

RESEARCH ARTICLE

The development of a South African Early Learning Outcomes Measure: A South African instrument for measuring early learning program outcomes

Matthew Snelling¹  | Andrew Dawes^{2,3} | Linda Biersteker⁴ | Elizabeth Girdwood⁴ | Colin Tredoux^{2,5}

¹Alacrity Development, Cape Town, South Africa

²Department of Psychology, University of Cape Town, South Africa

³Young Lives, Department of International Development, University of Oxford, Oxford, England

⁴Innovation Edge, Cape Town, South Africa

⁵Department of Psychology Université de Toulouse, France

Correspondence

Matthew Snelling (MPhil), Alacrity Development, Cape Town, South Africa.
Email: matthew@alacritydevelopment.org

Funding information

Innovation Edge

Abstract

Background: Assessment of early childhood development programme effectiveness in South Africa is hampered by a lack of suitable measures that account for variations in cultural and socio-economic backgrounds and can be administered by non-professionals. This contribution reports the standardisation of the South African Early Learning Outcomes Measure (ELOM), an instrument designed for population level monitoring of the developmental status of children aged 50–69 months and for evaluation of early learning programmes.

Methods: The development of the ELOM was informed by South Africa's National Curriculum Framework from Birth to Four and its National Early Learning and Development Standards. ELOM items were drawn from reliable and valid instruments, particularly those used in Africa and other developing regions and were clustered in five domains: gross motor development, fine motor coordination and visual motor integration, emergent numeracy and mathematics, cognition and executive functioning, emergent literacy and language. The ELOM was standardised on a sample of 1,331 children aged 50–69 months, from five South African official languages and five socio-economic strata. Item Response Theory techniques were used to establish reliability, validity, and differential item functioning.

Results: Confirmatory Factor Analysis established that ELOM domains are unidimensional and internally consistent. Items discriminate reliably between more and less able children and do not discriminate unfairly between children of the same ability from different language backgrounds. Socio-economic gradients were evident in children's performance. South African Early Learning Development Standards (ELDS) based on standard scores were developed and set at the 60th percentile of the sample standard score distribution.

Conclusions: This research produced the first South African, age-validated population-level standardised instrument that can be administered relatively cheaply by trained non-professionals. This will facilitate the assessment of the efficacy of early learning programmes in enabling children to reach ELDS prior to entering Grade R and track progress toward Sustainable Development Goal 4.2.

KEYWORDS

Africa, early assessment, early learning, early learning program, outcome measure

1 | INTRODUCTION

The development of the Early Learning Outcomes Measure (ELOM) was prompted by the lack of an instrument aligned to South African Early Learning Development Standards (ELDS; UNICEF, 2016), that takes into account the cultural and economic background of the population, that is, inexpensive and can be administered by trained pre-school personnel, rather than professionals. The need for an instrument of this type is further prompted by developments in South African early childhood policy that recognise the critical importance of quality early education, prior to entering formal school (Republic of South Africa, 2015, p. 59; Berry, Dawes, & Biersteker, 2013). This goal is aligned in purpose and timing with United Nations Sustainable Development Goal (SDG) 4.2, which seeks to ensure access to quality early learning for all children by 2030.

Research on the ingredients of effective early learning programmes has been undertaken for many years (Bailey, Duncan, Odgers, & Yu, 2017; Duncan et al., 2007; Heckman, Pinto, & Saveliev, 2013). Although there is less evidence from the Global South than that available in developed economies, the importance of ensuring programme effectiveness is receiving increasing attention in societies with more limited resources (Aboud, 2006; Britto, Yoshikawa, & Boller, 2011; Dua et al., 2016; Engle et al., 2011; Kagitcibasi, Sunar, Bekman, Baydar, & Cemalcilar, 2009; Mwaura, Sylva, & Malmberg, 2008; Richter et al., 2017; Walker, Chang, Powell, & Grantham-McGregor, 2005).

Reliable, standardised, context appropriate instruments for measuring early learning programme outcomes and for monitoring whether children are developmentally “on track” prior to school, are required if we are to measure progress toward these goals and improve programme delivery. These need to be sensitive to local cultural and socio-economic variations. This is an important consideration for South Africa, a country with multiple ethnolinguistic communities and significant socio-economic disparities and where programmes for the most disadvantaged children tend to be of poorest quality (Biersteker, Dawes, Hendricks, & Tredoux, 2016). Several international population-based early learning outcome instruments that serve this purpose have appeared in recent years. One of the first widely used is the Early Development Index, in which teachers rate the developmental status of children (Fernald, Kariger, Engle, & Raikes, 2009; Janus, 2007). Save the Children's International Development and Early Learning Assessment (IDELA) has been tested in more than fourteen low and middle-income countries and assesses gross and fine motor development, social and emotional functioning, literacy, mathematics, and aspects of executive functioning in children 42 to 72 months (Dowd, Borisova, Amente, & Yenew, 2016; Pisani, Borisova, & Dowd, 2015). Measure of development and early learning module of the Measuring Early-Learning Development Outcomes (MELQO; early

childhood development [ECD] measure, 2018) has been developed for monitoring progress toward achievement of SDG 4.2 (access to quality education). Other instruments, largely covering the same domains, include the East Asia-Pacific Early Child Development Scales (Rao et al., 2014a) and the Tongan Early Human Capability Index (Brinkman & Vu, 2016).

One key challenge for test development in developing countries including South Africa, is the variety of cultures, traditions of child rearing, and the breadth of languages that must be accommodated (Serpell & Simatende, 2016). The cross-cultural developmental literature has engaged these problems for many years (e.g., Cole & Bruner, 1971; Dasen, 1984). Further, Mwaura and Marfo (2011) note (East African) “dependence on imported instruments, often adopted with little or no adaptations” (p. 138). This situation is changing as a number of instruments have been developed or adapted for African populations. We cannot be exhaustive here, but recent examples include the Ages and Stages Questionnaire, adapted for Zambia and South Africa (Hsiao et al., 2016). Other Zambian examples include a non-verbal reasoning test for rural children (Zuilkowski, McCoy, Serpell, Matafwali, & Fink, 2016), the Panga Munthu Test, and the Zambia Child Assessment Tool (Matafwali & Serpell, 2014).

1.1 | The standardisation of the South African ELOM

This contribution reports on the standardisation of the ELOM: a low-cost instrument designed to measure the extent to which South African children, aged 50 to 69 months are developmentally “on track” by the time they enter Grade R, the non-compulsory reception year of school. To be eligible, children must be 60 months by June of the Grade R year.

2 | METHODS AND MATERIALS

The development of the ELOM is described in Dawes, Biersteker, Girdwood, Snelling, and Tredoux (2019). Here, we restrict ourselves to its psychometry and standardisation.

ELOM design was underpinned by the ELDS literature (e.g., Kagan, Castillo, Gomez, & Gowani, 2015). Indicators were informed by South Africa's National Early Learning Development Standards for children from birth to 4 years (Department of Basic Education, 2009), the National Curriculum Framework from birth to 4 years (Department of Basic Education, 2015), and the Curriculum Assessment Policy Statements for Grade R (Department of Basic Education, 2011). ELOM items are grouped in domains, which in turn fall under selected South African ELDS, each with specified indicators and ELOM test items. Criteria for inclusion of an ELOM standard and indicator in a domain were based on

- a) alignment with national early learning policy is regarded as essential (e.g., Kagan, Castillo, Gomez, & Gowani, 2015; Zieky, Perie, & Livingston, 2008);
- b) high representation in South African, regional, and other international ELDS (ensuring alignment with commonly used approaches);
- c) appropriate for inclusion in the intended ELOM domains; and
- d) deemed essential by South African policy makers and educationists for readiness to benefit from Grade R (based on interviews prior to pilot).

