Art. #1979, 12 pages, https://doi.org/10.15700/saje.v41n4a1979

Development of an instrument to assess early number concept development in four South African languages

Elizabeth Henning
South Africa DSI-NRF Research Chair, Faculty of Education, University of Johannesburg, Johannesburg, South Africa
ehenning@uj.ac.za
Lars Balzer
Evaluation Unit, Swiss Federal University for Vocational Education and Training SFUVET, Zollikofen, Switzerland
Antje Ehlert
Department of Inclusive Education, Special Educational Needs, University of Potsdam, Potsdam, Germany and University of Johannesburg, Johannesburg, South Africa
Annemarie Fritz
Psychology Department, Faculty of Educational Sciences, Duisburg-Essen University, Duisburg, Germany; Faculty of Education, University of Johannesburg, Johannesburg, South Africa and Akademie Wort + Zahl – International Institute for Intercultural Education, Cologne, Germany

A recently published interview-based test, known by its partly German acronym, MARKO-D SA, is introduced in this article by way of a narrative of its development through various cycles of research. The 48-item test, in 4 South African languages, captures number concept development of children in the 6 to 8-year age group. The authors present their argument for the South African versioning and translation of the test for this country, where there is a dearth of suitable assessment instruments for gauging young children's mathematical concept development. We also present the findings of the research that was conducted to standardise and norm the local version of the test, along with our reasoning about the theoretical strength of the conceptual model that undergirds the test.

Keywords: early grades; early stimulation; isiZulu; MARKO-D SA; number concept development; Rasch-modelling; Sesotho

Introduction: The Genesis of the MARKO-D SA Test

In a search for a trustworthy instrument to capture grade R (kindergarten) children's number concept development, a research group in the Faculty of Education at the University of Johannesburg came across a German interview-based test. In collaboration with the original authors of this test, the team embarked on a process of translation and localisation, which culminated in the publication of the test in South Africa, presenting it in four local languages (Henning, Ehlert, Balzer, Ragpot, Herholdt & Fritz, 2019). Due to the dearth of research in cognitive developmental psychology in South Africa, the research group explored the international literature about numerical cognition to provide a theoretical basis for the test. Despite the public attention to weak mathematics performance of primary school learners, there is little theorising about mathematical cognition and the assessment of children's competence in South Africa. In the *South African Journal of Childhood Education*, which specialises in publishing childhood education research, the bulk of the articles theorise sociocultural and pedagogical aspects of early numeracy learning, with far fewer articles about mathematical, and specifically, numerical cognition (https://sajce.co.za/index.php/sajce).

In this article, we present an overview of the MARKO-D SA instrument as it morphed from a German test into one that was published for use in Sesotho, English, isiZulu and Afrikaans. We present a narrative of the development of the test, beginning with a discussion of the construct that the instrument assesses, namely early number concepts of children that are in transition from pre-school to the early grades.

Conceptual Framework: Numeracy Competence and Diagnostic Testing

The learning of mathematics, specifically numeracy, involves complex processes that begin long before the onset of formal schooling, when children learn to learn symbolically (Ansari, 2008; Bartelet, Vaessen, Blomert & Ansari, 2014; Carey, 2009; Dehaene, 2011; De Smedt, Noël, Gilmore & Ansari, 2013; Spelke & Kinzler, 2007; Wynn, 1992a, 1992b). Children's individual mathematical learning prerequisites differ both in quantity and quality (Aunio & Niemivirta, 2010). While some children come to school, already well prepared for primary school mathematics, others lack important prior exposure to mathematical concepts (Bezuidenhout, 2018). Some of this exposure may be associated primarily with sociocultural factors, which we do not include in this article, as this has been well-theorised in the South African research community. An example of this type of analysis and theorising is apparent in the work of Spaull and Kotze (2015) and the collaborators in the South African Chairs of numeracy learning of the National Research Foundation (Graven & Venkat, 2017).

The relevance of the early acquisition of mathematical competences has been substantiated by several studies that emphasise the importance of early number concept development specifically (Aunio, Korhonen, Ragpot, Törmänen & Henning, 2021; Aunio & Räsänen, 2016; Desoete, 2015). Young children's early numerical knowledge is regarded by many authors to be a strong predictor for later mathematical achievement, while pre-school performance predicts the mathematical performance during primary school (Aunola, Leskinen, Lerkkanen & Nurmi, 2004; Desoete, 2015). This means that children with prior knowledge have a fair chance of successfully engaging with what is offered at school when the transition to symbolic learning (Henning & Ragpot, 2015) begins in earnest (Leibovich & Ansari, 2016). Children with limited prior knowledge and numerical skills are at risk of developing arithmetic difficulties (Aunio & Räsänen, 2016*;* Aunola et al., 2004; Desoete, 2015; Herholdt, 2017; Mononen & Aunio, 2015). According to these and other authors (Baroody, J, Clements & Sarama, 2019; Chinn, 2015; Desoete, 2015), it is crucial to identify children who may start with a backlog as early as possible and to support them in forming number concepts and in developing overall arithmetical competencies. For that, one needs to administer a trustworthy measure.

