Parent-child numeracy activities are important predictors of a student's success in the development of numeracy skills (Kleemans, Peeters, Segers & Verhoeven, 2012), and the participation of children in numeracy-based activities at home (LeFevre et al., 2009) or in preschool (Jung, 2011) supports the learning of mathematics in school. However, many students with general or specific disabilities may not absorb numeracy concepts that are implicit in daily activities. They do not develop a *number sense* (Butterworth, 1999; Dehaene, 1997) or the early numeracy competencies that they need to establish before learning more complex mathematical concepts (Powell & Fuchs).

The role of the teacher

While teachers have a significant impact on the learning outcomes of their students, not all teachers are effective (Hattie, 2009) and teaching practices commonly promoted as effective are not used or understood by all teachers. The teaching practice of differentiation, for example, is much talked about (yet poorly understood) (Birnie, 2015) and has strong links to Vygotsky's (1923/1993) recommendation that teaching focus on finding the current ability of the student and targeting instruction accordingly (Birnie). This contrasts with evidence presented to a Federal Senate Committee investigation into education for students with additional needs (Commonwealth of Australia, 2016) that revealed teachers were modifying curriculum without first assessing the relevant skills of the student. This was further compounded by evidence that many teachers have low expectations for students and many programs for students with additional needs lack academic rigour (Commonwealth of Australia). These issues may be indicative of a wider and more critical issue: that teachers are unaware of how best to educate some of their students with additional needs.

One way to support teachers' understanding of their students as learners may be the provision and use of rigorously constructed learning progressions to guide instructional decisions (Heritage, 2008). The notion of building blocks that are foundational to more complex learning, or scaffolds that support students as they work towards more sophisticated levels of skills and understanding (Vygotsky, 1929/1993), are consistent with representations of learning as a pathway or progression. For example, Heritage described learning progressions as pathways along which students move in any particular knowledge domain. Further, Black, Wilson and Yao (2011) described the progression of learning in terms of a 'road map'. The emphasis is on a trajectory that sets out the ways in which student understanding unfolds, so that teachers can best decide on the next steps for promoting and monitoring student learning. This approach builds on the Vygotskian vision that teachers should first ascertain their students' strengths or starting point for learning, and then scaffold learning and target their teaching accordingly. A scaffold, as used here, represents a structure to assist a student to internalise tools and concepts provided by a more knowledgeable other (Vygotsky). A formalised scaffold might assist teachers in their instruction when presented as a progression of increasing skills and knowledge. Similarly, students can use learning progressions to gain a clear sense of the things they need to strive towards as they build their skills and understanding.

Building learning progressions from criterion-referenced frameworks

In previous research, a criterion-referenced framework was successfully used to design an assessment of early literacy skills for students with additional needs (Woods & Griffin, 2013). That study reported student outcomes as locations on a learning progression, and provided partial validation for the assessment by demonstrating the extent to which teachers could use their students' reports to plan, implement, and monitor programs of instruction (Woods & Griffin). The capacity of the assessment to describe the early literacy learning of students with additional needs, and the strong correlation between literacy and numeracy skills (Purpura & Napoli, 2015), supported development of a similar assessment, following a similar design methodology, to measure early numeracy for this cohort of students. The intended benefits of such an approach were to support teachers' understanding of the needs of their students and direct more sensitive and competent observations of student learning (Woods & Griffin).

Therefore, the aim of the study described in this paper was to design an assessment of early numeracy skills and understanding suited to the purpose of supporting teaching and learning for students with additional needs. The development of core concepts of number/symbol awareness is well documented in the literature and includes proposals of an innate pre-verbal ability to identify, represent, and manipulate quantities (Dehaene, 1997; Gallistel & Gelman, 1992, Gelman & Butterworth, 2005). It is this innate *number sense* that is built upon, through the use of language and symbolic representation, to develop the core skills of counting and arithmetic (Butterworth, 1999). An initial stage of the research drew upon the subject matter expertise of academics and specialist teachers to identify and refine a definition of the educational construct of interest and a set of representative capabilities to be covered by a pool of assessment items. The educational construct was defined as 'using symbolic representation of quantity, shape and pattern to bring order to the world'. Figure 1 shows a diagrammatic representation developed to describe the construct and hypothesised increasing levels of numeracy competence.

