



EXAMINING SUCCESSFUL  
NUMERACY PROGRAMS:  
EVALUATING LEARNER  
MATHEMATICS  
OUTCOMES IN AN RCT IN  
SOUTH AFRICA

*CIES Washington: February 21, 2023*

*Ingrid Sapiro and Roelien Herholdt*

*Funda Wandu; University of the  
Witwatersrand; JET*

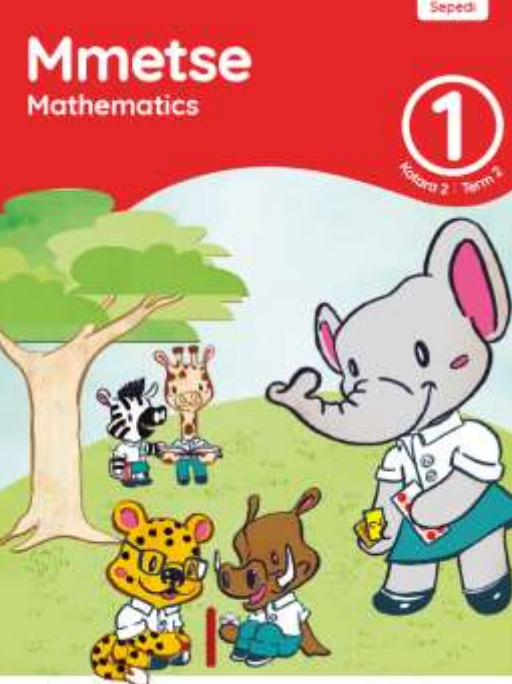
# BACKGROUND

- Early grades maths gaining attention in South Africa (tertiary, NGO, DBE)
- Bala Wandé programme – the maths arm of the Funda Wandé programme (started 2018)
- Active in 3 provinces in SA (Limpopo, Eastern Cape, Western Cape)
- Improve maths learning outcomes in low fee schools
  - Multilingual materials support (Teacher Guide, Learner Activity Book, Dictionary)
  - Manipulatives (concrete and print materials, integrated design)
  - Support (TA, mentor, coach, subject advisor)