Identifying competence

A principle of educational-psychological assessment is to identify individual prerequisites and (cognitive) conditions for learning and to design a learning environment that is conducive to that, based on test results. Typically, on the one hand, such testing aims to assess an individual's current competence and to compare that with the performance of peers in a once-off summative assessment. On the other hand, assessment also aims to monitor learning progress, to set learning goals, and to plan an intervention, which can be regarded as formative assessment (Clements & Sarama, 2014).

For the assessment of an individual's learning performance in mathematics, three approaches, based on different theories, can be distinguished. Learners' performance data can be collected via tests that are either 1) curriculum-based, i.e. they solely assess the expected outcomes required by the curriculum, or they 2) focus on basic arithmetical competences that lay the foundation for ensuing mathematical competencies, or 3) they are developmentally focussed, based on a conceptual developmental model that describes the successive acquisition of arithmetical competencies.

Curriculum-based tests

This group of tests includes mainly standardised instruments that assess the content of a given year level curriculum, such as the Annual National Assessments that were, until recently, employed in South Africa. In such tests, the sum score of a learner's performance is compared with the performance of peers in the norming sample. A learner's performance is considered low when the score lies within the 16% of the lowest performing students in the norming sample. This type of traditional assessment is problematic because low-performing learners will only be able to respond to a small number of items. Assumptions about the reasons for these learning difficulties, such as a lack of precursor competencies, or limited conceptual understanding of the reasoning for mathematical operations, cannot be supported and examined, based solely on the data.

Tests on basic arithmetic operations

In another approach, instruments focus on the basic arithmetic operations, such as addition, subtraction, multiplication and division. Such tests are often constructed as speed tests. Learners attempt to respond successfully to as many items as possible in a given period (typically 2–3 minutes), with items becoming increasingly difficult. The time is limited because proficiency is gauged by a test-taker's efficient strategies, as well as confident fact retrieval, culminating in faster performance. Speed components, such as rapid retrieval from memory, are important for the development of mathematical competencies in the primary school years when automatic retrieval is crucial. With these tests, proficient, high performing children can be identified early in their school career because they recall and use stored knowledge efficiently and automatically. Dehaene (2020:223) states that automatization is important as it "frees up the cortex's resources."

However, for low performing learners, the informative value of such tests is limited; no additional information can be gained beyond knowing that the strategies of these learners are not efficient and that they do not retrieve factual and procedural knowledge effectively, or that they do not have such stored knowledge. In early numeracy testing, this means that they still rely on slow counting strategies when solving basic arithmetical problems, which slows down their response.

Developmental tests

Instead of summative assessments that identify an individual's "competency status" at the time of testing, there has been a growing call for assessments that aim to recognise the learning process, and that can continuously monitor a child's learning pathways and growth as formative assessment. On this view, to find the level of a learner's competence with optimal precision, the individual's learning is captured and then described as a process of continuously gaining knowledge and competence. In this sense, learning is viewed as increasing expertise within a certain area, described by developmental psychologists such as Carey (2009) and Spelke (2000) as domain-specific knowledge. Tests that fulfil the requirements of developmental assessment aim to determine the level of proficiency of a specific developmental competency. Due to their theoretical foundation, such tests also fulfil the quality criterion of *prognostic validity,* because they go beyond a sole description of the current skills, to address questions about the next steps within the development as well. An example of a developmental assessment instrument is the MARKO-D test, which is based on a model of early number concept development.

A cognitive model of number concept development

The key assumption of all competence level models is that the acquisition of a certain competency can be described hierarchically, with specific indicators of knowledge that are observable at levels that build on one another. Thus, competencies in a certain learning domain can be understood as a continuum on which different levels of proficiency can be distinguished. Regarding early numeracy learning, it means that basic arithmetical concepts are acquired successively, with increasing conceptual sophistication as learning progresses. Although the notion of the incremental development of early number concepts is not new (e.g., Baroody, AJ & Wilkins, 1999; Carey, 2009; Fuson, 1988; Griffin & Case, 1996; Mix, Huttenlocher & Levine, 2002; Piaget, 1965; Resnick, 1989; Steffe, Cobb & Von Glasersfeld, 1988), a comprehensive empirically validated model, together with a concomitant test, has not yet been developed for this age group. The empirically, cross-sectional and longitudinally validated Model of Development for Arithmetic Concepts (Ricken, Fritz & Balzer, 2013:10) describes the successive acquisition of arithmetic competencies and concomitant conceptual knowledge of children aged 4 to 8.

The five levels of concept development do not refer to encapsulated single (modular) entities, but to overlapping development in a sequence of cognitive development (Siegler & Alibali, 2005). In the MARKO-D manual (Henning et al., 2019) the developmental levels feature in tasks in the number range up to 20.

Level I: Counting

Children’s first experience of natural numbers is coupled with their development of language around the age of 2. In different languages and in different sociocultural settings, children learn number words and soon also the sequence of the words in a number word “line.” At first, number words are non-semantic (Carey, 2009; Fuson, 1988; Le Corre, Van de Walle, Brannon & Carey, 2006) and remain a “list of meaningless lexical items” (Carey, 2009:308), which is sometimes recited in a random sequence. Gradually, knowledge of the number word line stabilises, but children are still not able to utilise number words for counting actions. At first they are not able to select only one object when asked to do so (Wynn, 1992a), even if they are able to recite the number word line up to 10. Gradually, they learn the meaning of the word numerals one, two, three and four. They are now able to meter and enumerate up to four objects, by assigning each object to a number word, and by relying on the strategic alignment of one-to-one correspondence of word and object.