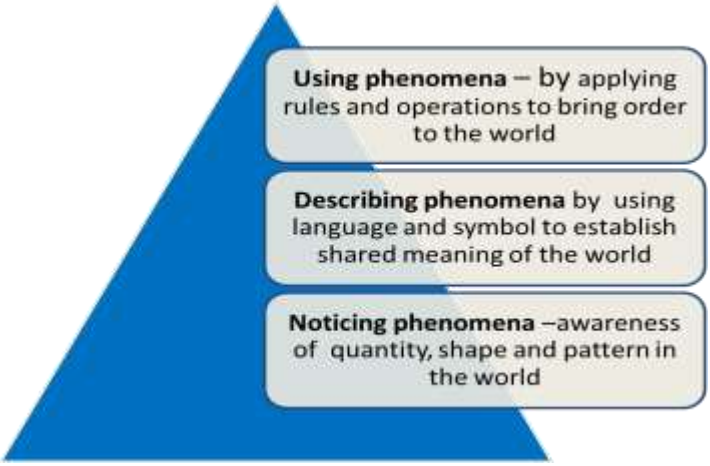
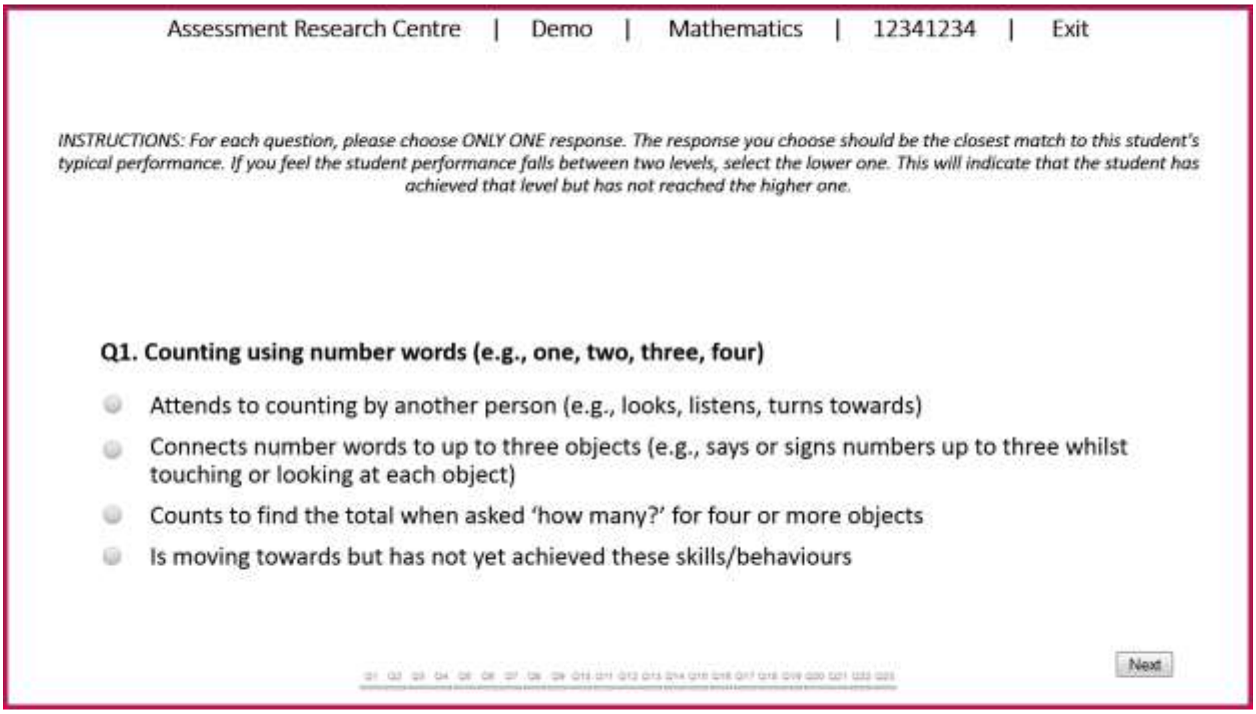


Figure 1. Hypothesised taxonomy of numeracy progression

As shown in Figure 1, the construct of interest included an initial capacity to notice and show awareness of numeracy concepts such as quantity, shape and pattern. It was important to develop an item pool that was inclusive of all learners, including those with severe and complex additional needs who may require high levels of support to engage in learning. The initial level - ‘noticing phenomena’ – was intended to describe the skills of students who are typically learning to react or respond to numeracy activities. This could, for example, involve repeated exposure to numeracy concepts or experiences, or highly explicit teaching. It was hypothesised as a first step towards engaging independently in numeracy learning. The middle level - ‘describing phenomena’ - referred to an ability to demonstrate understanding of numeracy concepts through the use of receptive language (e.g., acting on instructions to halve a group or selecting the biggest item from a collection of objects), and also sharing knowledge of numeracy concepts and experience through the use of expressive language (e.g., being able to describe one object as longer or shorter than another, or use number words to count a set of objects). This linked to the construct definition of ‘bringing order’ to experience, as students can make requests and follow instructions pertaining to everyday numeracy concepts more successfully if they have access to a more comprehensive and specific vocabulary. At the top level of the diagram is the capability of ‘using phenomena’, which described students who are learning to manipulate numbers and apply operations to solve problems. A pool of assessment items was then written purposively to describe the use of numeracy skills and concepts to ‘bring order’ to students’ experiences of daily activities with emphasis placed on the functional application of numeracy skills and understanding. Notably, the pool of items did not include skills with little or no functional purpose, such as rote counting for example.

Each item in the draft item pool was presented in a format suited to partial credit scoring, with a question (described as a behavioural indicator) and a corresponding series of response choices (described as quality criteria) that were ordered in increasing complexity. The pool of items was piloted in schools with 14 teachers of students with additional needs, and then prepared for trial in an online format. An example question is shown in Figure 2. A list of the 23 skills in the item pool is shown in Table 1.



INSTRUCTIONS: For each question, please choose ONLY ONE response. The response you choose should be the closest match to this student's typical performance. If you feel the student performance falls between two levels, select the lower one. This will indicate that the student has achieved that level but has not reached the higher one.

Q1. Counting using number words (e.g., one, two, three, four)

○ Attends to counting by another person (e.g., looks, listens, turns towards)
○ Connects number words to up to three objects (e.g., says or signs numbers up to three whilst touching or looking at each object)
○ Counts to find the total when asked 'how many?' for four or more objects
○ Is moving towards but has not yet achieved these skills/behaviours

Figure 2. Example of the presentation of a numeracy item

Table 1
The pool of numeracy items

Item number	Item description
1	Counting using number words
2	Conservation of number
3	Recognising quantities without counting
4	Connecting numerals to number words
5	Ordering numerals
6	Dividing an object
7	Dividing a collection of objects
8	Describing a change in quantity
9	Adding items to a group
10	Subtracting items from a group
11	Skip counting
12	Working with common 2D shapes and objects
13	Completing written number patterns
14	Comparing length of common objects
15	Comparing the size of common objects
16	Sorting objects into categories
17	Describing relative position
18	Recognising and making patterns
19	Using ordinal number
20	Predicting outcomes of events
21	Interpreting representations of information
22	Working with concepts of time/duration
23	Identifying/naming Australian coins

As shown in Table 1, the skills included in the item pool related to students' early or foundational understanding of numeracy concepts, with an emphasis on number, shape, and pattern. Some items were included in the item pool to improve the face and content validity of the assessment for teachers. These related to students' understanding of concepts related to time or duration and the use of money.

