Marko-D validation in collaboration with Bala Wande

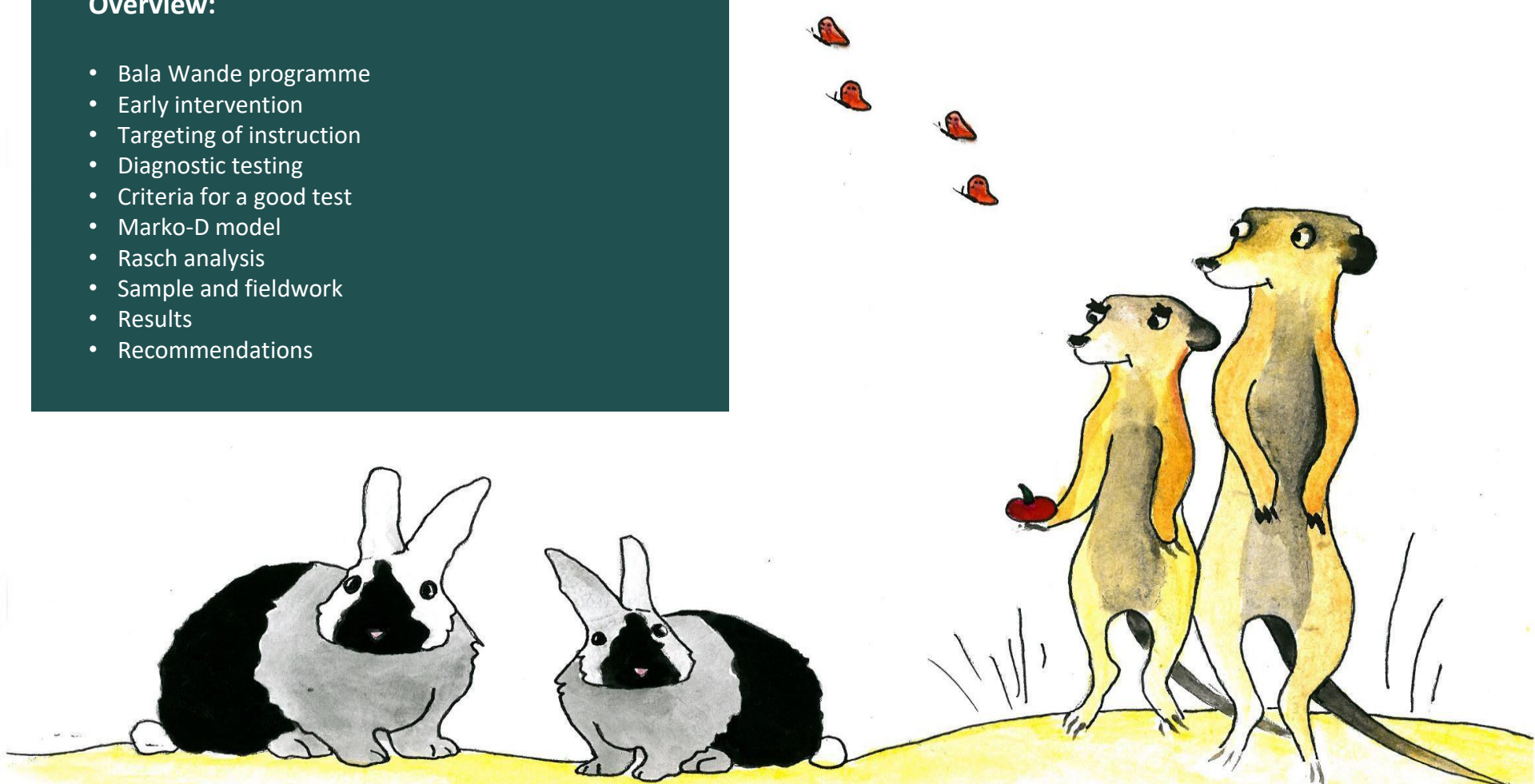
ROELIEN HERHOLDT (JET Education Services),
INGRID SAPIRE (Funda Wande, WITS University)
NWABISA MAKALUZA (Funda Wande)
5 SEPTEMBER 2023

JET EDUCATION SERVICES
THINK EDUCATION. THINK JET.