The acquisition of number words for bigger amounts, from five onwards, requires a new “stage” of development, which, according to Le Corre and Carey (2007), also manifests successively and which lasts minimally up to half a year for each number concept to solidify, until the number word “ten” has been solidified semantically. The metering and enumerating of quantities happen on this level only by counting in a fixed succession. Spelke (2017) argues that the verbal knowledge of numerals is the origin of real counting.

Level II: Ordinal number line

A change in the representation of numbers takes place when numbers become associated with the order of successive quantities and how these are represented on what has become known as the “mental number line.” Dehaene (2011) describes the mental number line as “a linear space extending continuously from small to larger numbers” (p. 264), on which “later in the list” is equated to “larger number” – but, that is all. The construction of a linear number line enables children to identify preceding and succeeding numbers. As the numerical quantity along the line becomes progressively larger, the numbers that appear further down on the line represent larger quantities. With the knowledge of the increasing value of the number word line, children begin to understand additive relations and begin to complete addition tasks. However, children who have developed the conceptual knowledge of this level are only able to compute by counting; if they do not form the concept of cardinality, they will remain “counters” only.

Level III: Cardinality

When children grasp the notion of cardinality, they are ready to embark on a numerical development journey. In the “counting out” process of, for example, seven objects, each object will be assigned a number word and the final counting word will capture the whole of the set. Once it is understood that a number is a composite unit that consists of individual elements, it becomes clear that numbers can be decomposed (broken up) as well. A set of seven elements can be partitioned into two subsets in different ways, while the whole quantity of elements does not change.

The cardinality principle in number concept development is the key prerequisite for the acquisition of effective calculation strategies. The addition of 7 + 8 = 15 is no longer a question of counting but can be done by decomposing the numbers adequately (8 is decomposed in 3 and 5), so the addition operation can be completed: 7 + 8 = 7 + 3 + 5 = 10 + 5 = 15.

Level IV: Part-part-whole relations (PPW)

With the development of cardinality, a child understands that each number is composed of various combinations of smaller numbers, so that each number can be decomposed systematically (5 = 5 + 0; 5 = 4 + 1; 5 = 3 + 2, etc.) Fuson (1992:95) considers the relation of subsets and totals as "numerical equivalence", since subsets combined are "equivalent to the sum." In this sense, the relation between the parts and the whole is determined. Individual parts and the whole together form a triadic relationship. For instance, the triad, 7–3–4 can be interpreted as follows: 7 (the whole set) has two subsets, consisting of 3 and 4 elements. The subsets of 3 and 4 together are equivalent to the whole 7 (3 + 4 = 7). If two quantities are known, the third one can be deduced, no matter which part is missing.

Level V: Equidistant number line intervals

Based on their cardinal knowledge of different numbers/quantities, children begin to realise that successive numerals "one", "two", "three", and "four" refer to sets that are related by +1. This realisation is coupled with an understanding that the magnitude of the difference between the numbers is always the same. Hence, in the number line representation, the distance between any two consecutive numbers is always an equal distance. With this knowledge, children have a type of scale at their disposal, which enables them to determine differences between two sets exactly.

This also means that distances of equal magnitude on the number line have the same cardinality. The child now understands that the distance between zero and five is equivalent to the distance between five and 10. The concept of the structured number line is a prerequisite for the understanding of multiplicative relationships as well as the concept of numerical place value.

This description of the five conceptual levels underlines that each level is characterised by a specific concept, which builds onto the previous concepts and prerequisites.

Assessing number concept development

The aim of the construction of the original MARKO-D in Germany was to locate a child's performance on one of the hierarchically sequenced levels by means of the test.

Two questions had to be addressed:

1) How can one be sure that the tasks truly operationalise the specific concept of each level?
2) How can one be sure that the test captures children's understanding of the five different number concepts reliably?

The researchers in Germany created an item pool according to the theoretical principles of Levels I to V with a total of 70 items. In several pilot studies they trialled the items to establish their empirical fidelity and then modified them where necessary. Using item response theory (IRT) as tool, they set out to determine whether all the items formed a one-dimensional cumulative scale, whether items on specific segments of the scale could be identified, and whether the sequence of these segments on the scale followed the sequence of levels in the model. Altogether, the researchers tested more than 3,000 children during the process of conducting several pilot studies. This led to the design of the MARKO-D test instrument (Ricken et al., 2013) with 55 items.

Based on the German MARKO-D, the test was translated for use in South Africa. We translated and adapted the test culturally for children in this country. The main aim was to confirm the validity of the test, to ensure that the instrument continued to assess the same constructs as in its original language – thus, whether the translated items retained the conceptual content of the original test. The first translation was from German to English. Children from three different age groups formed part of this first pilot; the groups comprised children from Grade R, Grade 1 and Grade 2 (n = 224). The English version of the test was then translated into three other South African languages, namely isiZulu, Afrikaans and Sesotho. The translations went through five iterations each during pilot studies. In the end, we were able to prove the validity of the model in all four languages. We had to remove some items, which were too easy or too difficult in one of the languages.