Following Grisham-Brown, Hallam, and Brookshire (2006) and Bagnato, Neisworth, and Munson (1997), ELOM domains and items were informed by a thorough scan of developmental assessment instruments. Particularly relevant were instruments with established validity and reliability developed for use in the sub-Saharan region or in other low-income country settings. These included Save the Children's IDELA (Pisani et al., 2015), the MELQO, and the Zambian Child Assessment Test (Fink, Matafwali, Moucheraud, & Zuilkowsk, 2012), among others. Final item selection followed consultations with educationalists and other experts.

A 31-item instrument was piloted in Cape Town in the three main languages of the area (English, Afrikaans, and isiXhosa). The pilot was administered to 42 preschool children (26 girls and 16 boys average age 63 months). The language breakdown was 12 Afrikaans, 17 English, and 13 isiXhosa speakers. In the analysis, several of pilot items showed highly skewed score distributions indicating both floor and ceiling effects. The process is outlined in Dawes et al., (2019).

Post-pilot, the ELOM consisted of 23 direct assessment items in five domains: (a) gross motor development (four items), (b) fine motor coordination and visual motor integration (four items), (c) emergent numeracy and mathematics (five items), (d) cognition and executive functioning (four items), and (e) emergent literacy and language (six items). The full list of items and their sources is presented in Tables 1a, 1b, 1c, 1d, and 1e.

Particular attention was paid to translation, cultural fairness, and metric equivalence across the five languages to be included in the standardization of the ELOM described below (Malda et al., 2008; Milfont & Fischer, 2015; Pena, 2007). Translations from the English were checked, revised, and back translated following input from indigenous language speakers, familiar with less formal versions of the languages. As the form and structure of the languages varies considerably, language development specialists also advised on item construction and cross-language item equivalence.

2.1 | Sample

G*Power 3.0 online software was used to determine sample size, with power set to 0.8 (Faul, Erdfelder, Lang, & Buchner, 2007) and Alpha set to 0.05 for all analyses except *t* tests, where this was reduced to 0.01 as a limited manual control for Type 1 error. A scaling factor of 1.4 was used to estimate *N* for Tukey's HSD tests.

Key Messages

- There are no standardised early childhood development (ECD) measures that account for South Africa's diverse cultural, linguistic, and socio-economic landscape and can be administered by non-professionals.
- The Early Learning Outcomes Measure (ELOM) is the first South African, age validated, standardised instrument for children between 50 and 69 months of age that can be used by non-professionals.
- Psychometric analyses indicate that the ELOM domains are unidimensional and internally consistent measures of their constructs; the items hold interval scale status and display adequate ability to discriminate reliably between more and less able children; and items do not discriminate unfairly between children of different socio-economic backgrounds when interpreted at the domain level.
- The ELOM enables both public and private South African organisations to better understand the performance of their ECD programmes in relation to national socioeconomic groups and the ELOM standard.
- The ELOM affords organisations the ability to collaborate, share their learnings, and refine their interventions so that scarce resources are used for the greatest benefit to the children, families, and communities that they serve.

A two-stage clustered sample design was employed. The study was conducted in three public school districts (clusters) in three of South Africa's nine provinces. These districts were purposefully selected on the basis of the languages spoken by children in reception year classes (Grade R) of schools in each district. This was done to ensure adequate representation of children speaking the five languages included in the study: Afrikaans, English, Setswana, isiZulu, and isiXhosa, which according to the 2011 census, are spoken by 70% of the South African population. In the first stage and in each district, probability sampling (proportional to Grade R population size) was used to randomly select schools within each of five School Quintile bands (hereafter Quintiles). Schools are classified based on the household income of their catchment area (Quintile 1 is poorest). Quintiles 1 to 3 schools are attended by poorer children and are not permitted to charge fees. Very few Quintile 1 schools were observed in the selected districts and so could not be randomly selected. A purposive sample of Quintile 1 schools was recruited from rural areas to examine the performance of the poorest children and explore whether the ELOM was a fair assessment of children from these more "traditional" African backgrounds. isiZulu and Setswana speaking children constituted this sample. In the second stage, children in all quintiles were recruited through simple random sampling.

TABLE 1A Domain 1: gross motor development (direct assessment)

National Curriculum Framework (Department of Basic Education, 2015): Early Learning and Development Area 1: well-being Aim: Children should be physically strong and show abilities and interest in physical activities. Standard: South African National Early Learning Development Standards (Department of Basic Education, 2009) Desired Result 6: Children begin to demonstrate physical and motor abilities and an understanding of a healthy lifestyle.			
Standard	Indicator	Assessment items	Item sources
1.1: Children are increasingly able to use their large (gross) muscle skills.	Child shows good control and coordination in large movements.	Direct assessment: ELOM Items 1, 2, 3, 4: 1: stand on one foot (timed). 2, 3, 4: catch a beanbag thrown by examiner with (a) two hands against their body, (b) preferred hand, (c) other hand.	1: ASQ3 (Squires & Bricker, 2009). 2, 3, 4: McCarthy Scales South African adaptation (Richter, Griesel, & Rose, 1994). Others using same or similar methods: Lesotho ELDS (Lesotho Ministry of Education and Training, 2014), CDAT (Rao, 2007), EAP ECDS (Rao et al., 2014b).

Note. ELOM: Early Learning Outcomes Measure; ELDS: Early Learning Development Standards

TABLE 1B Domain 2: fine motor development and visual motor integration (direct assessment)

National Curriculum Framework Early Learning and Development Area 1: well-being Aim: Children should be physically strong and show abilities and interest in physical activities. Standard: South African National Early Learning Development Standards Desired Result 6: Children begin to demonstrate physical and motor abilities and an understanding of a healthy lifestyle.			
Standard	Indicator	Assessment items	Item sources
2.1: Children are increasingly able to use their small muscle (fine motor) skills.	Child uses small muscles competently and demonstrates visual-motor integration.	Direct assessment ELOM Items 5, 6, 7, 8: 5 (a) Use two sticks to copy a cross made of two sticks; (b) use sticks to copy a picture of a square. 6: Use a pencil to copy a triangle. 7: Use a pencil to draw self. 8: Number of beads (out of 10) threaded in 40 s.	5 a and b: Beery VMI (Beery, Buktenica, & Beery, 2010) 6: Beery VMI. Others using same or similar methods: IDELA (Save the Children International, 2014), EAP ECDS, ZAMCAT (Fink et al., 2012). 7: scoring draws on EAP ECDS and IDELA. Others using similar methods: ASQ 3. 8: ZAMCAT; others using same or similar methods: ASQ 3, EAP ECDS.

Note. ELOM: Early Learning Outcomes Measure; IDELA: International Development and Early Learning Assessment

TABLE 1C Domain 3: emergent numeracy and mathematics (direct assessment)

National Curriculum Framework Early Learning and Development Area 4: Exploring Mathematics Aims: 1: Children show awareness of and are responsive to number and counting; 2: Children sort, classify, make comparisons and solve problems. Standard: South African National Early Learning Development Standards Desired Result 5: Children are learning about mathematical concepts.			
Standard	Indicator	Assessment items	Item sources
3.1: Children demonstrate an understanding of number concepts.	Child is able to count with one to one correspondence.	Direct assessment ELOM Item 9: 1: Counts three marbles; 2: Counts eight buttons; 3: Counts 15 objects from a mixed presentation of marbles, buttons, and small sticks.	Modification of IDELA 1:1 correspondence item using three classes of object. Others using similar methods: MODEL, EAP ECDS.
	Child is able to do simple calculations using addition and subtraction.	Direct assessment: ELOM Item 10: 1: Child sums items using a picture card stimulus. 2: Child subtracts items using a picture card stimulus.	PreGypt Battery (Herbst, no date); others using similar methods: IDELA, EAP ECDS, MODEL, CDAT.
3.2: Children begin to demonstrate an understanding of symbols, shapes, size, and space.	Child is able to sort, classify, and match shapes (sorting and classification).	Direct assessment ELOM Item 11: Child groups stars and circles by colour and shape.	IDELA.
	Child understands measurement terms to do with size and length, amount (spatial vocabulary).	Direct assessment ELOM Item 12: from a picture: Child identifies objects in a picture that are above, under, in front of, on the side.	IDELA and MODEL.
	Child is able to understand vocabulary for location (measurement vocabulary).	Direct assessment ELOM Item 13: From picture stimuli: Child identifies biggest, smallest, longest, and shortest.	IDELA and MODEL.