Bala Wande
Calculating with Confidence

Overview:

- Bala Wande programme
- Early intervention
- Targeting of instruction
- Diagnostic testing
- Criteria for a good test
- Marko-D model
- Rasch analysis
- Sample and fieldwork
- Results
- Recommendations



MARKO-D Collaboration

Bala Wande Programme

Bala Wande is the maths arm of Funda Wande

RCT interventions (EC and LP)

Limpopo intervention:

- Print materials and manipulatives
- Training (4 times per year)
- LTSM arm – no TA
- LTSM+TA – TA per teacher
- Monitoring and support

Collaboration – development of Sepedi version of the Marko-D instrument (evaluation/new language version)
MARKO-D Sepedi piloted in non-programme schools using teacher assistants and test administrators

Importance of early intervention

Early intervention and support for children facing academic difficulties is well documented.

Stormont et al. (2015), Parson and Brynner (1997)

Significant learning gaps as early as grade 3. The researchers advocate that academic gaps be identified as early as grade 1.

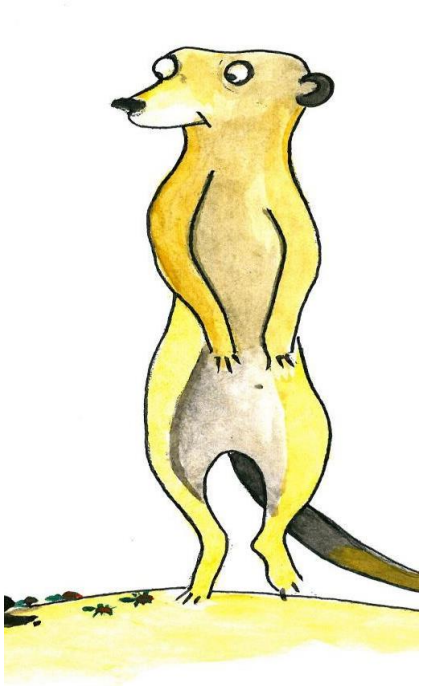
Spuall and Kotze (2015), Bryant (2005), Robinson, Menchetti & Torgesen (2002)

Academic difficulties might become persistent over time.

Racz et al. (2013) Herman et al. (2008), Geary, Hamson & Hoad (2000), Ashcroft, Kruase & Hopko, (2007)

Early intervention may mitigate the development of comorbid conditions like depressive symptoms or anxiety.

Herman et al. (2008), Ashcroft, Kause & Hopko (2007), Desoete (2008)



Link between teacher knowledge and diagnostic testing

Teachers possess 3 categories of knowledge:

1) knowledge of mathematics, 2) knowledge of children's mathematical development and their current position on the developmental trajectory, and 3) knowledge of instructional tasks and tools to aid their progress (Clements and Sarama, 2011)

The contribution of diagnostic test :

A well-designed diagnostic test based on a clear cognitive developmental theory indirectly enhances teaching practices by helping teachers understand the conceptual theory underlying the test and implement effective early identification and support strategies (Fritz, Balzer, Herholdt, Ragpot, & Ehlert, 2014)

The role of the “zone of proximal development” (Vygotsky, 1962, Veresov, 2009):

Properly utilizing diagnostic test results can help teachers identify a child's current level of functioning, enabling them to set appropriate tasks and provide suitable instructional guidance. Failure to do so may lead to a "mismatch"(Clements & Sarama’s , 2011, p. 970) in teaching,

The use of psycho-educational tests

Psycho-educational tests, when combined with informal assessments, observations, interviews, and other data sources, hold promise in comprehending a child's situation (Kilgus et al., 2014; Landsberg, Kruger & Swart, 2011; Sattler, 2008; Donald, Lazarus and Lolwana, 2010)

Validity and reliability

Reliable and validity of measures are central to teachers identifying children requiring extra assistance (Jenkins, Hudson, and Johnson, 2007; Klingbeil, McComas, Burns, and Helman, 2015).

More than one data source

Psycho-educational tests, when combined with informal assessments, observations, interviews, and other data sources, hold promise in comprehending a child's situation (Sattler, 2008; Donald, Lazarus & Lolwana, 2010).