The first set of pilot studies in all four languages was then conducted. The English and Afrikaans pilot studies showed similar results; the order of the items in these two languages corresponded to the order of the items in the German test. In the other two languages we conducted a couple of back-and-forth translations with guidance from African language linguists and assistance from a team of teachers.

The final test contains 48 items. The items form a one-dimensional cumulative scale with five distinguishable segments according to the levels of the theoretical model. In total, each segment (or level) includes items of the respective level. Thus, the validity of the model and the test can be considered as proven in all four languages, which means that the model and the test are culture and language independent (See examples of items in Figure 1).

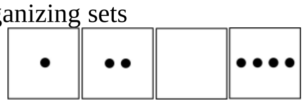
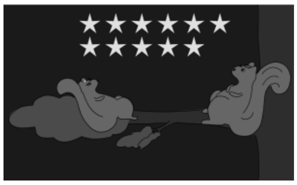
Level	Task	Picture	Instruction
I	Counting of sets		Sets with 6 and 9 nuts: "How many are there?"
II	Finding preceding and succeeding numbers		"What number comes before/after 5?" "What number is between 2 and 4 (5 and 7)?"
III	Organizing sets 		"Have a look at the square. How many chips go into this empty square? Put that many chips into the square!"
IV			"Give me 5 chips, 3 of them should be red!"
V	Recognising differences between sets		"Which row has more? How many more are there?"

Figure 1 Examples of items for each level of the model[i]

Special mention should also be made of the use of characters in a narrative that forms the backdrop of the test. In the original German version, squirrels were inserted as characters. This was not suitable for South African children. Therefore, the characters were changed to meerkats and the drawings made accordingly. For the drawings, the idea was to keep the illustrations non-invasive and minimalistic in terms of colour and background to not overload working memory.

Method and Results

Sample

All psychometrical and statistical information are based on one large sample for modelling and a representative subsample for norming purposes. The numbers for the population of interest were extracted from the 2014 Annual National Assessment database compiled by the Gauteng Department of Education. The total number of Grade 1 learners taught in Afrikaans, English, Sesotho and isiZulu in Gauteng was 175,033. The EMIS dataset of 2014 was used to source school background data that was missing from the Annual National Assessment database.

The population can be described as the total number of Grade 1 learners in Gauteng who were taught in their home languages (HL) of Sesotho, Afrikaans, English isiZulu, as well as children taught in English, but who were not first language English speakers – also referred to as English first additional language (EFAL) speakers in the South African education system, or English language learners (ELLs). The total number of learners in the test population was 163,226. The distribution in the population was as follows: 9.50% of Grade 1 learners were first language (L1) speakers of Afrikaans, 17.40% of isiZulu, 9.02% of Sesotho and 14.34% of English. The second language (L2) English speakers represented 49.73% of the population.

After discussion by the test developers (Henning et al., 2019), the decision was made to use school fees as a proxy for socio-economic factors. Table 1 shows that "no-fee" schools are attended by a substantial number of African language speakers in the first grade. For the English and Afrikaans group, the opposite pattern is observed.

Table 1 School fees and language groups

Number of cases in population and percentages (in brackets) attending fee or no-fee schools per language			
Language	Fee	No fee	Total
Afrikaans	15,135 (97.58)	376 (2.42)	15,511
isiZulu	720 (2.53)	27,688 (97.47)	28,408
Sesotho	416 (2.82)	14,315 (97.18)	14,731
English HL	21,694 (92.67)	1,716 (7.33)	23,410
English FAL	51,030 (62.87)	30,136 (37.13)	81,166
Total	88,995	74,231	163,226

In 2014 and 2015, 1,200 Grade 1 learners were randomly selected from conveniently sampled schools from the south west, south east, south and central districts of Gauteng. These learners were tested in February and early March of their Grade 1 year. One thousand one hundred and eighty six children from this sample completed the test. This sample (Table 2) was further used for testing the model (we refer to this as the modelling sample).

Table 2 Modelling sample

Number and percentage (in brackets) of cases in modelling sample			
Language	Fee	No fee	Total
Afrikaans	192 (95.05)	10 (4.95)	202
isiZulu	37 (15.23)	206 (84.77)	243
Sesotho	29 (12.50)	203 (87.50)	232
English HL	204 (95.33)	10 (4.67)	214
English FAL	198 (64.08)	111 (35.92)	309
Total	660	540	1,200

In order to give representative information about number concept development and its measured competencies, it is important to depict the proportionally represented population regarding language and fee/no-fee schools as described above. We randomly selected 602 learners, taking the language and fee/no-fee distributions into account, ending up with a representative sample regarding these criteria (Table 3).

Table 3 Language and school fees

Norming sample size per language and per fee/no-fee						
			Fee		No-fee	
	Number of cases in the sample	% in sample	Number of cases in sample	% in language group	Number of cases in sample	% in language group
Afrikaans	57	9.47	56	98.25	1	1.75
isiZulu	105	17.44	3	2.86	102	97.14
Sesotho	55	9.13	2	3.64	53	96.36
English HL	86	14.29	80	93.02	6	6.98
English FAL	299	49.67	188	62.88	111	37.12
Total	602		329	54.65	273	45.35

Three hundred and twenty two (53.5%) were female and 280 (46.5%) were male. The age at the time of testing ranged from 62 to 102 months, with a mean of 77.5 months (*SD* = 6.4).