Note. ELOM: Early Learning Outcomes Measure; IDELA: International Development and Early Learning Assessment

Although the sample's construction was not proportional to the languages of each district (these data were not available), it is (apart from Quintile 1 schools—see below), representative of children in

grade R in these districts. Given that school quintiles are assigned on the basis of national poverty data, it is very probable that the sample is typical of the socio-economic range of the country.

TABLE 1D Domain 4: cognition and executive functioning (direct assessment)

National Curriculum Framework Early Learning and Development Area 6: knowledge and understanding of the world			
Aims: 1: Children explore and investigate their life world; 2: Children explore design, make items, and use technology; 3: Children explore and investigate time and place.			
Standard: South African National Early Learning Development Standards Desired Result 1: Children are learning how to think critically, solve problems, and form concepts.			
Standard	Indicator	Assessment items	Item sources
4.1: Children are learning how to think critically, solve problems, and form concepts.	Child demonstrates cognitive flexibility (flexible rule use and working memory).	Direct assessment ELOM Item 14: Child sorts six cards, first according to colour and then switches to shape.	Dimensional Change Card Sort: Zelazo (2006). Also used in EAP ECDS.
4.2: Children are developing the ability to attend to instructions, remember them, and control impulses while performing a task.	Child demonstrates auditory discrimination, focused attention, working memory, and inhibitory control.	Direct assessment ELOM Item 15: Child copies the examiner's exact sequence of taps on the table with a pencil.	Pencil tapping task: ZAMCAT; Brooker et al. (2010). Others using similar methods EAP ECDS, Diamond and Taylor (1996). MODEL uses the head, toes, knees, and shoulders task of Ponitz et al. (2008).
	Child demonstrates short-term memory.	Direct assessment ELOM Item 16: Child is asked to repeat series of verbally presented numbers.	Digit span (forward) IDELA. Instructions based on the Children's Memory Scale (Cohen, 2011). Others using same or similar methods: MODEL and CDAT.
	Child demonstrates problem solving ability and working memory.	Direct assessment ELOM Item 17: Child assembles seven puzzles of increasing levels of difficulty.	Picture puzzle completion: Junior South African Individual Scale (Human Sciences Research Council, 1981; Herbst & Huisamen, 2000). IDELA uses one puzzle.

Note. ELOM: Early Learning Outcomes Measure; IDELA: International Development and Early Learning Assessment

TABLE 1E Domain 5: emergent literacy and language (direct assessment)

National Curriculum Framework: Early Learning and Development Area 3: communication			
Aims: 1: Children listen to sounds and speech; 2: Children listen with understanding; 3: Children speak using different styles of communication; 4: Children make meaning by "reading" what they see, hear, feel, taste, and touch.			
Standard: South African National Early Learning Development Standards desired Result 4: Children are learning to communicate effectively and use language.			
Standard	Indicator	Assessment items	Item sources
5.1: Children are able to communicate effectively and use language.	Child is able to speak in full sentences.	Direct assessment ELOM Item 18: In response to a stimulus picture of a girl crying, the child is asked to describe her feelings and actions to be taken to "help her feel better."	IDELA (modified picture); others using similar methods EAP ECDS. Prior to age validation, Items 18 and 19 were intended to measure empathic response and awareness of own feelings in the social and emotional awareness domain. On factor analysis, these items loaded with language items indicating strong reliance on expressive language.
	Child is able to relate an account of events that is logical and with correct language usage (expressive language).	Direct assessment ELOM Item 19: Child is asked to describe (a) what makes her/him feel sad, and what can be done to feel better; (b) what makes her/him feel happy.	
	Child is able to name common objects (expressive vocabulary)	Direct assessment ELOM Item 20: Child describes what she/he does when getting up in the morning.	Items 20 and 21 were developed for ELOM following pilot of IDELA and MELQO Version 6 Form A.
	Child shows understanding of stories told to her/him (listening comprehension)	Direct assessment ELOM Item 21: Child is asked to name items to be seen inside and outside their home.	MODEL and IDELA.
	Child recognises initial sounds of words (initial sound discrimination)	Direct assessment ELOM Item 22: Assessor reads a story, after which questions are asked of the child to gauge understanding.	MODEL and IDELA.
		Direct assessment ELOM Item 23: Child is asked to state which word commences with a particular phoneme.	MODEL, IDELA; others using similar methods: EAP ECDS.

Note. ELOM: Early Learning Outcomes Measure; IDELA: International Development and Early Learning Assessment; MELQO: Measuring Early-Learning Development Outcomes

2.2 | Preparation of data for analysis

Records were screened for duplicates, incomplete assessments, outliers, and influential values. A check for outliers and influential values was used to detect groups of scores that had an undue influence on

Quintile group mean scores. Schools were compared (using 95% confidence intervals) to identify those within quintiles, where confidence intervals did not overlap. Two Quintile 1 schools were identified, and their influence confirmed. These children were excluded. Four items from the World Health Organization Ten Point Disability Screen

(Durkin et al., 1995) were used to identify participants with problems in vision, hearing, intellectual functioning, and mobility during the assessment (see Table 2). Children with these challenges were removed from the sample. Quintiles 4 and 5 schools, and Quintiles 2 and 3 schools showed significant overlap of confidence intervals, indicating that they did not meaningfully differ in ELOM performance. This was confirmed during some preliminary multilevel modelling using all Quintile groups. Following this exploration, Quintiles 4 and 5 were combined, and 2 and 3 were combined resulting in three distinct socio-economic groups: Quintile 1, Quintiles 2/3, and Quintiles 4/5. Table 3a presents the final sample and Table 3b the purged records. Final sample records were imported into an IBM SPSS v22 data file for analysis.

3 | PROCEDURE

3.1 | Ethics

The study was approved by the University of Cape Town, Psychology Department Research Ethics Committee (Reference number PSY2015-048). Provincial Education Departments and schools gave their permission to assess children. Consent for child assessment was sought from parents. Children were asked to assent to participation and were informed that they could cease their participation at any time.

3.2 | ELOM administration

Native speakers of the languages of administration who had worked with or been involved in research with preschool children received a 5-day standardised training programme to become ELOM assessors. Inter-assessor reliability was established as satisfactory as presented in Table 4.

TABLE 2 Early Learning Outcomes Measure disability screen items

1	Did this child seem to have difficulty seeing?
2	Did this child appears to have difficulty with hearing?
3	When you told this child to do something, did he/she seems to have difficulty understanding what you are saying?
4	Did this child have difficulty in walking or moving his/her arms, or did he/she have weakness and/or stiffness in the arms or legs?

TABLE 3A Early Learning Outcomes Measure sample by quintile, language, and province

	Western Cape			Total	North West	KwaZulu-Natal	Total	All Target (%)
	English	Afrikaans	isiXhosa		Setswana	isiZulu		
Quintiles 4 and 5	112	38	46	196	81	184	461	154
Quintiles 2 and 3	38	97	201	336	202	218	756	126
Quintiles 1	—	—	—	—	114	—	114	95*
Total	150	135	247	532	397	402	1,331	130

*On the basis of the 120 target for Quintile 1 Setswana.

TABLE 3B Purged record frequencies by category

Category	Frequency (n)
Duplicated records	14
Incomplete records	18
Disability	85
Outliers	0
Influential values	42
Total	159

TABLE 4 Early Learning Outcomes Measure inter-assessor reliability by provincial team

	Fleiss' Kappa
Western Cape assessors (n = 8)	$\kappa = 0.795$
North West assessors (n = 6)	$\kappa = 0.684$
KwaZulu Natal assessors (n = 8)	$\kappa = 0.916$

The study was conducted within the first 6 weeks of the Grade R year. Assessments were undertaken individually at schools in a quiet venue and using a digital tablet-based manual and scoring system. Mean testing time was 58 min. Relevant information on the child (e.g., school, age, and gender) and his/her scores were digitised and coded into a tablet programme (SurveyCTO, 2016). This reduced the probability of error and missing values. Scores were uploaded daily to the master database.