Situational awareness

The influence of socio-economic, cultural, and language backgrounds on children's learning and cognition needs to be carefully considered. E.g. the practical difficulties in implementing **differential teaching in large and under-resourced classes** (Spaull, 2013)

Diagnostic tests, according to Sattler (2008) and Stoiber (2014), play a key role in gauging a **child's proficiency level, guiding effective support and interventions**, and **monitoring progress**. Long and Dunne (2014) contend that diagnostic tests can enhance teaching by providing teachers with an additional resource.

Why the MARKO-D?

South African scholars highlighted **deficiencies in the mathematics education system** (Fleisch, 2008; Schollar, 2015; and Taylor, Van der Berg, and Mabogoane, 2013)

The **necessity of robust measurement** tools for early identification and support is evident (Jenkins et al., 2007).

The **availability of norm-referenced and diagnostic tests** targeting grade 1 students in South Africa is limited, raising questions about their applicability based on criteria of cognitive model alignment, standardization, South African norms, and linguistic diversity.

In this context, the Mathematik und Rechenkonzepte für de 1. Klasse-Diagnose (MARKO-D) test, originally developed and validated in Germany, emerges as a potential solution. The MARKO-D test assesses number concepts and arithmetic skills (Fritz, Ehlert, & Balzer, 2013a; Ricken, Fritz & Balzer, 2013). **The test was validated in Afrikaans, English, Sesotho and isiZulu** from 2013 to 2018 (Henning, Ehlert, Balzer, Ragpot, Herholdt, & Fritz, 2019).

Research question

This study explores the validity of the South African MARKO-D among grade 1 children who are Sepedi home language users.

Special features of the MARKO-D

Individual oral interview based test

Better rapport between child and enumerator. Reduces test-anxiety (Neisworth & Bagnato, 2004)

Test items are presented within a story-based narrative about meerkats and their friends

Enhances child-friendliness and engagement (De Villiers, 2015)

Semi-realistic illustrations are used

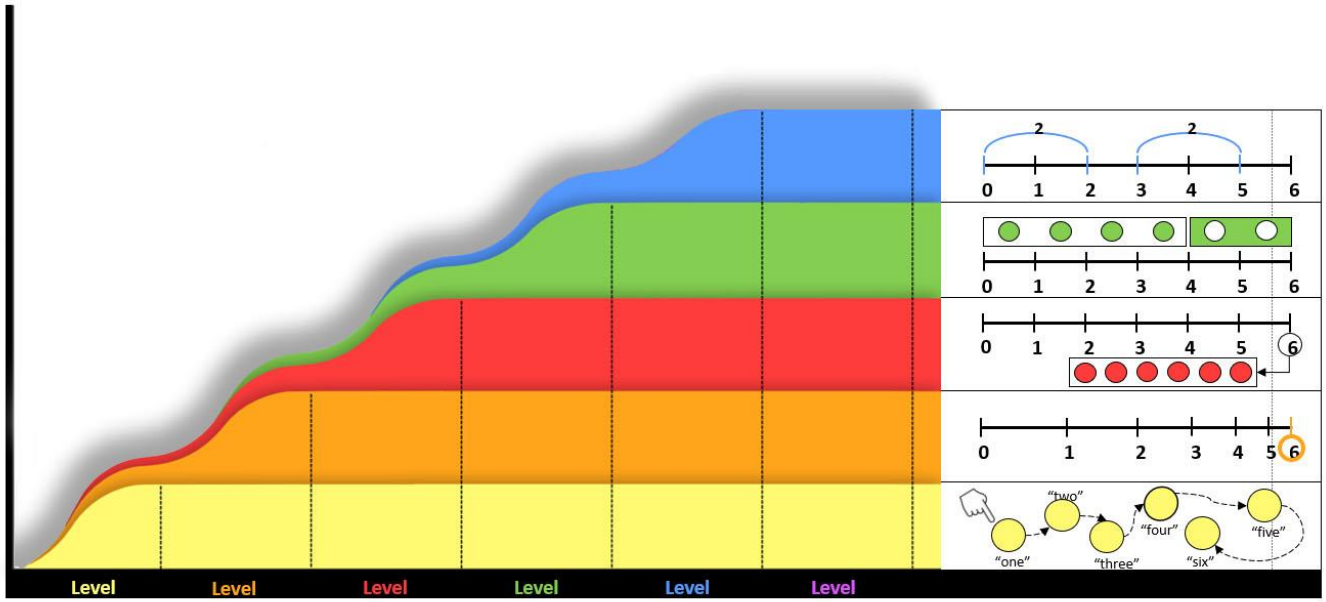
Assists to make the test contextually relevant