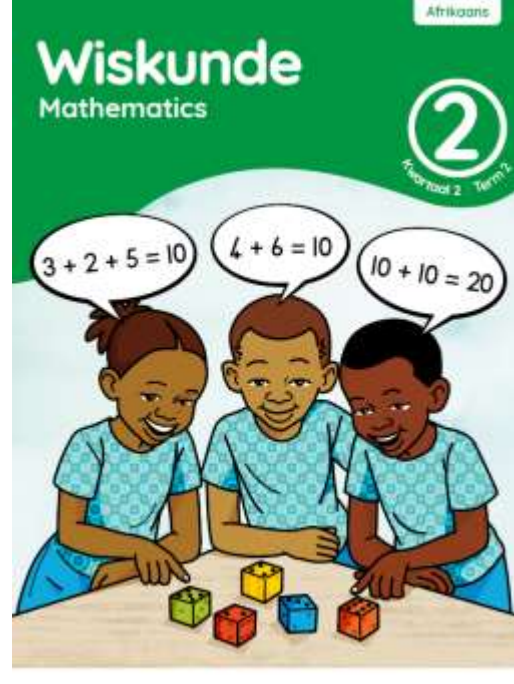


# **Bala Wande**

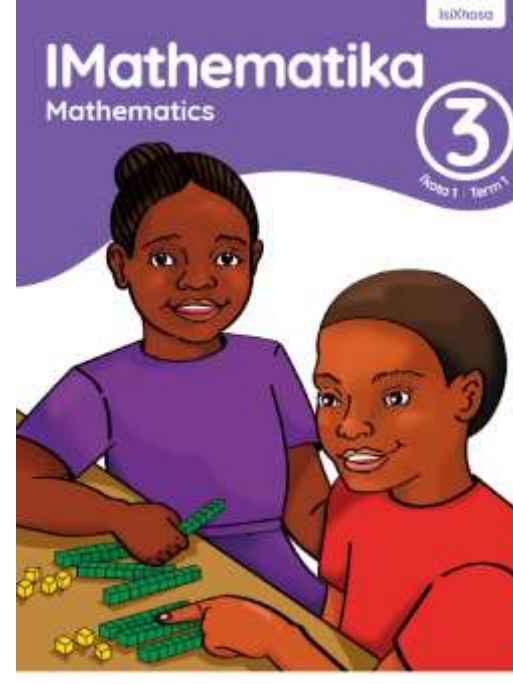
Calculating with Confidence



Morutwana : Learner: \_\_\_\_\_



Leerder : Learner: \_\_\_\_\_

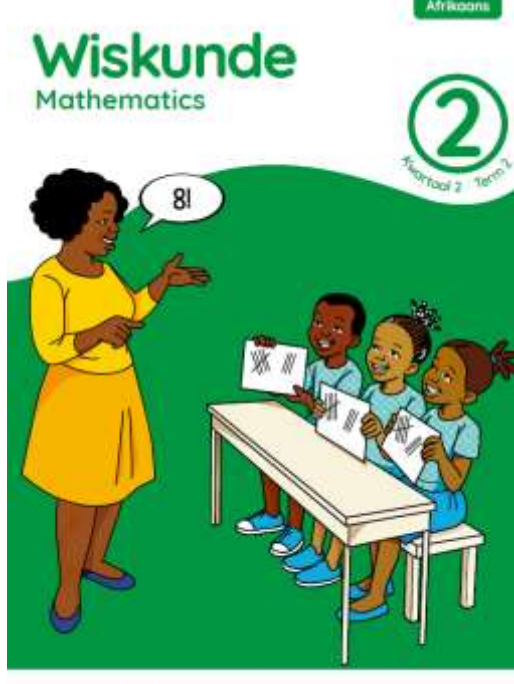


Umfundi : Learner: \_\_\_\_\_

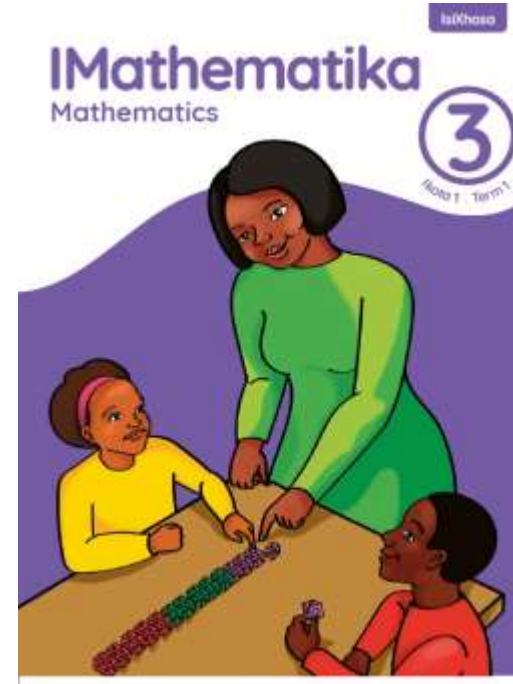
# LEARNER ACTIVITY BOOKS, TEACHER GUIDES, DICTIONARIES, MANIPULATIVES, VIDEOS



Morutlhi : Teacher: \_\_\_\_\_



Onderwyser : Teacher: \_\_\_\_\_



Utitshaka : Teacher: \_\_\_\_\_



## LIMPOPO INTERVENTION

- The impact evaluation in Limpopo uses an RCT with 120 no-fee schools in Capricorn North and Capricorn South districts randomized into one of three arms (40 schools per arm/2400 learners).
  - Materials only (LTSM)
  - Materials and TA support (LTSM+TA).
  - Control group
- The LTSM and LTSM+TA arms also receive **training** and are monitored by the programme team.

# “PREVENTION IS BETTER THAN CURE”

Numeracy skills are important – in formal and informal contexts.

Mathematical learning begins before formal schooling begins.