Measurement Model

When developing a test, it is important to examine assumptions about the relationship between the performance captured by the test and the latent person ability, which is of interest, but not directly observable. The MARKO-D SA test comprises five theoretical levels, which have been operationalised with items assigned to the respective levels (See Figure 1). To empirically prove items and to check the validity of a "level" concept, one can use IRT, and if the construct in question is one-dimensional (for example, early numerical competence), the One-dimensional (1PL) Rasch Model is appropriate.

It would be beyond the scope of this text to describe Rasch modelling (Wright & Linacre, 1994) in detail, but our procedure should be explained. A common way of working in test development is to start with an extensive set of items and to conduct various empirical tests to prove empirical fit. Often, most of the initial items are omitted during this process, due to insufficient *fit statistics*. As there is no possibility to develop and modify items continuously, non-fitting items are removed, provided the remaining item-pool remains large enough. This normally results in a test with strong statistical rigour, and with competence levels that can be described, because the content of the items fit a particular level.

For the MARKO-D SA, we used Rasch modelling differently, starting with a strong theory, which included a stringent notion of competence levels and of competencies required for each level, so that corresponding items could be built and tested empirically. Sufficiently fitting statistics, as well as an appropriate location on the levels on the scale, were needed for each of the items to be included in the final test. If one, or both, of the conditions were not met, corresponding items were analysed in detail, asking questions such as: What might be wrong with this item? Is there a wording problem? Might it be possible to understand the item in a different way to what the test designers had intended? Why is the task in the item too easy or too hard to be completed, compared with the competence level to which it should belong? This procedure often results in a modification of the wording of an item, coupled with repeated empirical testing. The outcome of such questioning is usually a test with a strong connection to theoretical assumptions and is, therefore, suitable for robust interpretation of test scores. Figure 2 shows an item-person map based on the data of the modelling sample.

```
                  PERSON - MAP - ITEM
                     <more>|<rare>
          6                +
                        .  |
                           |
                           |
                           |
          5             .  +
                           |
                           |
                        .  |T 21
                           |  23
          4             .  +
Level 5                    |
                        .  |  42
                        .  |  30
                           |  25
          3            .#  +
                       .# T|
                       .#  |  29    31   35
                       .#  |  32
                       ##  |S
-----------------------------------------------------------
          2           .##  +  41
Level 4               ###  |  28    40
                     .###  |  26
                .######### S|  39
-----------------------------------------------------------
                     ####  |  48
          1         .####  +  27    44
Level 3            .#####  |
              .###########  |  33    47
                  .######  |  11    13    15   16
-----------------------------------------------------------
                 .#######  |  14    34    4
          0 .############ M+M 9     3     43
Level 2            .#####  |
                   ######  |  12    2
                 .#######  |  10
                 .#######  |  8
         -1 .############  +
                   .#####  |
                  .###### S|
-----------------------------------------------------------
                    .####  |  46
                      .##  |
         -2           ###  +  22     45
                       .#  |S
                        .  |  18     24   5
                        .  |  20     37
Level 1                   T|
         -3            .#  +  36     38
                        .  |  7
                        .  |  17     19   6
                           |
                        .  |
         -4                +
                           |
                        .  |T 1
                           |
                           |
         -5                +
                     <less>|<freq>
           EACH "#" IS 8: EACH "." IS 1 TO 7
```

Figure 2 Person-item map of the whole sample

In this map, [#] represents eight persons and [.] represents one to seven persons; [M], [S] and [T] (left of the line for subjects and right of the line for items) represent the mean [M], as well as one [S] or two [T] standard deviations from the mean of the respective distribution. The mean of the scale is fixed at 0 logit. The unit of measurement on this item-person map, which represents both the item difficulty and person ability, is defined as a logit ("log odds unit"). Measures are expressed on the logit scale with the average item measure or person ability arbitrarily set to 0. On the common interval-scaled ability/difficulty scale, the score distribution of the children is indicated on the left (person ability) and the position of the individual items on the scale on the right (item difficulty). The higher a child's position, the higher her or his ability, and the higher the position of an item, the more difficult it is. The higher the position of a child, compared to that of a given item, the higher the probability that this child will solve the item correctly.

The items of the One-dimensional Dichotomous Rasch Model show satisfactory values (weighted infit MNSQ 1 ± 0.2 for 45 out 48 items; weighted infit MNSQ 1 ± 0.3 for 3 items) and the person reliability is at .90. Seven items from the original German MARKO-D have been omitted due to unsatisfactory item fit. The validity requirements for a One-dimensional Rasch Model can be considered as fulfilled for a 48-item version of the MARKO-D SA: the items form a one-dimensional cumulative scale. The item-person map indicates that the items cover the entire range of numerical concepts described in the developmental model. Children's range of conceptual numerical abilities is covered appropriately as well. The horizontal boundary lines between levels were added by allocating the items to the respective levels for which they had been constructed, based on the model. The grouping of items according to model levels was successful. Results show that the empirical allocation of almost all items corresponds to the theoretically predicted proficiency levels. Thus, segments on the scale can be identified that include items which can be solved based on one of the five numerical concepts proposed in the model, and the sequence of these segments on the scale follows the sequence of levels in the model.