Author Name: Jane Strickland
Contact Email: jane.strickland@unimelb.edu.au

Some of the more demanding steps or criteria within items drew upon more complex operations, such as addition, subtraction, or solving number problems, while others asked about a student's understanding of the concept of probability to predict the likely outcome of situations. Together, these 23 items were intended to describe the early numeracy skills and understanding of students with additional learning needs. The research questions were thus linked to the development of an assessment instrument, where student outcomes could be reported as a location on a learning progression. They were articulated as follows:

- To what extent can an assessment of early numeracy skills and understanding be designed and mapped to a learning progression for students with additional learning needs?
- How well did the pool of assessment items describe learning for students with different types of additional learning needs?

Method

Participants

Teachers from 65 schools assessed 2597 students using the pool of items. Participants included teachers and students from specialist and mainstream schools in Victoria and Western Australia, for whom the principal consented to participation of the school in the research. Teachers and parent/guardians of the students also gave written consent to participate after ethics approval was gained through the University ethics approval process.

The students assessed in the study ranged in age from 2 to 24 years, with the majority aged between 7 and 16 years and with an average age of 12 years (SD = 3.5 years). All students in the study had been identified by their teachers as requiring additional support in numeracy, and a large proportion had co-occurring conditions that impacted their learning. Teachers selected up to five students to assess, and they were asked to choose students who were representative of the range of abilities, ages, gender, and disabilities in their classroom. Table 2 displays the characteristics of students in the study, including the number of students identified by their classroom teachers as having different sub-types of additional learning needs. It should be noted that students could be identified as having more than one disability or form of additional need.

Table 2
Characteristics of students assessed in the trial of the item pool

	Number of students	% (of 2597 students)
Gender		
Boys	1870	72
Girls	727	28
School type		
Mainstream	99	4
Specialist	2498	96
Disability type		
Intellectual disability	1870	72
Autism spectrum disorder	1402	54
Severe language disorder	649	25
Vision impairment	208	8
Hearing impairment	130	5

Materials

Teachers responded to the pool of 23 items that were presented in an online format, provided to schools as part of the Victorian Department of Education and Training's (2016) Abilities Based Learning and Education Support (ABLES) materials. An example of an item is shown in Figure 2 above. Teachers were instructed to read each of the items presented on screen, and asked to select the option or criterion that best described the proficiency of the student under consideration. Items provided three or four response options from which teachers could choose, ranging from a beginning skill for each item to a relatively high level of proficiency. Teachers could also select an option that indicated the student had not yet achieved any of the skills described in the item.

Analysis

In order to provide evidence of construct validity, data were analysed using a partial credit approach to item response modelling (Masters, 1982). This allows for analysis of the items to identify how well each fit the underlying Rasch (1960) model and to examine the overall reliability of the instrument. Through use of Conquest IRT software (Adams, Wu, Macaskill, Haldane & Xun Sun, 2015), estimates of item and step difficulty, standard errors of measurement for these estimates, mean square fit statistics, Cronbach's alpha reliability indices, and item and person separation reliability indices were generated. Item difficulty parameters included Thurstone thresholds. These mark points along the ability continuum at which the cumulative probability of being assigned the successive score category reaches 0.5 (or a pre-determined response probability). The analysis generates a variable map to plot both the student ability and item difficulty estimates on one axis. Mapping both ability level of students and item difficulty onto one scale supports interpretation of the data as a continuum or progression of early numeracy skills.

A second analysis is an investigation into differential item function (DIF). DIF refers to a family of statistical methods used to detect performance differences on individual items, and it is an important part of scale development and the validation process. Historically, DIF analysis was used to detect differences between group (such as gender, race, or language background) performance on an item that cannot be explained by the underlying trait measured by items. However, the presence of DIF is not always evidence of item bias. DIF can be the result of the presence of a secondary dimension in the assessment that is important for the trait being measured. An item is considered biased if differences between group performance on the item are caused by sources that are irrelevant to the measured trait. Since it is not always obvious what causes difference in performance on an item, those flagged for the presence of DIF must be further examined by relevant subject matter experts. Where some items are identified as more or less difficult, relative to the other items and for a specific group in the sample, then this can hold implications for the interpretation of a measured trait. In the application of IRT, DIF occurs when the conditional probability of a correct response differs for two or more groups in the population. Therefore, DIF can be investigated by comparing item characteristic curves of two or more groups in the sample. Since characteristic score curves are completely defined by corresponding item parameters, DIF can be detected by comparing item parameters. This can be done by comparing differences in item difficulty parameters obtained by calibrations of items using responses from different sub-groups. This method is suitable when observed DIF is uniform and adequate fit for individual group calibrations is established.

Results and discussion

Initial analysis showed that the pool of items had a high level of internal consistency, as demonstrated by a Cronbach's alpha reliability of 0.98. As shown in Table 3, calibration of item estimates produced a range of 1.34 to 0.69 for the weighted mean squares (MNSQ) fit statistics of the 23 items. These statistics are based upon the difference between observed and expected scores (known as the residual) and indicate how well the expected observations fit the Rasch model (Adams & Khoo, 1996). There are different opinions on the upper and lower limits for good weighted MNSQ fit statistics (Wilson, 2005), with Adams and Khoo suggesting a lower limit of 0.75 and an upper limit of 1.33 while Wright, Linacre, Gustafson and Martin-Lof (1994) argued that a lower limit of 0.5 and an upper limit of 1.7 was reasonable depending on the type and purpose of measurement. Only two items in the pool showed a weighted MNSQ fit statistic above 1.3. While the source of misfit could not be determined with certainty, one possibility was that they referred to skills that were more difficult for teachers to observe in classrooms. Item 20 was designed to assess how well students used concepts of chance and probability to predict the outcome of events. The purpose of item 21 was to assess how well students