3.3 | Analyses

Psychometric analyses were conducted to establish the age validity, reliability, validity, and fairness of the ELOM. First, unidimensionality and internal consistency of ELOM domains was assessed using confirmatory factor analysis (CFA; Byrne, 2010; Kline, 2011). Next, we undertook Rasch Analysis to establish a uniform interval scale for each domain (Bond & Fox, 2015). Finally, differential item functioning (DIF) techniques were used to ensure that ELOM items did not discriminate against children of the same potential ability from different backgrounds (Bond & Fox, 2015).

3.3.1 | Age and quintile differences

A multilevel linear model consisting of two levels was constructed with the upper representing the school and the lower representing

the children. Quintile, age (split into two categories—50 to 59 months and 60 to 69 months), and gender (male and female) were included as independent variables. As discussed earlier, three quintile groups were constructed. The dependent variables for each model were ELOM total score and each of five ELOM domains.

4 | RESULTS

4.1 | Unidimensionality and internal consistency of ELOM domains

CFA met all standards outlined by Byrne (2010) and Kline (2011). The results of the final CFA are presented in Table 5.

Table 6 summarises the results of the Rasch Modelling for each domain. All domains are adequately represented by their items and are reliable for prediction of a child's position above and below a standard (Bond & Fox, 2015; Linacre, 2016).

Rasch standard scores were constructed to make all Rasch logit scores positive and to rescale the logit scores so that they start at 0 and end at a maximum score of 20 for each domain—for more on this process, see Bond and Fox (2015) and Linacre (2016). The maximum score for each domain was chosen in order to create an equal weighting between the domains. The total contribution of each domain produces a maximum score of 100 for the ELOM. Tables 7a, 7b, 7c, 7d, and 7e show the original score and the transformed standard score—this also shows how ordinal original scores collapsed to produce a smaller group of ratio standard scores in many cases. This allows a more accurate and valid comparison of children with different levels of ability, based on their scores.

TABLE 7A Early Learning Outcomes Measure items and standard scores: gross motor development

Item number	Description	Original scoring	Standard score
1	Standing on one leg for 10 s	0	0
		1	1.36
		2	2.91
2	Catch bean bag both hands	0	0
		1	1.00
		2	2.07
3	Catch bean bag preferred hand	0	0
		2	2.22
		4	3.69
4	Catch bean bag nonpreferred hand	0	0
		2	2.62
		4	4.25
6		6	6.30
		6	7.20

4.2 | Differential item functioning

Table 8 presents the DIF for each domain and item, for gender (included to safeguard against bias caused by differential development in male and female children). No DIF was evident.

Socio-economic group and language group were also included to examine whether the performance of children from different socio-economic backgrounds or language groups was disadvantaged on particular tasks. Figure 1a, 1b, 1c, 1d, and 1e present DIF by socio-economic group, as represented by Quintile 1 (Q1), Quintiles 2 and 3 (Q2/3), and Quintiles 4 and 5 (Q4/5). Figure 2 presents DIF by the five language groups.

DIF is only a concern if there is no opposing DIF in other items to counterbalance the effect within the domain (Bond & Fox, 2015;

TABLE 5 Confirmatory factor analysis fit statistics by Early Learning Outcomes Measure domain

ELOM domains	χ^2	<i>p</i>	SRMR	GFI	CFI	RMSEA	10% CI RMSEA	90% CI RMSEA
Gross motor development	4.65	0.098	0.01	0.99	0.99	0.03	<0.01	0.07
Fine motor coordination and visual motor integration	0.45	0.801	0.01	1.00	1.00	<0.01	<0.01	0.03
Emergent numeracy and mathematics	1.65	0.439	0.01	0.99	1.00	<0.01	<0.01	0.06
Cognition and executive functioning	4.72	0.095	0.02	0.99	0.99	0.03	<0.01	0.07
Emergent literacy and language	4.14	0.126	0.01	0.99	0.99	0.03	<0.01	0.07

Note. ELOM: Early Learning Outcomes Measure; SRMR: Standard Root Mean Residual; GFI: Goodness of Fit Index; CFI: Comparative Fit Index; RMSEA: Root Mean Square Error of Approximation.

TABLE 6 Rasch statistics for Early Learning Outcomes Measure domains

Domain	Mean square		Point-measure correlation	Rasch reliability		Variance explained
	Infit	Outfit		Person	Item	
Gross motor development	0.91–1.2	0.86–0.97	0.47–78	0.75	1.00	75.3%
Fine motor coordination and visual motor integration	0.94–1.08	0.91–1.09	0.55–0.83	0.71	1.00	65.1%
Emergent numeracy and mathematics	0.85–1.20	0.74–1.25	0.49–0.73	0.63	1.00	66.8%
Cognition and executive functioning	0.68–1.17	0.60–1.19	0.57–0.81	0.65	1.00	60%
Emergent literacy and language	0.91–1.15	0.73–1.08	0.52–0.74	0.75	1.00	59.3%

TABLE 7B Early Learning Outcomes Measure items and standard scores: fine motor coordination and visual motor integration

Item number	Description	Original scoring	Standard score
5	Copy 3D shapes	0	0
		1	3.11
		2	4.20
6	Copy triangle	0	0
		1	
		2	4.58
7	Draw self	0	0
		1	
		2	
		3	3.12
		4	
		5	3.77
		6	
		7	4.74
8	String beads	0	0
		1	1.00
		2	2.03
		3	2.75
		4	3.41
		5	3.95
		6	4.51
		7	5.33
		8	6.48
		9	
10			

TABLE 7C Early Learning Outcomes Measure items and standard scores: emergent numeracy and mathematics

Item number	Description	Original scoring	Standard score
9	Counting in classes	0	0
		1	2.81
		2	4.05
		3	6.03
10	Addition and subtraction	0	0
		1	2.84
		2	4.51
11	Sorting and classification	0	
		1	2.47
		2	4.01
12	Spatial vocabulary	0	0
		1	
		2	1.38
		3	
13	Measurement vocabulary	0	
		1	
		2	1.00
		3	
4	2.65		

Linacre, 2016). After consulting the adjusted score totals associated with all noticeable cases of DIF, all DIF balanced out adequately to maintain the fairness of the domains.

4.3 | Age and quintile differences

The total ELOM score and each of the domain scores were modelled separately. Estimated marginal means are presented in Table 9. These results suggest that quintile and age groups differ as expected.

TABLE 7D Early Learning Outcomes Measure items and standard scores: cognition and executive functioning

Item number	Description	Original scoring	Standard score
14	Dimension change card sort	0	0
		1	1.00
		2	3.11
15	Pencil tapping test	0	0
		1	1.34
		2	
		3	
		4	2.68
		5	
		6	
		7	
		8	4.37
		9	
10	6.72		
16	Digits forward	0	0
		1	
		2	
		3	2.11
		4	4.03
17	Picture puzzle completion	0	0
		1	1.26
		2	3.02
		3	
		4	
		5	6.14
		6	
7			

Further, in some cases, older children in lower quintiles perform similarly to younger children in higher quintiles (total ELOM column).

In sum, psychometric analyses indicated that the direct assessment ELOM domains are

- unidimensional and internally consistent measures of their constructs;
- the items hold interval scale status and display adequate ability to discriminate reliably between more and less able children; and
- items do not discriminate unfairly between children of different socio-economic backgrounds when interpreted at the domain level.

5 | CONSTRUCTION OF ELOM STANDARDS AND NORMS

There are several methods for setting performance standards and their associated cut scores (Kane, 2011; Ricker, 2006; Zieky et al., 2008).