MARKO-D is based on an established developmental theory

Aligns with the call for typical cognitive developmental paths (Clement & Samara, 2011)

Conceptual theory of the MARKO-D

Levels in the Marko-D



Fritz and Ricken (2008)

Interpretation of Rasch analysis

1

When it comes to developing a test, it is important to examine assumptions about the relationship **between the performance captured by the test and the latent person's ability,** which actually is of interest but not directly observable.

2

One important characteristic of the Rasch model is in **case of adequate item fit**, it allows the creation of an **interval scale**, representing both item difficulties and person abilities on the same scale. Standardised **infit or outfit values** (MNSQ) close to 1 indicate a good model fit. Higher MNSQ values point to too low selectivity, while too low MNSQ values indicate too high selectivity and thus redundant items in the test. Wright and Linacre (1994) recommend 1 ± 0.5 for less demanding settings and 1 ± 0.3 for the identification of well-fitting items.

3

What one often finds is **the preparation of an extensive set of items** as a starting point of test development, and various empirical tests to prove empirical fit of items. Often, the majority of the initial items is omitted during this process, due to insufficient fit statistics.

4

In the case of the MARKO-D SA, we argue for a different procedure. We use a strong theory of numerical conceptual mathematical competence as starting point. This theory includes a stringent idea of competence levels and competencies necessary for each level. A smaller number of items which really fit the theoretical assumptions are then empirically tested. **Sufficiently fitting statistics as well as an appropriate location on the levels on the scale are needed for each item to be included in the final test.** If one or both of the conditions are not fulfilled, corresponding items are analysed in detail, asking questions

Sample

186 grade 1 learners

Ages ranged from 66 months to 96 months

Third MARKO-D Sepedi pilot

Results

52
items

Wright and Linacre (1994) 1 ± 0.3 for the identification of well-fitting items.

.88

The **person reliability** index in Rasch analysis indicates the replicability of the order of persons on the person-item map if this sample of persons were give a parallel set of items measuring the same construct.

.99

The **item reliability** indicates the replicability of the order of item if the same items were given to a different sample of respondents.

```
MEASURE    PERSON - MAP - ITEM
             <more>|<rare>
     4             +
                   |
                   |  Q42_5                                    11 items
                 X |  Q30_L5
                XX |  Q31_L5
                   |
                 X |  Q25_L5    Q29_L5
     3             T+  Q21_L5    Q23_L5    Q32_L5
                   |  Q26_L4    Q35B_L5
                   |  Q28_L4
                XX |-------------------------------------------
               XXXXX |S                                          Q35A_L5 -take out
              XXXXXX |  Q27_L3    Q35A_L5   Q41_L4    QNEW1_L5   Q27_L5 -take out
                 XXX |  Q39_L4
                   |  Q40_L4    QNEW2_L5
                   |
               XXXXX S|
                 XXX |                                          5 items
          XXXXXXXXXXX |-------------------------------------------
            XXXXXXXX |
               XXXXX |                                          6 items
     1         XXXXX  +
               XXXXX |  Q48_L3    Q4_L2                         Q4_L2 & Q2A_L2 - take out
          XXXXXXXXXXX |  Q2A_L2
                   |  Q16_L3
             XXXXXXX M|  Q47_L3
                 XXX |  Q14_L3    Q15_L3    Q3_L2_____________
              XXXXXXX |  Q33_L3    Q43_L2                        Q3_L2 on L2
     0      XXXXXXXXX  +M Q34_L3
            XXXXXXXXX |  Q12_L2
              XXXXXXX |
           XXXXXXXXXX |  Q11_L3
                   |  Q13_L3    Q2B_L2    Q44_L3
              XXXXXXX |  Q8_L2
             XXXXXXX S|  Q10_L2    Q9_L2                         12 items
    -1          XXXX  +-------------------------------------------
             XXXXXXXX |
                   |  Q46_L1                                   15 items
                 X |
               XXXXX |
                  XX |
                   |  Q45_L1
    -2           XXX  +
                 XX T|
                   |
                   |S Q37_L1
                   |  Q7_L1
                   |
                   |
    -3             +  Q22_L1
                 X |  Q20_L1
                   |  Q36_L1    Q38_L1
                   |
                   |
                   |  Q1_L1
                   |  Q17_L1    Q18_L1    Q6_L1
    -4             +
                   |  Q5_L1
                   |
                   |  Q19_L19
                   |
                   |
                   |  Q24_L1
    -5             +T
               <less>|<freq>
```