Author Name: Jane Strickland
Contact Email: jane.strickland@unimelb.edu.au

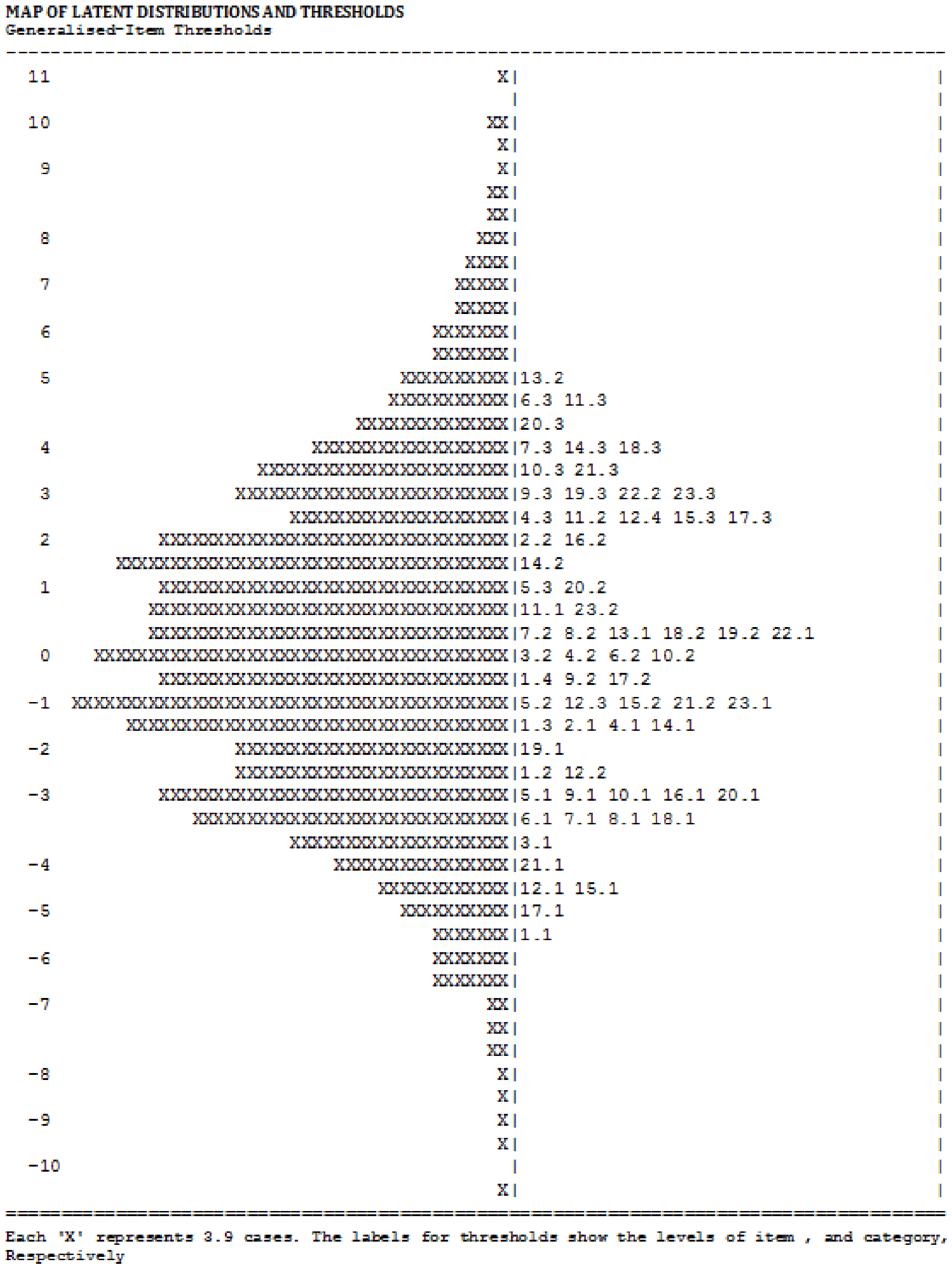
could interpret representations of information (e.g., charts, timetables, maps). Both items will be reviewed with subject matter experts to determine whether they can be improved, and then subjected to a further round of trials in schools before decisions are taken about their inclusion in the final version of the assessment instrument.

Table 3
Calibration estimates for the item pool

Number	Item Description	Estimate	Error	Weighted MNSQ
1	Counting using number words	-2.44	0.04	1.13
2	Conservation of number	0.30	0.04	0.84
3	Recognising quantities without counting	-1.92	0.05	1.00
4	Connecting numerals to number words	0.20	0.04	1.29
5	Ordering numerals	-1.02	0.04	0.90
6	Dividing an object	0.41	0.04	0.99
7	Dividing a collection of objects	0.24	0.04	0.89
8	Describing a change in quantity	-1.63	0.05	0.76
9	Adding items to a group	-0.18	0.04	0.73
10	Subtracting items from a group	0.17	0.04	0.74
11	Skip counting	2.58	0.05	0.69
12	Working with common 2D shapes and objects	-1.38	0.04	1.30
13	Completing written number patterns	2.75	0.06	0.79
14	Comparing length of common objects	1.35	0.04	0.82
15	Comparing the size of common objects	-1.08	0.04	0.95
16	Sorting objects into categories	-0.30	0.05	0.94
17	Describing relative position	-1.21	0.05	0.89
18	Recognising and making patterns	0.31	0.04	0.83
19	Using ordinal number	0.50	0.04	0.87
20	Predicting outcomes of events	0.69	0.04	1.34
21	Interpreting representations of information	-0.72	0.04	1.31
22	Working with concepts of time/duration	1.57	0.05	0.78
23	Identifying/naming Australian coins	0.81	0.04	1.03

A variable map (Figure 3) generated by the *ConQuest* IRT software (Adams et al., 2015) showed an overall good match between student ability and item difficulty. On the map, students are represented by X’s and the distribution of observed student abilities is spread along a vertical axis with students of low ability, as assessed by the pool of items, shown at the lower extremes of the axis and students of higher ability shown at relatively higher positions on the scale. The numbers on the right side of the map represent the relative difficulty of quality criteria (or steps) for each item. The lowest item, for instance, represents the first step/quality criterion of the first item (as shown in Figure 2) and was the easiest of the pool of items, therefore demanding the least student ability overall. Criteria requiring lower levels of proficiency are displayed at the bottom of the scale, and progressively higher positions of criteria are associated with higher relative difficulty.