International experts recommend that ELDS should be set at a representative sample median (Kagan & Britto, 2005a; Kagan & Britto, 2005b; Kagan & Britto, 2008a; Kagan & Britto, 2008b). However, for the ELOM, we regarded this to be too low because 65% of the ELOM sample represented children from disadvantaged backgrounds (Quintiles 1 to 3). This lowered the total sample median (Quintiles 4 and 5 children scored significantly higher than the other groups). For this reason, standards should be aspirational so that children who meet them are ready to learn from what is offered when they enter Grade

TABLE 7E Early Learning Outcomes Measure items and standard scores: emergent literacy and language

Item number	Description	Original scoring	Standard score
18	Expressive language: empathy	0	3.21
		1	
		2	
		3	
19	Expressive language: self-awareness	0	2.64
		1	
		2	
		3	
20	Expressive language	0	3.42
		1	
		2	
		3	
21	Expressive vocabulary	0	3.41
		1	
		2	
		3	
		4	
		5	
22	Oral comprehension	0	2.47
		1	
		2	
		3	
		4	
		5	
		6	
		7	
		8	
		9	
23	Initial sound discrimination	0	3.17
		1	
		2	
		3	

Standard score = (proposed maximum score – proposed minimum score) / (real maximum score – real minimum score) (Rasch logit – real minimum score) + proposed minimum score.

R. This should also be a realistic level of performance to expect from early learning programmes.

Rasch Analysis was used to derive standard score distributions for ELOM total scores and for each of the domains. ELOM performance standards were benchmarked at the standard score achieved by the top 40% of children in the sample each domain and for ELOM total.

Following Zieky et al. (2008), we presented the empirically derived standards to an expert group, including representatives of government departments responsible for early childhood policy and services (social development, basic education and planning, monitoring, and evaluation). This group approved the standard.

As part of this process, we provided quintile group performance profiles (with median standard scores) to permit early learning programme management to compare the performance of their children with that of children from similar backgrounds, as well as the standard they should aspire to achieve. Finally, as illustrated in Table 10, we categorised standard score bands as achieving the standard, falling behind, and at risk.

TABLE 8 Differential item functioning for each domain by gender

Item	Domain	Differential item functioning	
		DIF contrast	$p_{\text{Rasch-Welch}}^{\text{Rasch}}$
Gender (male: $n = 640$; female: $n = 691$)			
1	Gross motor development	0.17	0.25
2		-0.15	0.15
3		0.10	0.34
4		-0.02	0.85
5	Fine motor control and visual motor integration	-0.40	<0.01
6		0.00	1.00
7		0.20	0.06
8		0.00	1.00
9	Emergent numeracy and mathematics	0.19	0.09
10		0.13	0.26
11		-0.27	0.03
12		0.16	0.21
13		0.06	0.71
14	Cognition and executive functioning	0.13	0.24
15		0.12	0.08
16		0.00	1.0
17		-0.30	<0.01
18	Emergent language and literacy	0.27	0.07
19		0.00	1.00
20		-0.05	0.67
21		0.11	0.30
22		-0.16	0.14
23		-0.10	0.43

Note. DIF: differential item functioning

6 | DISCUSSION

With the adoption of South Africa's early childhood policy in 2016, the importance of a sound preparation for school during the early years is widely recognised. This is particularly important given evidence of the poor performance of economically disadvantaged children, particular in the early grades, with performance gaps between socio-economic strata widening as children progress through the education system (Department of Basic Education, 2014). The country has very limited evidence on the quality of early learning programmes of any type. One representative sample study of centre-based programmes indicates that their quality, particularly for economically disadvantaged children, is poor and unlikely to enhance readiness to learn in school (Biersteker et al., 2016). The vast majority of South African early care and education programmes are funded by parents receiving government subsidies for the poor. The state and private funders increasingly require evidence that their money is well spent on quality programmes. Furthermore, SDG 4.2 requires that countries track the quality and outcomes of early childhood programmes with the goal of improving readiness for primary education (Gove & Black, 2016). The SDGs present a huge opportunity for ensuring that no children are left behind on account of economic disadvantage and other impediments (Raikes, Britto, Yoshikawa, & Iruka, 2017).

Appropriate tools are required for measuring the quality of children's services and their effectiveness. We argue that a national population level instrument such as ELOM that accounts for context and

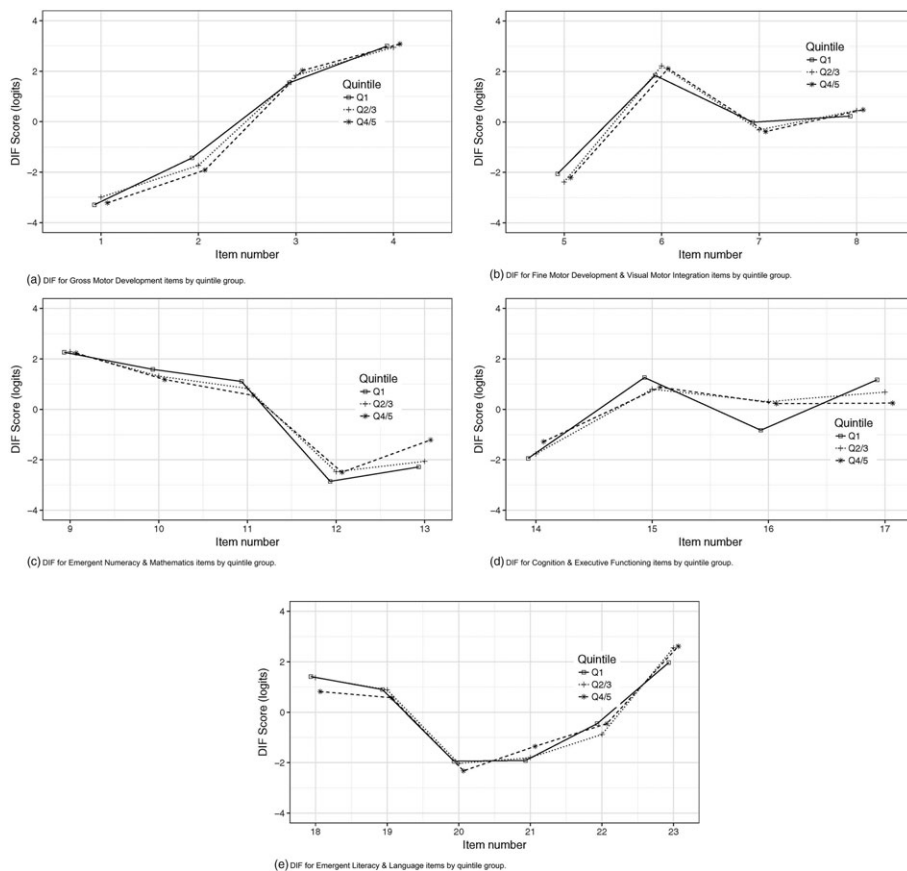


FIGURE 1 (a). Differential item functioning (DIF) for gross motor development items by quintile group. (b). DIF for fine motor development and visual motor integration items by quintile group. (c). DIF for emergent numeracy and mathematics items by quintile group. (d). DIF for cognition and executive functioning items by quintile group. (e). DIF for emergent literacy and language items by quintile group

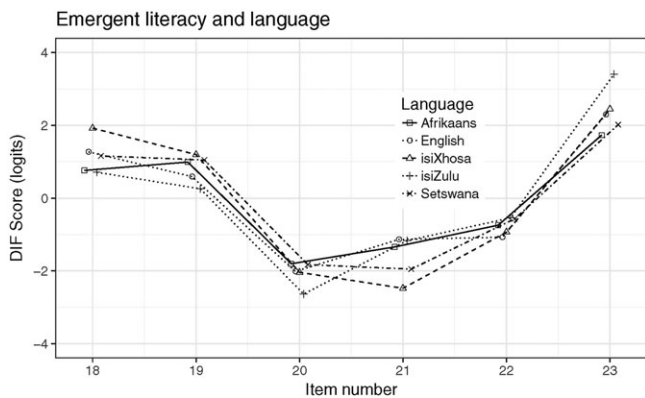


FIGURE 2 Differential item functioning for emergent literacy and language items by home language of child

population characteristics and is linked to national early learning development standards, would be most suited to this purpose. Prior to its development, South Africa lacked a locally standardised instrument for measuring the outcomes of early learning programmes for children. In our introduction, we argued for an instrument aligned to national policy, curriculum, and early learning development standards and that it should not unfairly discriminate between children from the different socio-economic and ethnolinguistic backgrounds assessed in this study. Psychometric analysis presented here confirms

that items in each domain measure the relevant construct, the ELOM discriminates between children of different abilities. It does not disadvantage children of similar ability who speak one of the five languages tested in this study. Children of the same ability perform similarly, regardless of their socio-economic status. Caution must be exercised in interpreting the results for children from Quintile 1 schools, which were not randomly selected. We await further studies to determine whether the ELOM provides a fair test of children speaking languages not assessed here.