Level 5

Level 4

Level 3

Level 2

Level 1

Item-person map

Removal of 4 items before final pilot.

Final validation pilot with 200 plus learners.

Longitudinal study of learner outcome changes over time from grade 1 to 3, focusing on predictive validity and diagnostic utility.

Concurrent validity study with EGMA.

Expansion to isiXhosa and other languages.

Recommendations

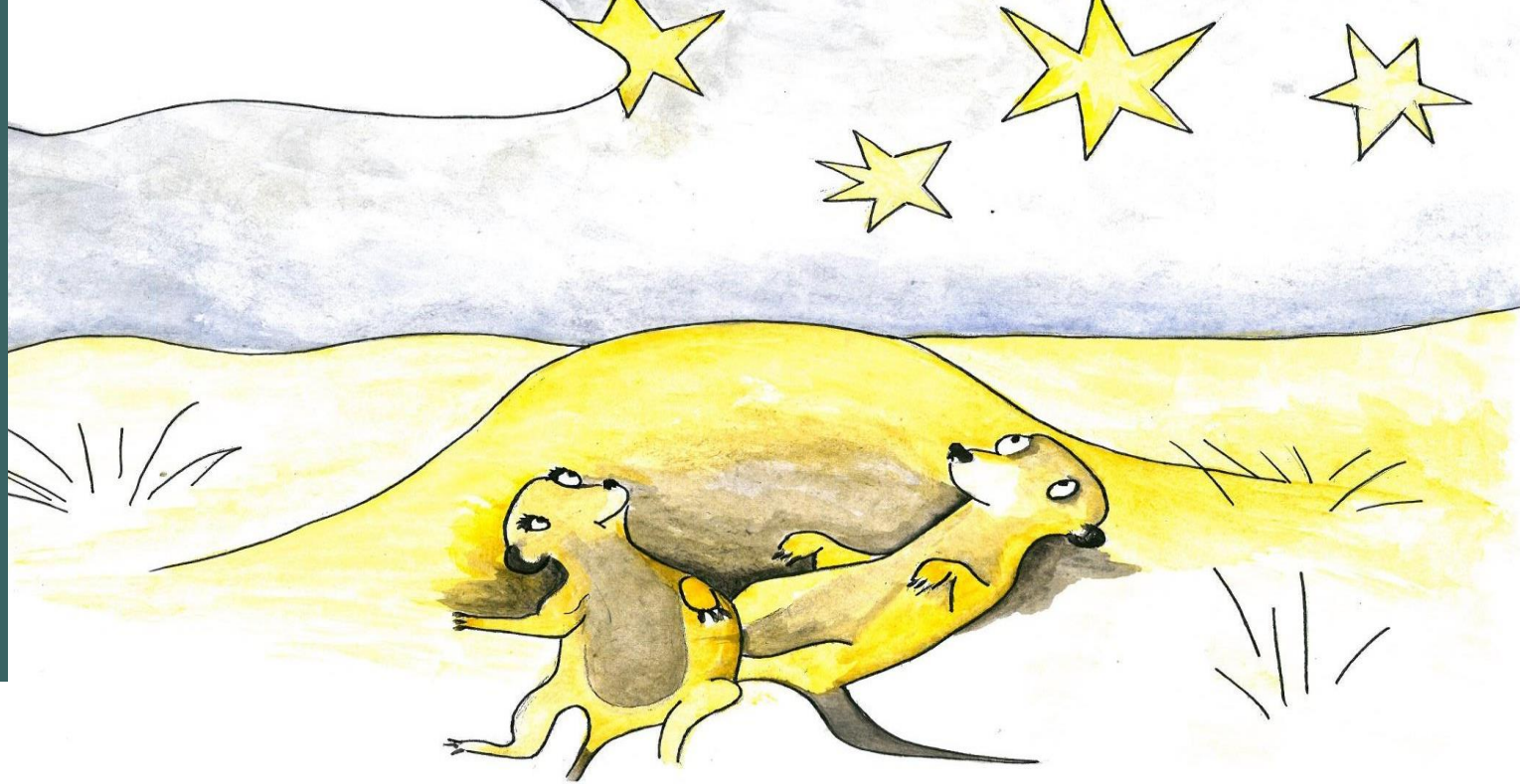
CONTACT US

Roelien Herholdt – rherholdt@jet.org.za
Ingrid Sapire – ingrid@fundawande.org
Nwabisa – nwabisa@fundaqwande.org

011 403 6401

www.jet.org.za

The Education Hub
6 Blackwood Avenue
Parktown
Johannesburg



THANK YOU

JET EDUCATION SERVICES
THINK EDUCATION. THINK JET.

References – Slide 4

Ashcroft, M. H., Krause, J. A., & Hopko, D. R. (2007). Predicting nonadherence to homework assignments in cognitive-behavioral therapy for depression. Journal of Cognitive Psychotherapy, 21(4), 332-337.

Bryant, B. R. (2005). Teaching Students with Learning Disabilities: A Life Centered Perspective. Boston: Pearson.

Desoete, A. (2008). Multi-method assessment of the arithmetical abilities of children with dyslexia. Journal of Learning Disabilities, 41(3), 294-304.

Geary, D. C., Hamson, C. O., & Hoad, D. (2000). Numerical and arithmetical deficits in learning-disabled children: Relation to dyscalculia and dyslexia. Aphasiology, 14(6), 635-647.

Herman, K. C., Lambert, S. F., Reinke, W. M., & Ialongo, N. S. (2008). Low academic competence in first grade as a risk factor for depressive cognitions and symptoms in middle school. Journal of Counseling Psychology, 55(3), 400-410.

Parsons, S., & Bynner, J. (1997). Does Numeracy Matter More? London: Basic Skills Agency.