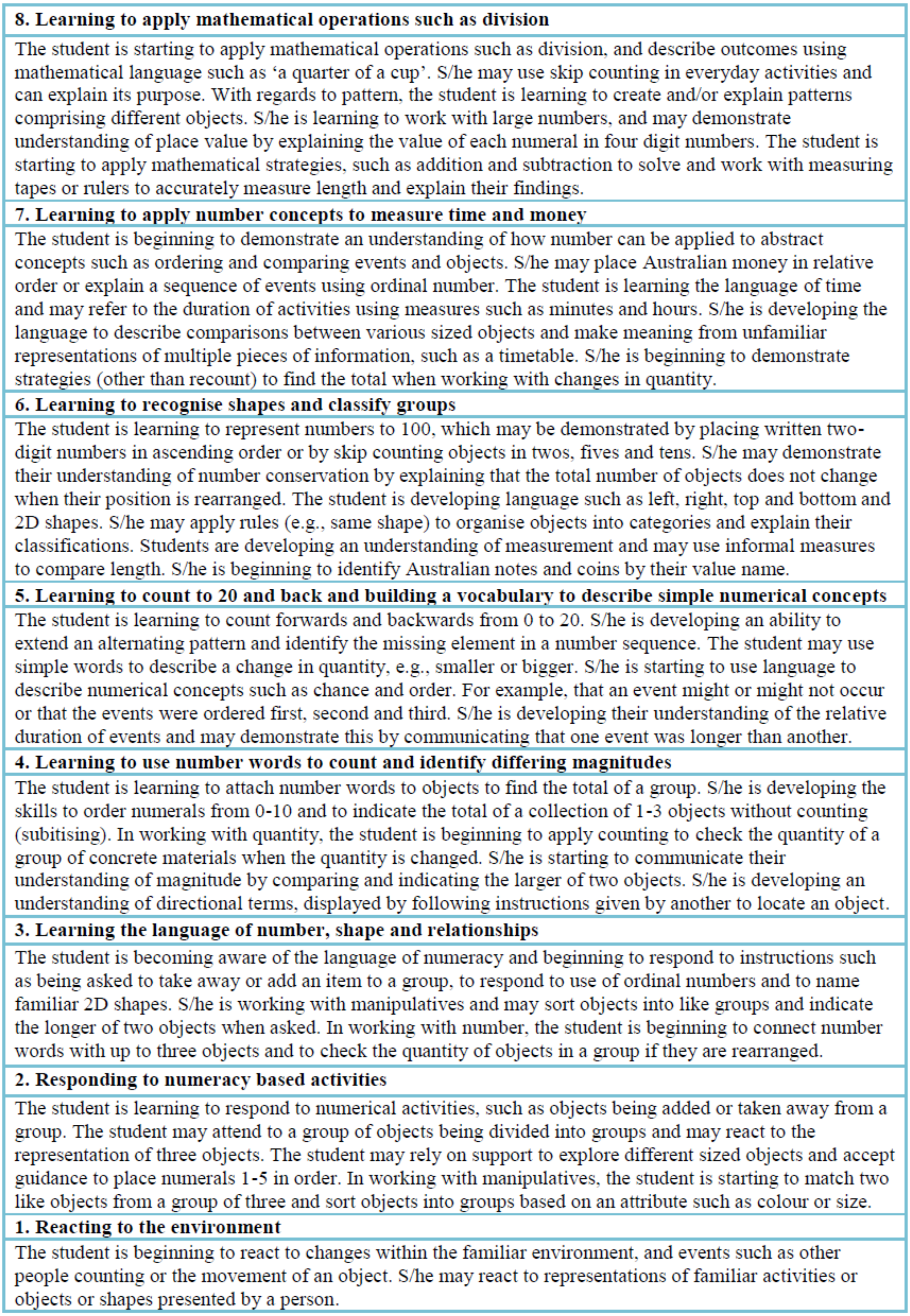
Author Name: Jane Strickland
Contact Email: jane.strickland@unimelb.edu.au



```
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
Generalised-Item Thresholds
---------------------------------------------------------------------------------------
  11                                         X|                                       |
                                              |                                       |
  10                                        XX|                                       |
                                             X|                                       |
   9                                         X|                                       |
                                            XX|                                       |
                                            XX|                                       |
   8                                       XXX|                                       |
                                          XXXX|                                       |
   7                                     XXXXX|                                       |
                                         XXXXX|                                       |
   6                                   XXXXXXX|                                       |
                                       XXXXXXX|                                       |
   5                                XXXXXXXXXX|13.2                                   |
                                   XXXXXXXXXXX|6.3 11.3                               |
                                XXXXXXXXXXXXXX|20.3                                   |
   4                        XXXXXXXXXXXXXXXXXX|7.3 14.3 18.3                          |
                       XXXXXXXXXXXXXXXXXXXXXXX|10.3 21.3                              |
   3                 XXXXXXXXXXXXXXXXXXXXXXXXX|9.3 19.3 22.2 23.3                     |
                          XXXXXXXXXXXXXXXXXXXX|4.3 11.2 12.4 15.3 17.3                |
   2          XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX|2.2 16.2                               |
          XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX|14.2                                   |
   1          XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX|5.3 20.2                               |
             XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX|11.1 23.2                              |
             XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX|7.2 8.2 13.1 18.2 19.2 22.1            |
   0    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX|3.2 4.2 6.2 10.2                       |
              XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX|1.4 9.2 17.2                           |
  -1  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX|5.2 12.3 15.2 21.2 23.1                |
           XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX|1.3 2.1 4.1 14.1                       |
  -2                 XXXXXXXXXXXXXXXXXXXXXXXXX|19.1                                   |
                     XXXXXXXXXXXXXXXXXXXXXXXXX|1.2 12.2                               |
  -3          XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX|5.1 9.1 10.1 16.1 20.1                 |
                 XXXXXXXXXXXXXXXXXXXXXXXXXXXXX|6.1 7.1 8.1 18.1                       |
                          XXXXXXXXXXXXXXXXXXXX|3.1                                    |
  -4                          XXXXXXXXXXXXXXXX|21.1                                   |
                                  XXXXXXXXXXXX|12.1 15.1                              |
  -5                                XXXXXXXXXX|17.1                                   |
                                       XXXXXXX|1.1                                    |
  -6                                   XXXXXXX|                                       |
                                       XXXXXXX|                                       |
  -7                                        XX|                                       |
                                            XX|                                       |
                                            XX|                                       |
  -8                                         X|                                       |
                                             X|                                       |
  -9                                         X|                                       |
                                             X|                                       |
 -10                                          |                                       |
                                             X|                                       |
=======================================================================================
Each 'X' represents 3.9 cases. The labels for thresholds show the levels of item , and category,
Respectively
```