The ELOM learning community has been established to support programmes to improve practice and enhance the developmental child outcomes of poor children prior to the reception year. The ELOM provides a means for ECD programme staff to establish children's level of performance relative to expected standards (whether they are "on track" or not). It also makes it possible to monitor their performance over the course of the programme and share what they have learned about delivering ECD services, at scale, within some of the most disadvantaged communities in South Africa.

In terms of further research, several initiatives are under way. First, the properties of the ELOM are not fully established, and studies are in preparation to establish concurrent, test-retest, and criterion-related validity. Second, as Quintile 1 children were not well represented in the study, data from more representative samples will be gathered. ELOM is currently being used in a comparative study of

TABLE 9 Estimated marginal means by age and quintile for Early Learning Outcomes Measure (ELOM) total and each ELOM domain

Quintile	Age	Total ELOM ^{ab}	GMD ^c	FMC&VMI ^{ab}	EN&M ^b	C&EF ^{ab}	EL&L ^b
Q1 (n = 114)	<5 (n = 53)	37.15 ± 3.58	6.72 ± 1.10	11.88 ± 0.88	7.97 ± 1.10	4.77 ± 1.10	5.81 ± 1.17
	>5 (n = 61)	41.13 ± 3.33	8.21 ± 1.03	12.50 ± 0.82	8.08 ± 1.02	5.95 ± 1.02	6.40 ± 1.10
Q2/3 (n = 756)	<5 (n = 115)	41.27 ± 2.43	7.62 ± 0.75	11.11 ± 0.60	7.99 ± 0.74	5.58 ± 0.75	8.97 ± 0.80
	>5 (n = 641)	49.89 ± 1.03	9.98 ± 0.32	12.87 ± 0.25	9.03 ± 0.32	7.78 ± 0.32	10.24 ± 0.34
Q4/5 (n = 461)	<5 (n = 90)	48.80 ± 2.75	8.39 ± 0.85	11.81 ± 0.67	9.34 ± 0.84	7.93 ± 0.84	11.32 ± 0.90
	>5 (n = 371)	54.29 ± 1.35	9.14 ± 0.42	14.18 ± 0.33	9.84 ± 0.41	9.29 ± 0.42	11.83 ± 0.44

Note. ELOM: Early Learning Outcomes Measure; FMC&VMI: fine motor coordination and visual motor integration; EN&M: emergent numeracy and mathematics; C&EF: cognition and executive functioning; EL&L: emergent literacy and language

^a<5" differ significantly from ">5" ($p < 0.05$).

^bBoth Q1 and Q2/3 differ significantly from Q4/5 ($p < 0.05$).

^cOnly Q2/3 differed significantly from Q4/5 ($p < 0.05$).

TABLE 10 Early Learning Outcomes Measure (ELOM) standards and performance bands: ELOM total and domains

	50–59 months			60–69 months		
	At risk	Falling behind	Achieving the standard	At risk	Falling behind	Achieving the standard
ELOM total	0–35.68	35.69–45.42	45.43–100	0–43.57	43.58–54.66	54.67–100
Gross motor development	0–4.97	4.98–8.66	8.67–20	0–7.19	7.20–10.74	10.75–20
Fine motor coordination and visual motor integration	0–10.17	10.18–11.55	11.56–20	0–11.37	11.38–14.20	14.21–20
Emergent numeracy and mathematics	0–6.49	6.50–9.07	9.08–20	0–6.64	6.65–10.54	10.55–20
Cognition and executive functioning	0–4.02	4.03–7.13	7.14–20	0–6.12	6.13–8.96	8.97–20
Emergent literacy and language	0–6.80	6.81–10.47	10.48–20	0–7.95	7.96–12.31	12.32–20

the effectiveness of different types of early learning programmes for low-income children. Finally, the ELOM is now available in all eleven official languages.

7 | LIMITATIONS

The ELOM standardization is not based on a nationally representative sample but rather on random samples of children in Grade R classrooms in three education districts in three of South Africa's nine provinces. However, given that school quintiles are assigned nationally using the same metric and given random selection of schools within each quintile (except Quintile 1), we would expect children from the same quintiles in other districts and provinces to perform similarly to those in this study. Although we regard ELOM as suitable for use nationally at this stage, adjustments may be required as we accumulate more data from around the country.

The Quintile 1 sample is not representative of children in that income band due to the districts not having enough schools in this category for randomization. However, our results confirm that they perform poorest relative to the higher quintile children on ELOM, as would be expected. On the basis of our experience, we view the selected schools as typical of the lowest economic stratum. Further studies of Quintile 1 children are required to confirm our profile for this group.

Children were not sampled proportionally by language spoken in each education district as data on the language distribution were not available. Future studies are required to address this limitation so that

our findings can be updated, as required. As the ELOM is used, we will be able to add to the database and examine the performance of children speaking other home languages.

Finally, the standards benchmarked through the use of the ELOM must be regarded as preliminary and may be revised as the instrument is used in the field. This may occur as more data is gathered to assess the extent to which programmes are successful in assisting children to reach expected levels of performance. It may be the case that the bar has been set too high, given the limited quality of many early learning programmes and the significant levels of child poverty (Biersteker et al., 2016; Children Count, 2017). Further investigation will tell.

ACKNOWLEDGEMENTS

The authors are grateful to the Investment Committee of Innovation Edge for funding the study; to the Departments of Education in the Western Cape, Free State, and KwaZulu-Natal provinces of South Africa for permission to conduct the research; to participating schools, teachers, and children, and to the many colleagues who provided advice and critical comment during the development of the ELOM. Finally, the authors are grateful to Andrew Snelling, for assisting with preparations for the submission of this article.

DECLARATION OF INTEREST

Innovation Edge (www.innovationedge.org.za) funded the study but had no role in its design, the analysis of the data, or the preparation

of this article. Innovation Edge approved the submission of this article for publication. All the authors provide research consultancy services to Innovation Edge from time to time.