Programme goal –  
improve maths  
learning outcomes

Implementation  
evaluation  
(external)

Baseline &  
Midline 1

- One-on-one numeracy assessment; EGMA (hybrid) – written – 10 learners

Midline 2

- EGMA (hybrid); MARKO D (pilot validation study)

Endline

- EGMA (hybrid)

## MARKO-D BACKGROUND

- Early mathematical foundations are critical.
- The Marko D test is based on an **empirically validated model of children's progressive understanding of numerical concepts**. The model was developed in Germany, based on theoretical suppositions and empirical data (Fritz, Ehlert & Leutner, 2018).
- In the model, number concept is viewed as a requirement for the construction of arithmetical skills.
- **These concepts gradually build upon each other, creating a continuum (or a pathway of learning progression) of increasing complexity.**
- This is in line with the South African school curriculum for the foundation phase (elementary school) that requires basic number concepts.

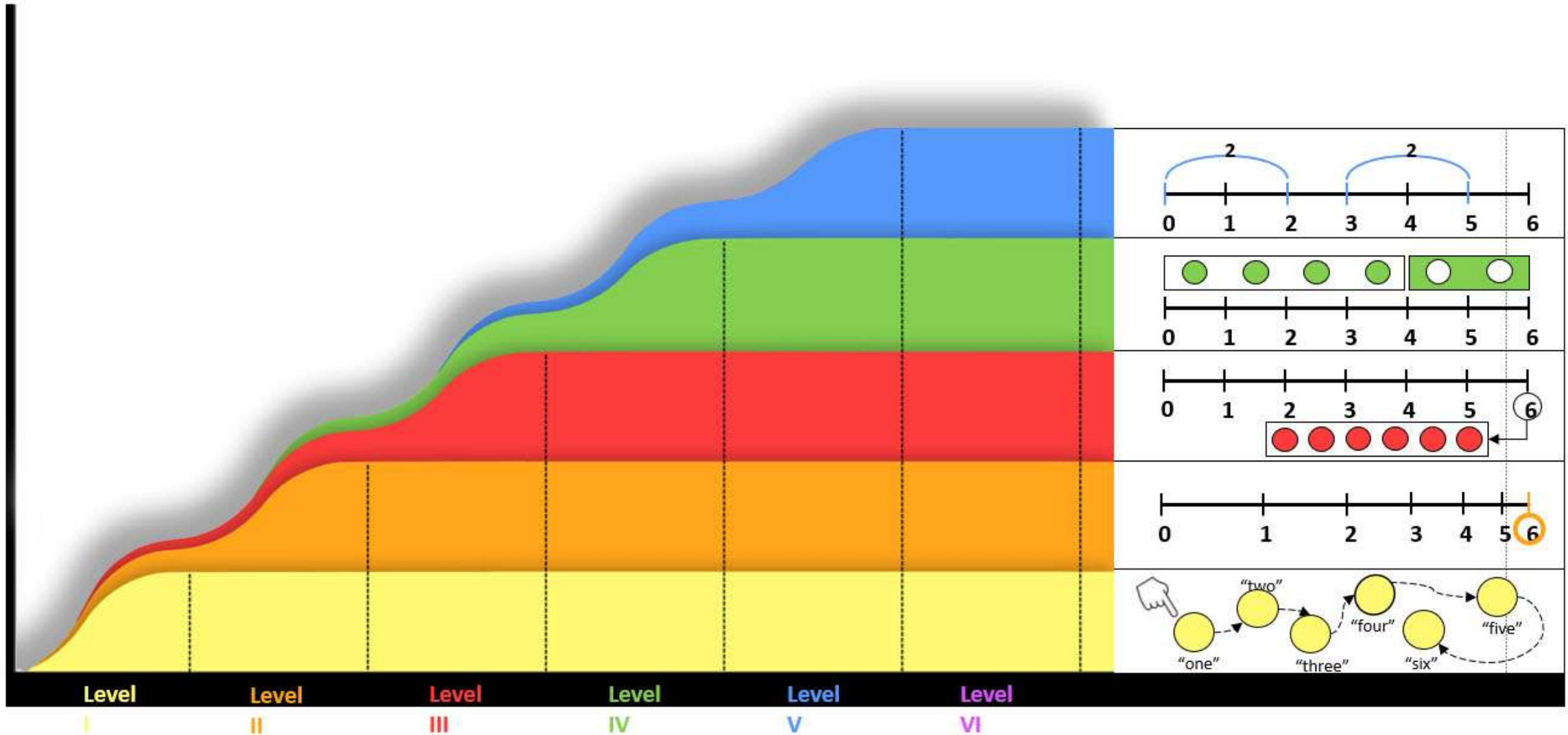
# DEVELOPMENT OF THE SA MARKO-D

- Translated items should retain the conceptual content of the original test.
- Several translation iterations until equivalence of items was established.
- English, isiZulu, Afrikaans and Sesotho.
- The challenge was to find out whether the item did not 'fit' the model because of level inconsistency, translation, or cultural relevance.
- Special mention should also be made of the use of drawings of meerkats as characters in the story that forms the framework for the test.