Figure 3. Variable map of latent distributions and thresholds

In addition to the spread of item difficulties and student abilities shown in Figure 3, the item separation reliability coefficient was 0.99 and the person separation reliability coefficient was 0.97. Together these statistics demonstrated that items were well spread out along the proficiency continuum and were able to adequately discriminate between students of varying ability (Wright & Stone, 1999). However, review of the match between student ability and item difficulty identified that there were insufficient items to represent the numeracy skills of students at the highest and lowest levels of skills. Further development of the item pool will, therefore, include attempts to write new items, or steps within items, to better represent the numeracy skills of these students. The variable map was used to support an understanding of the order of emergence of skills, and this was then interpreted as a learning progression. The progression derived from a review of the map is shown in Figure 4.



8. Learning to apply mathematical operations such as division

The student is starting to apply mathematical operations such as division, and describe outcomes using mathematical language such as 'a quarter of a cup'. S/he may use skip counting in everyday activities and can explain its purpose. With regards to pattern, the student is learning to create and/or explain patterns comprising different objects. S/he is learning to work with large numbers, and may demonstrate understanding of place value by explaining the value of each numeral in four digit numbers. The student is starting to apply mathematical strategies, such as addition and subtraction to solve and work with measuring tapes or rulers to accurately measure length and explain their findings.

7. Learning to apply number concepts to measure time and money

The student is beginning to demonstrate an understanding of how number can be applied to abstract concepts such as ordering and comparing events and objects. S/he may place Australian money in relative order or explain a sequence of events using ordinal number. The student is learning the language of time and may refer to the duration of activities using measures such as minutes and hours. S/he is developing the language to describe comparisons between various sized objects and make meaning from unfamiliar representations of multiple pieces of information, such as a timetable. S/he is beginning to demonstrate strategies (other than recount) to find the total when working with changes in quantity.

6. Learning to recognise shapes and classify groups

The student is learning to represent numbers to 100, which may be demonstrated by placing written two-digit numbers in ascending order or by skip counting objects in twos, fives and tens. S/he may demonstrate their understanding of number conservation by explaining that the total number of objects does not change when their position is rearranged. The student is developing language such as left, right, top and bottom and 2D shapes. S/he may apply rules (e.g., same shape) to organise objects into categories and explain their classifications. Students are developing an understanding of measurement and may use informal measures to compare length. S/he is beginning to identify Australian notes and coins by their value name.

5. Learning to count to 20 and back and building a vocabulary to describe simple numerical concepts

The student is learning to count forwards and backwards from 0 to 20. S/he is developing an ability to extend an alternating pattern and identify the missing element in a number sequence. The student may use simple words to describe a change in quantity, e.g., smaller or bigger. S/he is starting to use language to describe numerical concepts such as chance and order. For example, that an event might or might not occur or that the events were ordered first, second and third. S/he is developing their understanding of the relative duration of events and may demonstrate this by communicating that one event was longer than another.

4. Learning to use number words to count and identify differing magnitudes

The student is learning to attach number words to objects to find the total of a group. S/he is developing the skills to order numerals from 0-10 and to indicate the total of a collection of 1-3 objects without counting (subitising). In working with quantity, the student is beginning to apply counting to check the quantity of a group of concrete materials when the quantity is changed. S/he is starting to communicate their understanding of magnitude by comparing and indicating the larger of two objects. S/he is developing an understanding of directional terms, displayed by following instructions given by another to locate an object.

3. Learning the language of number, shape and relationships

The student is becoming aware of the language of numeracy and beginning to respond to instructions such as being asked to take away or add an item to a group, to respond to use of ordinal numbers and to name familiar 2D shapes. S/he is working with manipulatives and may sort objects into like groups and indicate the longer of two objects when asked. In working with number, the student is beginning to connect number words with up to three objects and to check the quantity of objects in a group if they are rearranged.

2. Responding to numeracy based activities

The student is learning to respond to numerical activities, such as objects being added or taken away from a group. The student may attend to a group of objects being divided into groups and may react to the representation of three objects. The student may rely on support to explore different sized objects and accept guidance to place numerals 1-5 in order. In working with manipulatives, the student is starting to match two like objects from a group of three and sort objects into groups based on an attribute such as colour or size.

1. Reacting to the environment

The student is beginning to react to changes within the familiar environment, and events such as other people counting or the movement of an object. S/he may react to representations of familiar activities or objects or shapes presented by a person.

Figure 4. Derived progression of early numeracy development

Author Name: Jane Strickland
Contact Email: jane.strickland@unimelb.edu.au

DIF analysis

The population of students for whom the assessment was intended consists of students with various types of disabilities that could have an impact on their performance on the items. Since the main goal of DIF analysis at this stage of instrument development was to detect possible item bias, items were investigated for DIF for students with physical disability, vision impairment, hearing impairment, severe language disorder, or autism spectrum disorder (ASD). Presence of DIF was examined by comparing item parameters estimated from responses from one group versus the other. Any difference in item parameters was further investigated for significance and effect size. No difference was found in item performance for students with physical disability, hearing impairment, or severe language disorder. Small DIF with absolute values of 0.396, 0.404, and 0.412 for items 3, 5, and 8 respectively was found for students with vision impairment or ASD. Table 4 shows the summary results of DIF analysis.