ORCID

Matthew Snelling  <https://orcid.org/0000-0001-8925-7408>

REFERENCES

- About, F. E. (2006). Evaluation of an early childhood preschool program in rural Bangladesh. *Early Childhood Research Quarterly*, 21(1), 46–60. <https://doi.org/10.1016/j.ecresq.2006.01.008>
- Bagnato, S. J., Neisworth, J. T., & Munson, S. M. (1997). *Linking assessment and early intervention: An authentic curriculum-based approach*. Baltimore: Paul H. Brookes.
- Bailey, D., Duncan, G. J., Odgers, C. L., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, 10(1), 7–39. <https://doi.org/10.1080/19345747.2016.1232459>
- Beery, K. E., Buktenica, N. A., & Beery, N. A. (2010). *Beery-Buktenica developmental test of visual-motor integration* (6th ed.). Minneapolis, MN: NCS Pearson.
- Berry, L., Dawes, A., & Biersteker, L. (2013). Getting the basics right. An essential package of services and support for ECD. In L. Berry, L. Biersteker, A. Dawes, L. Lake, & C. Smith (Eds.), *South African child gauge* (pp. 26–33). Cape Town: Children's Institute, University of Cape Town. Retrieved from https://open.uct.ac.za/bitstream/handle/11427/4089/CI_chapters_sachildgauge13_essentialpackage_2013.pdf?sequence=1
- Biersteker, L., Dawes, A., Hendricks, L., & Tredoux, C. (2016). Center-based early childhood care and education program quality: A South African study. *Early Childhood Research Quarterly*, 36, 334–344. <https://doi.org/10.1016/j.ecresq.2016.01.004>
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences*. New York: Routledge, Taylor and Francis Group.
- Brinkman, S., & Vu, B. T. (2016). *Early childhood development in Tonga: Baseline results from the Tongan early human capability index*. Washington, DC: World Bank. Retrieved from <https://openknowledge.worldbank.org/handle/10986/25674>: World Bank Publications. License: CC BY 3.0 IGO
- Britto, P., Yoshikawa, H., & Boller, K. (2011). Quality of early childhood development programs in global contexts. *Social Policy Report*, 25(2), 1–30. Retrieved from <https://eric.ed.gov/?id=ED519240>
- Brooker, S., Okello, G., Njagi, K., Dubeck, M. M., Halliday, K. E., Inyega, H., & Jukes, M. C. (2010). Improving educational achievement and anaemia of school children: Design of a cluster randomised trial of school-based malaria prevention and enhanced literacy instruction in Kenya. *Trials*, 11, 93. <https://doi.org/10.1186/1745-6215-11-93>
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. New York, NY: Routledge.
- Cohen, M. J. (2011). Children's Memory Scale. In J. S. Kreutzer, B. Caplan, & J. DeLuca (Eds.), *Encyclopedia of clinical neuropsychology* (pp. 556–559). New York, NY: Springer.
- Cole, M., & Bruner, J. S. (1971). Cultural differences and inferences about psychological processes. *American Psychologist*, 26(10), 867–876. <https://doi.org/10.1037/h0032240>
- Children Count. (2017). Statistics on children in South Africa. Retrieved from <http://childrencount.uct.ac.za/domain.php?domain=2>.
- Dasen, P. R. (1984). The cross-cultural study of intelligence: Piaget and the Baoule. *International Journal of Psychology*, 19(1–4), 407–434. <https://doi.org/10.1080/00207598408247539>
- Dawes, A., Biersteker, L., Girdwood, E., Snelling, M. J. T. L., & Tredoux, C. G. (2019). *Early learning outcomes measure. Technical manual (Revision 2)*. Claremont, Cape Town: The Innovation Edge. Retrieved from <http://www.innovationedge.org.za/>
- Department of Basic Education (2009). *National Early Learning and Development Standards for children birth to four years (NELDS)*. Pretoria: Department of Basic Education. Retrieved from <http://www.gauteng.gov.za/government/departments/education/Documents1/NELDS%20guidelines%208%20Jan.pdf>
- Department of Basic Education (2011). *Grade R curriculum and assessment policy statement Grades R-3*. Pretoria: Department of Basic Education. Retrieved from <http://www.education.gov.za/>
- Department of Basic Education (2014). *Report on the annual national assessments of 2014*. Pretoria, South Africa: Department of Basic Education. Retrieved from http://www.saqqa.org.za/docs/rep_annual/2014/REPORT%20ON%20THE%20ANA%20OF%202014.pdf
- Department of Basic Education (2015). *The South African National Curriculum Framework for children from birth to four*. Pretoria: Department of Basic Education. Pretoria: Department of Basic Education. Retrieved from <https://www.education.gov.za/Portals/0/Documents/curriculum%20docs/NCF%202018/NCF%20English%202018%20web.pdf?ver=2018-05-14-124718-317>.
- Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to “Do as I say, not as I do”. *Developmental Psychobiology*, 29(4), 315–334. [https://doi.org/10.1002/\(sici\)1098-2302\(199605\)29:4<315::aid-dev2>3.3.co;2-c](https://doi.org/10.1002/(sici)1098-2302(199605)29:4<315::aid-dev2>3.3.co;2-c)
- Dowd, A. J., Borisova, I., Amente, A., & Yenew, A. (2016). Realizing capabilities in Ethiopia: Maximizing early childhood investment for impact and equity. *Journal of Human Development and Capabilities*, 17(4), 477–493. <https://doi.org/10.1080/19452829.2016.1225702>
- Dua, T., Tomlinson, M., Tablante, E., Britto, P., Yousfzai, A., Daelmans, B., & Darmstadt, G. L. (2016). Comment. Global research priorities to accelerate early child development in the sustainable development era. *The Lancet Global Health*, 4(12), 887–889. [https://doi.org/10.1016/s2214-109x\(16\)30218-2](https://doi.org/10.1016/s2214-109x(16)30218-2)
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... Sexton, H. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428. <https://doi.org/10.1037/0012-1649.43.6.1428-1446>.
- Durkin, M. S., Wang, W., Shrout, P. E., Zaman, S. S., Hasan, Z. M., Desai, P., & Davidson, L. L. (1995). Evaluating a ten question screen for childhood disability: Reliability and internal structure in different cultures. *Journal of Clinical Epidemiology*, 48(5), 657–666. [https://doi.org/10.1016/0895-4356\(94\)00163-k](https://doi.org/10.1016/0895-4356(94)00163-k)
- ECD measure (2018). What is MELQO. Retrieved from: <http://ecdmeasure.org/about-melqo/what-is-melqo/>
- Engle, P. L., Fernald, L. C., Alderman, H., Behrman, J., Ogara, C., Yousafzai, A., ... Iltus, S. (2011). Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *The Lancet*, 378(9799), 1339–1353. [https://doi.org/10.1016/s0140-6736\(11\)60889-1](https://doi.org/10.1016/s0140-6736(11)60889-1)
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. Retrieved from <https://link-springer-com.ezproxy.uct.ac.za/article/10.3758/BF03193146>
- Fernald, L. C. H., Kariger, P., Engle, P., & Raikes, A. (2009). *Examining early child development in low-income countries: A toolkit for the assessment of children in the first five years of life*. Washington DC: The World Bank. Retrieved from <http://siteresources.worldbank.org.ezproxy.uct.ac.za/>