Racz, S. J., King, K. M., Wu, J., Witkiewitz, K., & McMahon, R. J. (2013). The predictive utility of early childhood disruptive behaviors for school-age social functioning. Journal of Abnormal Child Psychology, 41(5), 679-691.

Robinson, M. M., Menchetti, B. M., & Torgesen, J. K. (2002). Toward a two-tiered model of remediation. In B. R. Foorman (Ed.), Preventing and remediating reading difficulties: Bringing science to scale (pp. 79–91). Baltimore, MD: York Press.

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.

Stormont, M., Herman, K., Reinke, W., King, K., & Owens, L. (2015). Early identification and intervention for students with academic difficulties: South African teachers' perceptions and practices. South African Journal of Education, 35(3), 1-12.

References – Slide 5

Clements, D. H., & Sarama, J. (2011). Early childhood mathematics intervention. Science, 333(6045), 968-970.

Fritz, J. H., Balzer, L., Herholdt, R. F., Ragpot, L., & Ehlert, A. (2014). Perspectives of educators on barriers to academic learning in the Foundation Phase. South African Journal of Childhood Education, 4(2), 46-68.

Veresov, N. N. (2009). Vygotsky's and A. R. Luria's joint research on higher mental functions. In A. Yasnitsky & R. van der Veer (Eds.), Revisionist Revolution in Vygotsky Studies (pp. 99-120). New York: Routledge.

Vygotsky, L. S. (1962). Thought and Language (E. Hanfmann & G. Vakar, Trans.). MIT Press. (Original work published 1934).

References – Slide 6

Donald, D., Lazarus, S., & Lolwana, P. (2010). Educational Psychology in Social Context: Ecosystemic Applications in Southern Africa. Oxford University Press.