Table 4
Summary of DIF analysis

Item	Grouping variable	DIF description	Skill assessed	Criterion	DIF explanation
Item 3	Vision impairment	Small, uniform - item more difficult for students with vision impairment	Recognising quantities of 1-3 without counting (e.g., using words, signs or symbols)	Reacts to (e.g., looks, listens, turns towards) a representation of 1-3 objects	Vision is largely relied upon to subitise. Although haptic (tactile) subitising exists (Plaisier, Bergmann & Kappers, 2009) it is not well described in the criterion.
Item 5	ASD	Small, uniform -item easier for students with ASD	Ordering numbers	ALL	Ability to work with number symbols is commonly a preserved strength in students with ASD (Hiniker, Rosenberg-Lee & Menon, 2015).
Item 8	ASD	Small, uniform - item more difficult for students with ASD	Describing changes in quantity	Responds (e.g., looks, gestures, smiles, protests) when an object is added or removed from the group	Impaired verbal communication is commonly identified in lower functioning students with ASD (Baron-Cohen, Leslie & Frith, 1985).

The presence of items with DIF can have the following unwanted effects: impact on estimates of individual student ability, impact on group ability estimates, and change in the hierarchy of item parameter estimates underpinning a learning progression. Impact on individual ability is calculated as the sum of DIF effects divided by the number of items. It is obvious that for students with ASD the net DIF is close to zero and therefore there is no impact on either individual or group ability estimates. DIF associated with vision impairment is also small and when divided by the number of items (59) it is close to zero. Moreover, the intended use of the assessment instrument is not to compare performance of students with different disabilities but to support their learning through the use of an associated learning progression. In order to examine whether a learning progression derived from the item pool would also be valid for use to describe the learning of students with sub-types of disability such as ASD, for example, item parameters can be compared to the item parameters underpinning the learning progression. If parameter invariance holds, item parameter estimates for students with ASD should be strongly correlated to the item parameters used to develop the learning progression. This can be graphically presented using bi-variate plots. Figure 5 displays a bi-variate plot of the parameter estimates obtained from students with ASD compared to the whole sample. It shows a correlation of 0.99 between parameter estimates for students with ASD and the parameter estimates underpinning the learning progression.

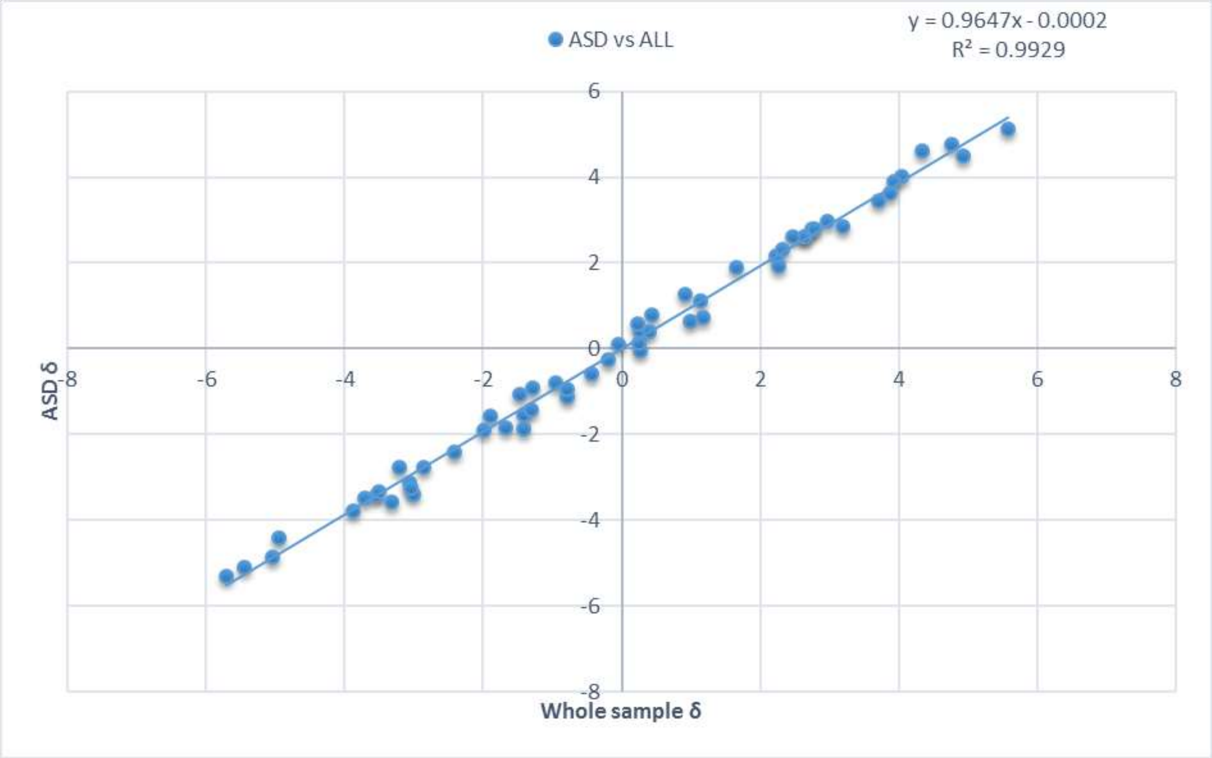


Figure 5. Comparison of item step difficulties obtained by calibrating responses for students with ASD to parameters for the whole sample

In sum, the results of the analysis suggested that the item pool could be meaningfully used to develop a learning progression as the basis of reports for teachers on the numeracy skills and understanding of their students with additional needs. The order of emergence was generalisable to a range of students with different types of disability as the analysis found negligible DIF. Further, the taxonomy used in the item-writing phase of instrument development (as shown in Figure 1 above) was reproduced in the learning progression derived from the analysis (Figure 4). Students moved from noticing and engaging with numeracy activities and phenomena in their everyday activities (Levels 1 and 2 on the derived learning progression), to learning how to use the language of numeracy to describe their experiences and observations (Levels 3 to 5), to using and applying their numeracy skills and understanding to manipulate numbers and solve problems (Levels 6 to 8).