- EXTAFRREGTOPEDUCATION/Resources/444707-1291071725351/ExaminingECDtoolkitFULL.pdf
- Fink, G., Matafwali, B., Moucheraud, C., & Zuilkowsk, S. S. (2012). The Zambian early childhood development project 2010 assessment final report. Center on the Developing Child at Harvard University. Retrieved from <http://www.poverty-action.org/study/zambian-early-childhood-development-project>.
- Gove, A., & Black, M. M. (2016). Measurement of early childhood development and learning under the Sustainable Development Goals. *Journal of Human Development and Capabilities*, 17(4), 599–605. <https://doi.org/10.1080/19452829.2016.1243520>
- Grisham-Brown, J., Hallam, R., & Brookshire, R. (2006). Using authentic assessment to evidence children's progress toward early learning standards. *Early Childhood Education Journal*, 34(1), 45–51. <https://doi.org/10.1007/s10643-006-0106-y>
- Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *The American Economic Review*, 103(6), 2052–2086. <https://doi.org/10.1257/aer.103.6.2052>
- Herbst, I., & Huysamen, G. K. (2000). The construction and validation of developmental scales for environmentally disadvantaged preschool children. *South Africa Journal of Psychology*, 30(3), 19–26. <https://doi.org/10.1177/008124630003000304>
- Hsiao, C., Richter, L., Makusha, T., Matafwali, B., Heerden, A., & Mabaso, M. (2016). Use of the ages and stages questionnaire adapted for South Africa and Zambia. *Child: Care, Health and Development*, 43(1), 59–66. <https://doi.org/10.1111/cch.12413>
- Human Sciences Research Council (1981). Handleiding vir die Junior Suid-Afrikaanse Individuele Skaal (JSAIS). Pretoria: South African Institute for Psychological and Psychometric Research.
- Janus, M. E. (2007). The Early Development Instrument: A tool for monitoring children's development and readiness for school. In M. E. Young, & L. M. Richardson (Eds.), *Early child development from measurement to action: A priority for growth and equity* (pp. 183–202). Washington DC: World Bank Publications. <https://doi.org/10.1596/978-0-8213-7086-5>
- Kagan, S., & Britto, P. R. (2008a). Early Learning & Development Standards: Validation review ELDS. Workshop. Bangkok, Thailand. February 11th – 15th, 2008.
- Kagan, S., & Britto, P. R. (2008b). Validation to Implementation. Workshop. Bangkok, Thailand. February 11th – 15th, 2008.
- Kagan, S. L., & Britto, P. R. (2005a). *Going global with indicators of child development. UNICEF Final Report*. New York, NY: UNICEF.
- Kagan, S. L., & Britto, P. R. (2005b). *Process guidelines: Steps to developing standards for early learning*. New York, NY: UNICEF.
- Kagan, S. L., Castillo, E., Gomez, R. E., & Gowani, S. (2015). Understanding and using early learning standards for young children globally. *International Journal of Child Care and Education Policy*, 7(2), 53–66. <https://doi.org/10.1007/2288-6729-7-2-53>
- Kagitcibasi, C., Sunar, D., Bekman, S., Baydar, N., & Cemalcilar, Z. (2009). Continuing effects of early enrichment in adult life: The Turkish Early Enrichment Project 22 years later. *Journal of Applied Developmental Psychology*, 30(6), 764–779. <https://doi.org/10.1016/j.appdev.2009.05.003>
- Kane, M. T. (2011). So much remains the same: Conception and status of validation in setting standards. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 53–88). Taylor and Francis e-library.
- Kline, R. B. (2011). *Principles and practice of structural equation modelling*. New York, NY: Guilford Publications.
- Lesotho Ministry of Education & Training (2014). *Early learning and development standards for early childhood education and development in Lesotho Revised Version*. Maseru: Ministry of Education and Training.
- Linacre, J. M. (2016). Index. Retrieved from <http://www.winsteps.com/index/htm>.
- Malda, M., van de Vijver, F. J. R., Srinivasan, K., Transler, C., Sukumar, P., & Rao, K. (2008). Adapting a cognitive test for a different culture: An illustration of qualitative procedures. *Psychology Science Quarterly*, 50(4), 451–468. Retrieved from http://www.psychologie-aktuell.com/fileadmin/download/PsychologieScience/4-2008/01_Malda.pdf
- Matafwali, B., & Serpell, R. (2014). Design and validation of assessment tests for young children in Zambia. *New Directions for Child and Adolescent Development*, 2014(146), 77–96. <https://doi.org/10.1002/cad.20074>
- Milfont, T. L., & Fischer, R. (2015). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 111–130. <https://doi.org/10.21500/20112084.857>
- Mwaura, P. A., & Marfo, K. (2011). Bridging culture, research, and practice in early childhood development: The Madrasa Resource Centers in East Africa. *Child Development Perspectives*, 5(2), 134–139. <https://doi.org/10.1111/j.1750-8606.2011.00168.x>
- Mwaura, P. A., Sylva, K., & Malmberg, L. E. (2008). Evaluating the Madrasa preschool programme in East Africa: A quasi-experimental study. *International Journal of Early Years Education*, 16(3), 237–255. <https://doi.org/10.1080/09669760802357121>
- Pena, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, 78(4), 1255–1264. <https://doi.org/10.1111/j.1467-8624.2007.01064.x>
- Pisani, L., Borisova, I., & Dowd, A. J. (2015). International Development and Early Learning Assessment Technical Working Paper. Retrieved from http://resourcecentre.savethechildren.se/sites/default/files/documents/idela_technical_working_paper_v3_nodraft.pdf.
- Ponitz, C. E. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly*, 23(2), 141–158. <https://doi.org/10.1016/j.ecresq.2007.01.004>
- Raikes, A., Britto, P., Yoshikawa, H., & Iruka, I. (2017). Children, youth and developmental science in the 2015–2030 global sustainable development goals. *SRCD Social Policy Report*, 30(3), 1–23. Retrieved from https://www.srcc.org/sites/default/files/documents/spr_30_3.pdf
- Rao, N. (2007). *Cambodian developmental assessment test*. Cambodia: UNICEF.
- Rao, N., Sun, J., Ng, M., Becher, Y., Lee, D., Ip, P., & Bacon-Shone, J. (2014a). Report on validation, finalization and adoption of the East Asia-Pacific Early Child Development Scales (EAP-ECDS). UNICEF. Retrieved from <http://www.arnec.net/wp-content/uploads/2015/07/EAP-ECDS-Final-Report1.pdf>.
- Rao, N., Sun, J., Ng, M., Becher, Y., Lee, D., Zhang, L., & Lau, C. (2014b). *East-Asia Pacific early child development scales*. Hong Kong: Faculty of Education, the University of Hong Kong.
- Republic of South Africa (2015). *National Integrated Early Childhood Development Policy*. Pretoria: Government Printers. Retrieved from http://www.dsd.gov.za/index2.php?option=com_docman&task=doc_view&gid=652&Itemid=39
- Richter, L. M., Daelmans, B., Lombardi, J., Heymann, J., Boo, F. L., Behrman, J. R., ... Darmstadt, G. L. (2017). Investing in the foundation of sustainable development: Pathways to scale up for early childhood development. *The Lancet*, 389(10064), 103–118. [https://doi.org/10.1016/s0140-6736\(16\)31698-1](https://doi.org/10.1016/s0140-6736(16)31698-1)

- Richter, L. M., Griesel, R. D., & Rose, C. B. (1994). The McCarthy scales of children's abilities: Adaptation and norms for use amongst Black South African children. *The South African Journal of Occupational Therapy*, 24(1), 17–30.
- Ricker, K. L. (2006). Setting cut-scores: A critical review of the Angoff and modified Angoff methods. *Alberta Journal of Educational Research*, 52(1), 53–64. Retrieved from <https://search-proquest-com/docview/228639019?pq-origsite=gscholar>
- Save the Children International (2014). International Development and Early Learning Assessment (IDELA). Save the Children.
- Serpell, R., & Simatende, B. (2016). Contextual responsiveness: An enduring challenge for educational assessment in Africa. *Journal of Intelligence*, 4(1), 3. <https://doi.org/10.3390/jintelligence4010003>
- Squires, J., & Bricker, D. (2009). *Ages & stages questionnaires: A parent-completed child monitoring system* (Third ed.). Baltimore: Paul H. Brookes Publishing.
- SurveyCTO (2016). Home. Retrieved from <https://www.surveyccto.com/index.html>.
- UNICEF (2016). Early Learning and Development Standards (ELDS) and school readiness. Evaluation report. New York, UNICEF. Retrieved from https://www.unicef.org/evaldatabase/index_ELDS.html.
- Walker, S. P., Chang, S. M., Powell, C. A., & Grantham-McGregor, S. M. (2005). Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: Prospective cohort study. *The Lancet*, 366(9499), 1804–1807. [https://doi.org/10.1016/s0140-6736\(05\)67574-5](https://doi.org/10.1016/s0140-6736(05)67574-5)
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1(1), 297–301. <https://doi.org/10.1038/nprot.2006.46>
- Zieky, M. J., Perie, M., & Livingston, S. (2008). *Cutscores: A manual for setting standards of performance on educational and occupational tests* Educational Testing Service. Retrieved from http://www.ets.org/research/policy_research_reports/publications/book/2008/gvcw
- Zuilkowski, S. S., McCoy, D. C., Serpell, R., Matafwali, B., & Fink, G. (2016). Dimensionality and the development of cognitive assessments for children in Sub-Saharan Africa. *Journal of Cross-Cultural Psychology*, 47(3), 341–354. <https://doi.org/10.1177/0022022115624155>

How to cite this article: Snelling M, Dawes A, Biersteker L, Girdwood E, Tredoux C. The development of a South African Early Learning Outcomes Measure: A South African instrument for measuring early learning program outcomes. *Child Care Health Dev.* 2019;45:257–270. <https://doi.org/10.1111/cch.12641>