Jenkins, L., Hudson, R. F., & Johnson, C. J. (2007). Screening for at-risk readers in a response to intervention framework. School Psychology Review, 36(4), 582-600.

Kilgus, S. P., Riley-Tillman, T. C., Chafouleas, S. M., Christ, T. J., & Welsh, M. E. (2014). Developing a measure of school-based consultation competence for school psychologists. School Psychology Quarterly, 29(1), 35-48.

Klingbeil, D. A., McComas, J. J., Burns, M. K., & Helman, L. A. (2015). Curriculum-Based Measures for Science in the Early Grades: Reading Across Modalities. School Psychology Review, 44(1), 45-62.

Landsberg, E., Krüger, D., & Swart, E. (2011). Addressing Barriers to Learning: A South African Perspective. Van Schaik Publishers.

Long, J. D., & Dunne, F. P. (2014). Functional Assessment of Behavior and Curriculum-Based Measurement in Mathematics for Students with Emotional and Behavioral Disorders: A Preliminary Investigation. Journal of Behavioral Education, 23(2), 258-277.

Sattler, J. M. (2008). Assessment of Children: Cognitive Applications. Jerome M. Sattler Publisher.

Spaull, N. (2013). Poverty & privilege: Primary school inequality in South Africa. International Journal of Educational Development, 33(5), 436-447

Stoiber, K. C. (2014). Math instruction and interventions in an RTI model: Methodologies and evidence-based practices. Routledge.

References – Slide 7

Fleisch, B. (2008). Primary Education in Crisis: Why South African Schoolchildren Underachieve in Reading and Mathematics. Juta and Company Ltd.

Fritz, A., Ehlert, A., & Balzer, L. (2013a). Mathematik und Rechenkonzepte für de 1. Klasse-Diagnose (MARKO-D) – Manual. Hogrefe AG

Henning, E., Ehlert, A., Balzer, L., Ragpot, L., Herholdt, R. & Fritz, A.(2019). MARKO-D SA - Assessment of Number Concept Development, University of Johannesburg.

Ricken, G., Fritz, A., & Balzer, L. (2013). Mathematik und Rechenkonzepte für de 1. Klasse-Diagnose (MARKO-D) – Manual. Hogrefe AG.

Schollar, E. (2015). Finding the Teacher Self: A Study of the Professional Identity of Graduates from a Bachelor of Education Programme. University of the Free State.

Taylor, S., Van der Berg, S., & Mabogoane, T. (2013). Graduation and Dropout Rates in South Africa: A Reconciliation of Data from the Higher Education Management Information System (HEMIS), the National Income Dynamics Study (NIDS) and the Labour Force Survey (LFS). Stellenbosch Economic Working Papers, 13(13).

References – Slide 8

Clements, D. H., & Sarama, J. (2011). Early childhood mathematics intervention. Science, 333(6045), 968-970.

De Villiers, M. (2015). Language as a window into preschoolers' developing number sense. Journal of Numerical Cognition, 1(2), 95-120.

Neisworth, J. T., & Bagnato, S. J. (2004). Child-Directed Play-Based Assessment. Brookes Publishing.

Republic of South Africa. (2006). Children's Act 38 of 2005

References – Slide 9

Dehaene, S. (2011). The number sense: How the mind creates mathematics. Oxford University Press.

Fritz, A., & Ricken, G. (2008). What difference does mathematical thinking make? A conceptual analysis. Mathematical Thinking and Learning, 10(3), 173-195.

Fritz, A., Ehlert, A., & Balzer, L. (2013b). Mathematik und Rechenkonzepte für de 1. Klasse-Diagnose (MARKO-D) – Manual. Hogrefe AG.

Rapin, L. (2016). Mathematical development and plasticity. Child Development Perspectives, 10(2), 85-90.

References – Slide 10 - 13

Anastasi, A. & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River: Prentice Hall.

Linacre, J.M. 2002. What do infit and outfit, mean-square and standardized mean? Rasch Measurement Transactions 16:878.

Wright, B.D. & Stone, M.H. 1979. Best test design. Chicago: MESA Press.

Wright, B.D. & Linacre, J.M. (1994). Reasonable mean-square fit values. Rasch Measurement Transactions 8:370.