Compared to the original German test, in total, seven items had to be removed for the South African version. This had to be done because the operationalisation of these items did not hold in all four of the South African languages. Only a few items differed in their level allocation from the German test; keeping in mind the probabilistic test model, which is acceptable.

Norming

The following table shows how children of the norming sample performed. This allows representative information about the competence level of Grade 1 pupils (for more details and interpretation of results, see Henning et al., 2019).

Table 4 Competence level achieved

	Frequency	Percent
Level 1	72	12.0
Level 2	274	45.5
Level 3	170	28.2
Level 4	53	8.8
Level 5	33	5.5
Total	602	100.0

The individual performance can be interpreted by the comparison to the data of the norming sample. In the manual (Henning et al., 2019), we present a norming table which allows comparing the test score and the corresponding conceptual level of an individual child with the levels of a representative sample.

Research Ethics

The manuscript is the result of a number of research projects, which have been granted ethical clearance by the Faculty of Education of the University of Johannesburg, under the permission umbrella of the *Institute of Childhood Education* at the time when this research was conducted.

Discussion

In the last few decades, the importance of early acquisition of mathematical competences has been emphasised by many authors in different fields, including cognitive developmental psychology, mathematical cognition and mathematics education. Researchers agree that there is a need to assess mathematical competencies as early as possible in order to identify those children who do not have the necessary prerequisites to cope with the school curriculum. There is also a drive to advance early mathematics as part of the social capital of an emerging economy in a developing country (JET Education Services & Kelello Consulting, 2018). With the MARKO-D SA, we present a test for Grade R and Grade 1 learners that is not curriculum-based, but developmentally oriented, according to a model that describes the successive acquisition of early arithmetical competencies.

As a first step, we had to show that the MARKO-D test, which was developed in Germany, was also valid in South Africa. The main aim was to ensure that the instrument assesses the same constructs as in its original language and that the translated items retained the conceptual content of the original test.

Based on the data of the research for this test, the validity requirements for a One-dimensional Rasch Model can be considered as fulfilled for the 48-item version of the MARKO-D SA. This means that the MARKO-D SA test can be used in South Africa, to identify children's levels of proficiency. When entering primary school, children's number concept needs to be on, or close to, Level III of the model presented in this article. Level III competence means conceptual understanding of cardinality, which is a prerequisite for moving beyond simple counting strategies to perform calculation tasks, and for the understanding of all further arithmetical operations.

Children who had only reached Level I (12% in our study) and were only able to count and enumerate small sets, were at risk and urgently need support. Those who had only reached Level II (45.5%), were in the "danger zone" of cementing counting as computational skill and making it a habit. According to Dehaene (2011), an understanding of numerical cardinality is neither innate, nor does it "emerge" naturally, but requires instruction. Therefore, schooling in Grade 1 should encourage children to complete arithmetic tasks with understanding, using strategies that are based on conceptual knowledge of number.

Due to its theoretical foundation, the MARKO-D SA test also fulfils the quality criterion of *prognostic validity*, because it goes beyond a sole description of the current skills, but addresses questions about the next steps of concept development as well. The test can, firstly, be used to assess the knowledge of children when they start school. Our norming sample, as well as a recent study by Bezuidenhout (2018), found that the young learners were not yet ready for the challenges of the fast-paced Grade 1 curriculum in South Africa.

The MARKO-D SA, furthermore, can be used by learning support specialists to assess children with possible mathematical learning difficulties, which place them at risk. It has also been shown to be suitable to evaluate a programme for teacher development (JET Education Services & Kelello Consulting, 2018) and to assess Grade R learners' progress over 1 school year (Bezuidenhout, 2018). In a country that has a shortage of mathematically literate citizens who will be eligible for the Fourth Industrial Revolution (4IR) employment market, early diagnosis of competence can contribute to improved teaching and improved learner outcomes. To date, primary school learning of mathematics has been identified as one of the weak learning outcomes (Radebe, 2017; Spaull & Kotze, 2015). For education, including curriculum adjustment and pedagogy, there are implications to consider with regard to laying the foundations of mathematical competence.

Educational Implications

Mathematical knowledge builds systematically on existing knowledge. By validating the underlying model of the MARKO-D, we have shown that children successively build conceptual knowledge and related strategies incrementally. They acquire competencies increasingly by connecting new knowledge with existing knowledge – the basis of constructivism as an epistemology. Dehaene (2020:221) argues for "consolidation" in this regard. Expertise, or competences, are not simply constructed by a quantitative increase in knowledge, but, above all, by knowledge that is qualitatively organised differently through new learning experiences, thus becoming richer, more flexible, more effective, and, importantly, consolidated. This newly acquired expertise in turn forms the basis for the acquisition of further expertise or knowledge in a specific domain.

