



# Managing large-scale data collection: Experience and challenges (Cas H Prinsloo)

Launch of Early Grade Reading  
Study North-West Findings,  
(16-)17 August 2017,  
DBE Conference Centre



Education and Skills  
Development

# Overview

- Introduction - Brief
- Overall perspective:
  - Description of experience
  - Identification of important components
- Illustration using selected EGRS components
- Areas for future consideration, including challenges

# Introduction – Brief

- Practical realities of large-scale data collection
- Overall successes
- Challenges
- Areas of consideration for future research / studies
- Scoping the input across overall HSRC work and the EGRS study specifically

# Overall perspective - Experience

Study	Funder	Client	Size
Personality tests	Baseline, ESKOM	Psychologists	8-10 yrs, n=2-10 000 (→2003)
Qual Learn Proj	Business Trust	JET (DoE)	70 interv/16 contr; national, 5yrs
Mveledzandivho	BHP Billiton	DoE	28 schools, 5 provinces
LDoE Literacy	Irish Aid	Limpopo DoE	25 schools, FP, qualitative only
Teacher leave	UNICEF	DoE	50 schools, national
Gr8 achievemnt	WCED	WCED	+/- 87 500 learners; Mths, Lang
Paradigms/pract	Nat Res Fnd-NRF	DBE (GP, LP)	Lit/lang t&l, +/- 10, qualitative
Extra tuition	Shuttleworth Fnd	WCED	Impact evaluation, 8 schools
Lrn timer regeneratn	USAID	Limpopo DoE	EFAL impact evaluation, 36 sch
TIMSS '11 & '15	DBE	DBE	Mths/Sc Gr 9, 12 500 lnrs, nat
Setswana Teach EGRS I	3ie	DBE (-NW)	RCT, 230 schools, impact eval, 4 600 learners
Few other	DoL, etc.	Miscellany	Non-education

# Overall perspective - Components

- **Contractual:** budget, project proposal, legal agreements, multi-stakeholder interaction\* and platforms (\*typically Steering Committee – funder, client, partners, recipient/s)
- **Team capacity:** Fields, proficiencies, languages
- **Design and methodology:** involved, complex, varied
- **Baselines:** In advance (if not → retrospective routes?)
- **Ethics clearance/s:** gatekeeper permissions; informed consent of participants (if minors, added implications)
- **Instruments:** development, theory of change, constructs
- **Research materials:** production, with unique identifiers
- **Logistics:** Packing, batching, routing, teams and travel

# Overall perspect. (2) - Components

- **Data collection:** Administration manual, training, quality assurance
- **Procurement:** printing, data collection, & capturing (tender lead time – 4 months)
- **Marking / scoring:** Recruit, train and moderate
- **Data management:** code instruments for capturing; instrument preparation for capture, training, QA, cleaning
- **Analysis and report writing:** proficiency and time; executive/summary, short, & technical versions
- **Dissemination:** client presentations, policy briefs, conference papers, peer reviewed articles, research reviews, media (radio, TV; including local and regional)

# Illustrations from EGRS I

- A model project: for design (RCT); multi-institutional, -sectoral and -stakeholder interaction; theory of change (Pre-Analysis Plan – PAP)
- Individual oral assessment at Foundation Phase level (even at Gr R, & Jan/Feb of Gr 1): non-existent literacy proficiency and reading culture; test anxiety
- Setswana linguistics expertise: in-house, departmentally, and externally
- Data collectors: recruitment and logistical capacity
- Dissemination / implementation networks: Policy Influence Plan (PIP); provincial recipient and national department platforms; conferences; peer-reviewed and other articles; international opportunities



# Challenges, future considerations

- Preferred service providers:
  - Technical expertise / track record (minimum required experience, demonstrated capacity) → recruitment, selection, training @
  - Rotated procurement / open tendering (for capacity development purposes and broadening the pool): → data-quality risks
- Deep, intensive linguistic\* foundations and underpinnings
- Time-consuming periods: procurement, instrument development, data analysis, and report writing
- Assessment issues: standardised instruments, norms, individual oral mode (time demands, text anxiety)
- Impatience in relation to ready policy inputs, not allowing sufficiently for deep evidence base, and rigour
- Size/depth of capacity pool; institutional infrastructure

\* And cognitive neuroscience



Thank you!

Questions / Discussion

# Theory of change - basis

- Reading acquisition - recent UNESCO report (Wagner, 2011):
  - Alphabetical principle: knowledge of letter names and sounds
  - Phonemic awareness: relation between phonemes and graphemes; decoding skill
  - Oral reading fluency (ORF): coordination of several automated decoding skills developed through practice
  - Vocabulary: understanding words in oral or written form (receptive and expressive components or modes)
  - Reading comprehension: retention of meaning.

Wagner, D.A. (2011). Smaller, quicker, cheaper: Improving learning assessments for developing countries. Paris: UNESCO / International Institute for Educational Planning.