- The final test contained 48 items. The items form a one-dimensional cumulative scale with five distinguishable segments, according to levels of the theoretical model. The sequence of these segments on the scale follows the sequence of levels in the model.





# Levels in the Marko-D

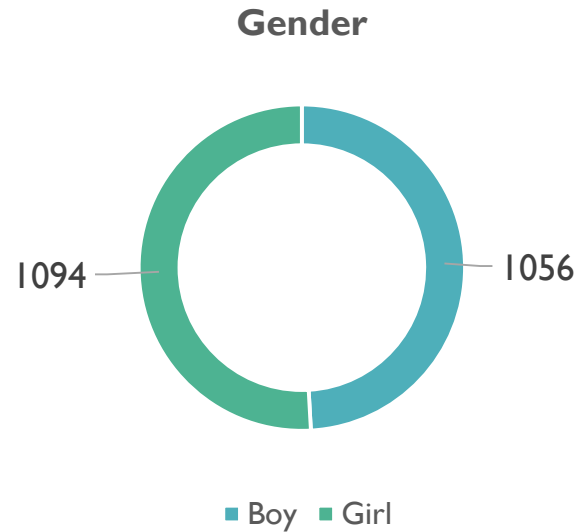


# BALA WANDE - MARKO D COLLABORATION

- Sepedi version development
  - Round 1 pilot (May 2022) – Bala Wandé team
  - Round 2 pilot (Oct 2022) – External evaluation team
- Validation process
- Validation contribution to evaluation
- Sepedi instrument development post pilot (add to English, IsiXhosa, Afrikaans developed so far in South Africa)

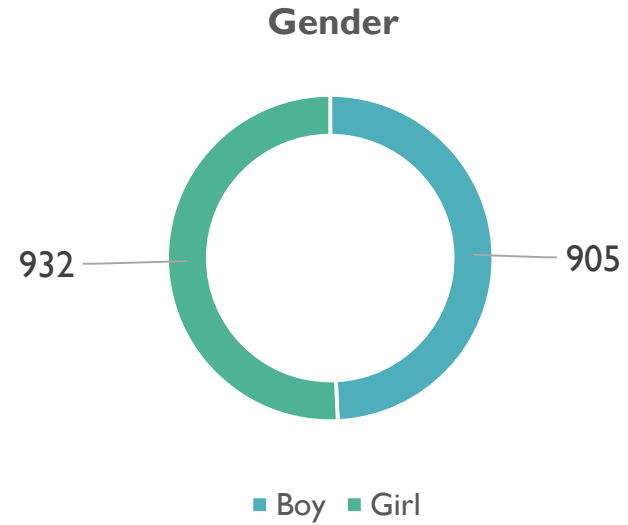
# TEST SAMPLE

- n=2510 Sepedi speaking grade 1 learners
- Tested in 3<sup>rd</sup> term of 2022
- Age range:
  - from 5 to 10 years
  - Or 60 to 95 months
- Slightly more girls than boys



# VALIDATION SAMPLE

- n = 1837 Sepedi speaking grade 1 learners
- Tested in 3<sup>rd</sup> term of 2022
- Age range:
  - from 6y 0m to 7y 11m
  - or 72 to 95 months
- Slightly more girls than boys



# RASCH ANALYSIS

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 Stone (1999) suggest the range of  $1 \pm 0.2$  as limiting values for MNSQ for tests in the high-stakes region, and Wright and Linacre (1994) recommend  $1 \pm 0.5$  for less demanding settings and  $1 \pm 0.3$  for the identification of well-fitting items.