Conclusion

This paper has evidenced the design of a developmental assessment of early numeracy and a derived learning progression intended for students with additional learning needs. These additional needs were viewed more specifically as various disability diagnoses and these groups were used to examine items for potential bias against or towards these groups. The analysis of DIF showed minor bias in individual items, particularly for students with ASD, but an overall extremely high correlation for the larger cohort in the trial. This allows for generalisation of the learning progression to students not assessed in the data set, and may be helpful for teachers to assess and understand early numeracy skills for many students with additional learning needs.

Author Name: Jane Strickland
Contact Email: jane.strickland@unimelb.edu.au

References

Adams, R., & Khoo, S.T. (1996). *Quest*. Melbourne, Australia: Australian Council for Educational Research.

Adams, R., Wu, M., Macaskill, G., Haldane, S., & Xun Sun, X. (2015). *ConQuest*. University of California, Berkley: Australian Council for Educational Research.

Australian Curriculum, Assessment and Reporting Authority (2015*). NAPLAN achievement in reading, persuasive writing, language conventions and numeracy: National report for 2015*, ACARA, Sydney.

Baron-Cohen, S., Leslie, A.M., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition, 21*, 37–46.

Birnie, B. F. (2015). Making the case for differentiation, *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *88*(2), 62-65. doi: 10.1080/00098655.2014.998601

Black, P., Wilson, M., & Yao, S. Y. (2011). Road maps for learning: A guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspective, 9*(2-3), 71-123.

Butterworth, B. (1999). *The mathematical brain*. London: Macmillan.

Commonwealth of Australia. (2016). *Access to real learning: the impact of policy, funding and culture on students with disability, being a report of the Senate Standing Committee on Education*. Canberra: Australian Government Publishing Service. Retrieved 18 January 2016 from http://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Education_and_Employment/students_with_disability/Report.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York, USA: Oxford University Press.

Fuson, K. C., Clements, D. H., & Sarama, J. (2015). Making early math education work for all children: Prekindergarten teachers lay the foundation for later success in mathematics when they attend to the concepts that young children can and should learn, *Phi Delta Kappan*, *97*(3), 63-68.

Gallistel, R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Numerical Cognition, 44*(1-2), 43-74.

Gelman, R., & Gallistel, R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gelman, R., & Butterworth, B. (2005). Number and language: How are they related? *Trends in Cognitive Science*, *9*(1), 6-10.

Hattie, J (2009). *Visible learning*. USA & Canada: Routledge.

Heritage, M. (2008). *Learning progressions: Supporting instruction and formative assessment*. Washington, DC: Council of Chief State School Officers. Retrieved from www.ccsso.org/content/PDFs/FAST%20Learning%20Progressions.pdf.

Hiniker, A., Rosenberg-Lee, M., & Menon, V. (2016). Distinctive role of symbolic number sense in mediating the mathematical abilities of children with autism. *Journal of Autism Developmental Disorders*, *46*, 1268-1281. doi:10.1007/s10803-015-2666-4

Jung, M. (2011). Number relationships in preschool. *Teaching Children Mathematics*, *17*(9), 550-557.

Kleemans, T., Peeters, M. Segers, E., & Verhoeven, L. (2012). Child and home predictors of early numeracy skills in kindergarten. *Early Childhood Research Quarterly*, *27*, 471-477. doi:10.1016/j.ecresq.2011.12.004

LeFevre, J., Skwarchuk, S., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science*, *41*, 55–66. doi:10.1037/a0014532

Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149-174.

Powell, S.R., & Fuchs, L.S. (2012). Early numerical competencies and students with mathematics difficulty. *Focus on Exceptional Children*, *44*(5), 1-16.

Purpura, D. J., & Napoli, A. M. (2015). Early literacy and numeracy: Untangling the relation between specific components. *Mathematical Thinking and Learning: An International Journal*, *17*, 197-218.

Plaisier, M. A., Bergmann Tiest, W. M., & Kappers, A. M. L. (2009). One, two, three, many – Subitising in active touch. *Acta Psychologica*, *131*, 163-170.

Author Name: Jane Strickland
Contact Email: jane.strickland@unimelb.edu.au

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Victorian Department of Education and Training (DET) (2016). ABLES. Retrieved from http://www.education.vic.gov.au/school/teachers/teachingresources/diversity/Pages/ables.aspx.

Vygotsky, L.S. (1929/1993). *The collected works of L. S. Vygotsky, Volume 2: The fundamentals of defectology (abnormal psychology and learning disabilities*). (R. W. Rieber & A. S. Carton, Trans.). New York: Plenum Press.

Wilson, M. (2005). *Constructing measures: An item-response modeling approach*. Mahwah, NJ: Lawrence. Erlbaum Associates.

Woods, K., & Griffin, P. (2013). Judgement-based performance measures of literacy for students with additional needs: Seeing students through the eyes of experienced special education teachers. *Assessment in Education: Principles, Policy & Practice, 20*(3), 325-348.

Wright, B.D., & Stone, M.H. (1999). *Measurement essentials (2nd ed.)*. Wilmington, Delaware: Wide Range Inc.

Author's note

Jane Strickland, Assessment Research Centre, University of Melbourne; Dr Kerry Woods, Assessment Research Centre, University of Melbourne, Masa Pavlovic, Assessment Research Centre, University of Melbourne.

This research was supported by an Australian Research Council Linkage grant with the Victorian Department of Education and Training as industry partner.