Six decades ago, Ausubel (1968) emphasised the importance of previous (existing) knowledge for the acquisition of new knowledge. According to Ausubel, teaching processes should, therefore, do justice to this principle: "The most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly" (Ausubel, 1968:vi). This means that, in order to enable cumulative learning, lessons should be prepared according to known, hierarchically structured learning content. Even if there is no systematic instruction according to some valid model, learners acquire some kind of knowledge and skills in each individual learning experience. However, if they are unable to relate the newly acquired knowledge to existing knowledge, it remains "inert knowledge", according to Renkl, Mandl and Gruber (1996:118–119) and forms no recognisable basis for building further competence, unless the inertness is "attacked" (*sic*). Many of the gaps that are evident in weak mathematics performance of South African learners may be due to missing *building blocks* – the term that Clements and Sarama (2014) and Sarama and Clements (2009) use to describe gaps in learning pathways in early mathematics. New knowledge is not easily recognisable if there is limited connection to what a learner already knows. Such a learner then builds knowledge only superficially and with little lasting effect. Against this background, arithmetic difficulties arise: Children with arithmetic difficulties have not yet taken certain learning steps, which teachers (and the curriculum – see Fritz, Long, Herzog, Balzer, Ehlert & Henning, 2020) rely on in classroom learning. Such fragmentary basic knowledge impedes any further learning. Children who start school with gaps and who do not acquire viable basic knowledge during the first years of school face a widening gap between their abilities and those of their peers or the learning objectives of their grade.

Early education seems to be the answer to the pressing questions. In recent years, a growing body of research, in particular from neuroscience, sociology and psychology has proven that early childhood education and care provides a crucial foundation for future learning, upon which learning in schools can build and which is important for success later in life. However, for early education to develop effectively and sustainably and prepare children well for school, pre-school education must also follow a curriculum that is based on research about this age group.

The MARKO-D SA is registered as an educational assessment instrument with the intention that it can be used by teachers and other educational practitioners to record the current level of knowledge of children and to help teachers to purposefully adapt their teaching to the children's prior knowledge. It is by no means reserved for the clinical psychology or the educational psychology practitioner only.

Acknowledgements

The National Research Foundation (NRF Competitive Rated Researcher grant (Grant number 85344). The DSI-NRF South Africa Research Chair of E Henning: *Integrated studies of learning language mathematics and science in the primary school*. (Grant number 98573).
The Zenex Foundation, Johannesburg.

Authors' Contributions

Elizabeth Henning compiled all the authors' contributions for the paper and wrote the final manuscript. Lars Balzer was responsible for the statistical and modelling work and texts. All authors reviewed the final manuscript.

Notes

i. The figures and tables in this paper are cited from the MARKO-D SA Manual, originating from the joint work of the authors of the South African manual (Henning et al., 2019) and the German manual (Ricken et al., 2013). They are published with permission from the authors.

iii. DATES: Received: 13 October 2019; Revised: 11 June 2020; Accepted: 13 August 2020; Published: 30 November 2021.

References

Ansari D 2008. Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience*, 9(4):278–291. https://doi.org/10.1038/nrn2334

Aunio P, Korhonen J, Ragpot L, Törmänen M & Henning E 2021. An early numeracy intervention for first-graders at risk for mathematical learning difficulties. *Early Childhood Research Quarterly*, 55:252–262. https://doi.org/10.1016/j.ecresq.2020.12.002

Aunio P & Niemivirta M 2010. Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences*, 20(5):427–435. https://doi.org/10.1016/j.lindif.2010.06.003

Aunio P & Räsänen P 2016. Core numerical skills for learning mathematics in children aged five to eight years – a working model for educators. *European Early Childhood Education Research Journal*, 24(5):684–704. https://doi.org/10.1080/1350293X.2014.996424

Aunola K, Leskinen E, Lerkkanen MK & Nurmi JE 2004. Developmental dynamics of math performance from preschool to Grade 2. *Journal of Educational Psychology*, 96(4):699–713. https://doi.org/10.1037/0022-0663.96.4.699

Ausubel DP 1968. *Educational psychology: A cognitive view*. New York, NY: Holt, Rinehart & Winston.

Baroody AJ & Wilkins JLM 1999. The development of informal counting, number, and arithmetic skills and concepts. In JV Copley (ed). *Mathematics in the early years*. Reston, VA: National Association for the Education of Young Children.

Baroody J, Clements DH & Sarama J 2019. Teaching and learning mathematics in early childhood programs. In CP Brown, MB McMullen & N File (eds). *The Wiley handbook of early childhood care and education*. Hoboken, NJ: Wiley & Sons.

Bartelet D, Vaessen A, Blomert L & Ansari D 2014. What basic number processing measures in kindergarten explain unique variability in first-grade arithmetic proficiency? *Journal of Experimental Child Psychology*, 117:12–28. https://doi.org/10.1016/j.jecp.2013.08.010

Bezuidenhout HS 2018. Input for young children's number concept development. PhD thesis. Johannesburg, South Africa: University of Johannesburg.

Carey S 2009. *The origin of concepts*. Oxford, England: Oxford University Press.

Chinn S (ed.) 2015. *The Routledge international handbook of dyscalculia and mathematical learning difficulties*. London, England: Routledge.

Clements DH & Sarama J 2014. The importance of the early years. In RE Slavin (ed). *Science, technology and mathematics (STEM)*. Thousand Oaks, CA: Corwin.

Dehaene S 2011. *The number sense. How the mind creates Mathematics* (2nd ed). New York, NY: Oxford University Press.

Dehaene S 2020. *How we learn: The new science of education and the brain*. London, England: Penguin Books.

De Smedt B, Noël MP, Gilmore C & Ansari D 2013. How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, 2(2):48–55. https://doi.org/10.1016/j.tine.2013.06.001

Desoete A 2015. Predictive indicators for mathematical learning disabilities/dyscalculia in kindergarten children. In S Chinn (ed). *The Routledge international handbook of dyscalculia mathematical learning difficulties*. London, England: Routledge.