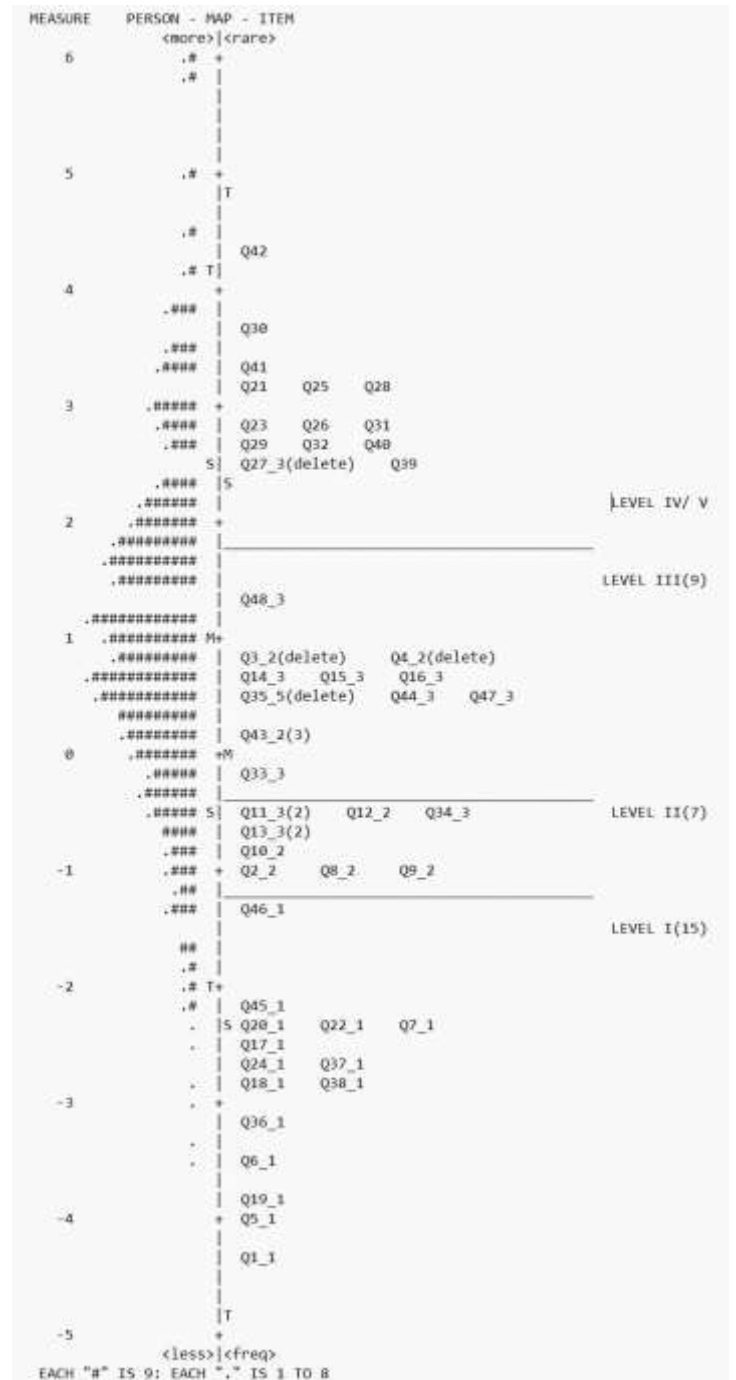
**The items of the one-dimensional dichotomous Rasch model show satisfactory values (weighted infit MNSQ  $1 \pm 0.2$  for 46 out of 48 items; weighted infit MNSQ  $1 \pm 0.3$  for 2 items.)**

# RELIABILITY

- Reliability addresses the extent to which the results are free from measurement error.
- 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. This assumes that the sample is big enough for persons to be spread along the ability continuum but also suitable and sufficient items that would demonstrate such a hierarchy of ability.
- The **item reliability** indicates the replicability of the order of item if the same items were given to a different sample of respondents with the same number of respondents and that behaved in the same way.
- The person reliability was 0.9 and the item reliability 1.0.

# ITEM-PERSON MAP

- LHS - [#] represents 9 persons and [.] represents 1- 8 persons (achievement spread)
- RHS - Items per level (items for exclusion/re-testing)
- Normal distribution (both sides)



# EXAMINING SUCCESS

- What works to improve maths learning and teaching?
- How can we use evaluation studies to inform the way forward?
- Collaboration is the key!