Fritz A, Long C, Herzog M, Balzer L, Ehlert A & Henning E 2020. Mismatch of the South African foundation phase curriculum demands and learners' current knowledge. *African Journal of Research in Mathematics, Science and Technology Education*, 24(1):10–20. https://doi.org/10.1080/18117295.2020.1724466

Fuson KC 1988. *Children's counting and concepts of number*. New York, NY: Springer. https://doi.org/10.1007/978-1-4612-3754-9

Fuson KC 1992. Research on learning and teaching addition and subtraction in whole numbers. In G Leinhardt, R Putnam & RA Hattrup (eds). *Analysis of arithmetic for mathematics teaching*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Graven M & Venkat H (eds.) 2017. *Improving primary mathematics education, teaching and learning: Research for development in resource-constrained contexts*. London, England: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-52980-0

Griffin S & Case R 1996. *Number worlds: Kindergarten level*. Durham, NH: Number World Alliance.

Henning E, Ehlert A, Balzer L, Ragpot L, Herholdt R & Fritz A 2019. *MARKO-D - South-Africa. Assessment of number concept development*. Johannesburg, South Africa: University of Johannesburg.

Henning E & Ragpot L 2015. Pre-school children's bridge to symbolic knowledge: First literature framework for a learning and cognition lab at a South African university. *South African Journal of Psychology*, 45(1):71–80. https://doi.org/10.1177/0081246314545199

Herholdt R 2017. The utility of the South African Marko-D test with regard to identifying grade 1 English second language speakers with mathematics difficulties. MEd dissertation. Johannesburg, South Africa: University of Johannesburg. Available at https://ujcontent.uj.ac.za/vital/%20access/manager/Repository/uj:25403/SOURCE1. Accessed 24 September 2021.

JET Education Services & Kelello Consulting 2018. *Evaluation of the Grade R mathematics project implemented in the Western Cape* (Endline report). Johannesburg, South Africa: JET Education Services.

Le Corre M & Carey S 2007. One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, 105(2):395–438. https://doi.org/10.1016/j.cognition.2006.10.005

Le Corre M, Van de Walle G, Brannon EM & Carey S 2006. Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology*, 52(2):130–169. https://doi.org/10.1016/j.cogpsych.2005.07.002

Leibovich T & Ansari D 2016. The symbol-grounding problem in numerical cognition: A review of theory, evidence, and outstanding questions. *Canadian Journal of Experimental Psychology*, 70(1):12–23. https://doi.org/10.1037/cep0000070

Mix KS, Huttenlocher J & Levine SC 2002. *Quantitative development in infancy and early childhood*. Oxford, England: University Press.

Mononen R & Aunio P 2015. Counting skills intervention for low-performing first graders. *South African Journal of Childhood Education*, 6(1):a407. https://doi.org/10.4102/sajce.v6i1.407

Piaget J 1965. *The child's conception of number*. New York, NY: Norton.

Radebe WF 2017. How early grade teachers identify children with mathematics learning difficulties. Master's dissertation. Johannesburg, South Africa: University of Johannesburg.

Renkl A, Mandl H & Gruber H 1996. Inert knowledge: Analyses and remedies. *Educational Psychologist*, 31(2):115–121. https://doi.org/10.1207/s15326985ep3102_3

Resnick LB 1989. Developing mathematical knowledge. *American Psychologist*, 44(2):162–169. https://doi.org/10.1037/0003-066X.44.2.162

Ricken G, Fritz A & Balzer L 2013. *MARKO – D: Mathematik- und rechenkonzepte im vorschulalter – diagnose; Manual* [Diagnostic test for mathematical and arithmetical concepts in the pre-school age-group; Manual]. Göttingen, Germany: Hogrefe.

Sarama J & Clements DH 2009. *Early childhood mathematics education research: Learning trajectories for young children*. London, England: Routledge.

Siegler RS & Alibali MW 2005. *Children's thinking* (4th ed). Englewood Cliffs, NJ: Prentice Hall.

Spaull N & Kotze J 2015. Starting behind and staying behind in South Africa: The case of insurmountable learning deficits in mathematics. *International Journal of Educational Development*, 41:13–24. https://doi.org/10.1016/j.ijedudev.2015.01.002

Spelke ES 2000. Core knowledge. *American Psychologist*, 55(11):1233–1243.

Spelke ES 2017. Core knowledge, language, and number. *Language Learning and Development*, 13(2):147–170. https://doi.org/10.1080/15475441.2016.1263572

Spelke ES & Kinzler KD 2007. Core knowledge. *Developmental Science*, 10(1):89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Steffe LP, Cobb P & Von Glasersfeld E 1988. *Construction of arithmetical meanings and strategies*. New York, NY: Springer-Verlag.

Wright BD & Linacre JM 1994. Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8:370–371.

Wynn K 1992a. Addition and subtraction by human infants. *Nature*, 358:749–750. https://doi.org/10.1038/358749a0

Wynn K 1992b. Children's acquisition of the number words and the counting system. *Cognitive Psychology*, 24(2):220–251. https://doi.org/10.1016/0010-0285(92)90008-P