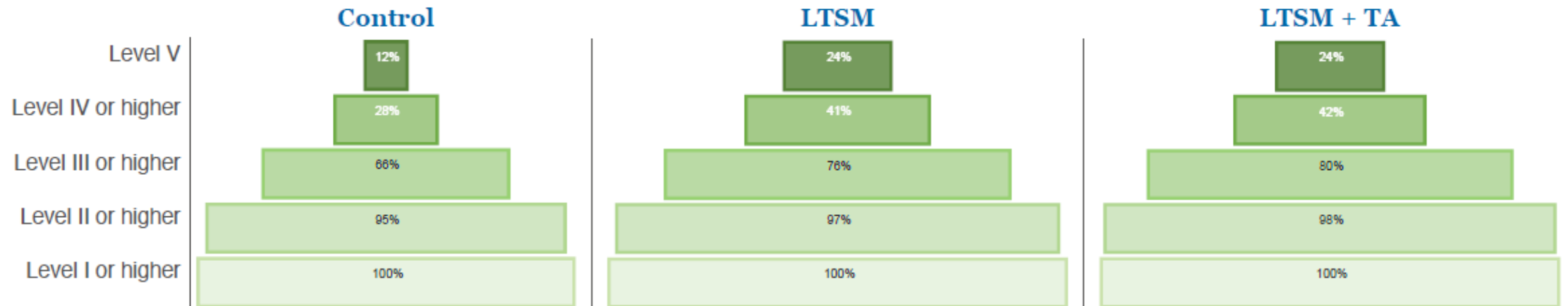


# MIDLINE FINDINGS: COMPOSITE SCORES

	Effect (standard deviations)		Effect (years of learning)		Significance of differences		
	LTSM	LTSM+TA	LTSM	LTSM+TA	LTSM vs Control	LTSM+TA vs Control	LTSM vs LTSM + TA
Grade 1:							
EGRA composite	-0.041	0.21		54%		*	**
EGMA composite	0.18	0.42	45%	106%	**	***	***
MARKO-D composite	0.344	0.367	86%	92%	***	***	
Grade 2:							
EGRA composite	0.22	0.50	54%	126%	*	***	**
EGMA composite	0.26	0.50	66%	124%	***	***	**



# MIDLINE FINDINGS: MARKO-D LEVELS



**MARKO-D SA**

ASSESSMENT OF NUMBER CONCEPT DEVELOPMENT

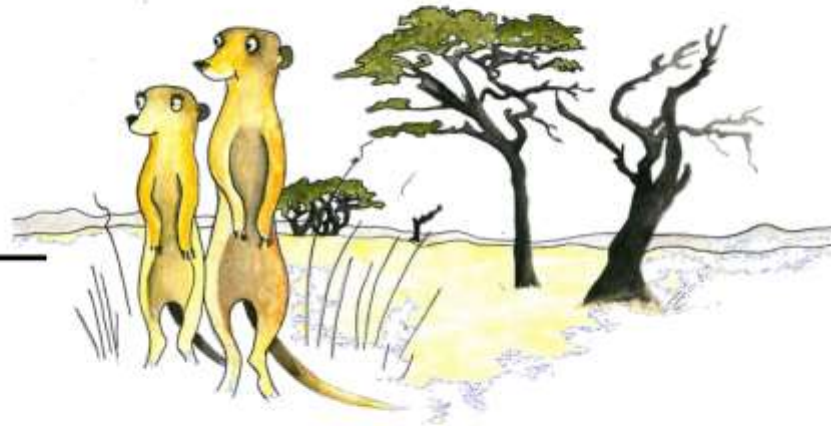
---

Herholdt, R., Ardington, C., & Sapire, I.

Instructions (**English/Sepedi**)

 **hogrefe**

---



THANK YOU

Ingrid Sapire

[ingrid@fundawande.org](mailto:ingrid@fundawande.org)

[ingrid.sapire@wits.ac.za](mailto:ingrid.sapire@wits.ac.za)

Roelien Herholdt

[rherholdt@jet.org.za](mailto:rherholdt@jet.org.za)

