Scalable **TR**ansdiagnostic **E**arly **A**ssessment of **M**ental Health

STREAM: Developing a Tablet-Based App to Assess Early Child Development in Low-Resource Settings

Gareth McCray & Nick Thompson

SEED Lab - Harvard Graduate School of Education

16th October 2023

MRC Medical Research Council

GCRF Global Challenges Research Fund

one Background

two The STREAM App

three Study Design

four Data Analysis

one

Background

Early Child Development

- Early child development and mental health form the foundations for positive outcomes in later life.
- Developmental disabilities reflect a range of conditions that impact a child's physical, cognitive, and/or behavioural functioning (Centers for Disease Control and Prevention, 2018).
- As of 2016, it is estimated that globally 8.3% of children under the age of 5 years (52.9 million) are living with developmental disabilities (Olusanya et al., 2018).

Neurodevelopmental Disorders in LMICs

- Increased NDD risk may relate to early adversity, including poverty, violence, inadequate hygiene and stimulation, perinatal issues, and limited access to good nutrition (Bitta et al., 2017; Sen, 2009; Toso et al., 2020; Byrne et al., 2017; Milner et al., 2015; Shonkoff et al., 2012).

- These risk factors can have detrimental effects on physical and mental health, hinder cognitive development, and lead to poor long-term outcomes (Shonkoff et al., 2012).

- It is estimated that around 95% of children with NDDs (50.2 million) reside in low and middle-income countries (LMICs) (Olusanya et al., 2018).

Early Detection & Intervention

- Developmental disabilities often go undetected in these settings (Dasgupta et al., 2016). Early detection of developmental disabilities in LMICs is complicated by various issues, notably:

  1) A reliance in many LMICs on time-intensive, specialist-dependent, proprietary assessment tools;

  2) scarcity of highly skilled health professionals, who tend to be concentrated in urban, often private, clinics.

- This presents a significant challenge to the implementation of effective interventions, especially during early childhood when brains are maximally plastic and responsive to change (Patel et al., 2018).

The STREAM Project

- There is a need for tools that address these issues; namely, tools that are open access, fit for use in the field by non-specialist-workers with minimal training, applicable across diverse cultural settings, and that cover the entire early developmental period (Olusanya et al., 2021).

- STREAM is a comprehensive tablet-based platform designed to be transdiagnostic, modular, and scalable. It aims to assess developmental abilities in 0-6 years old children.

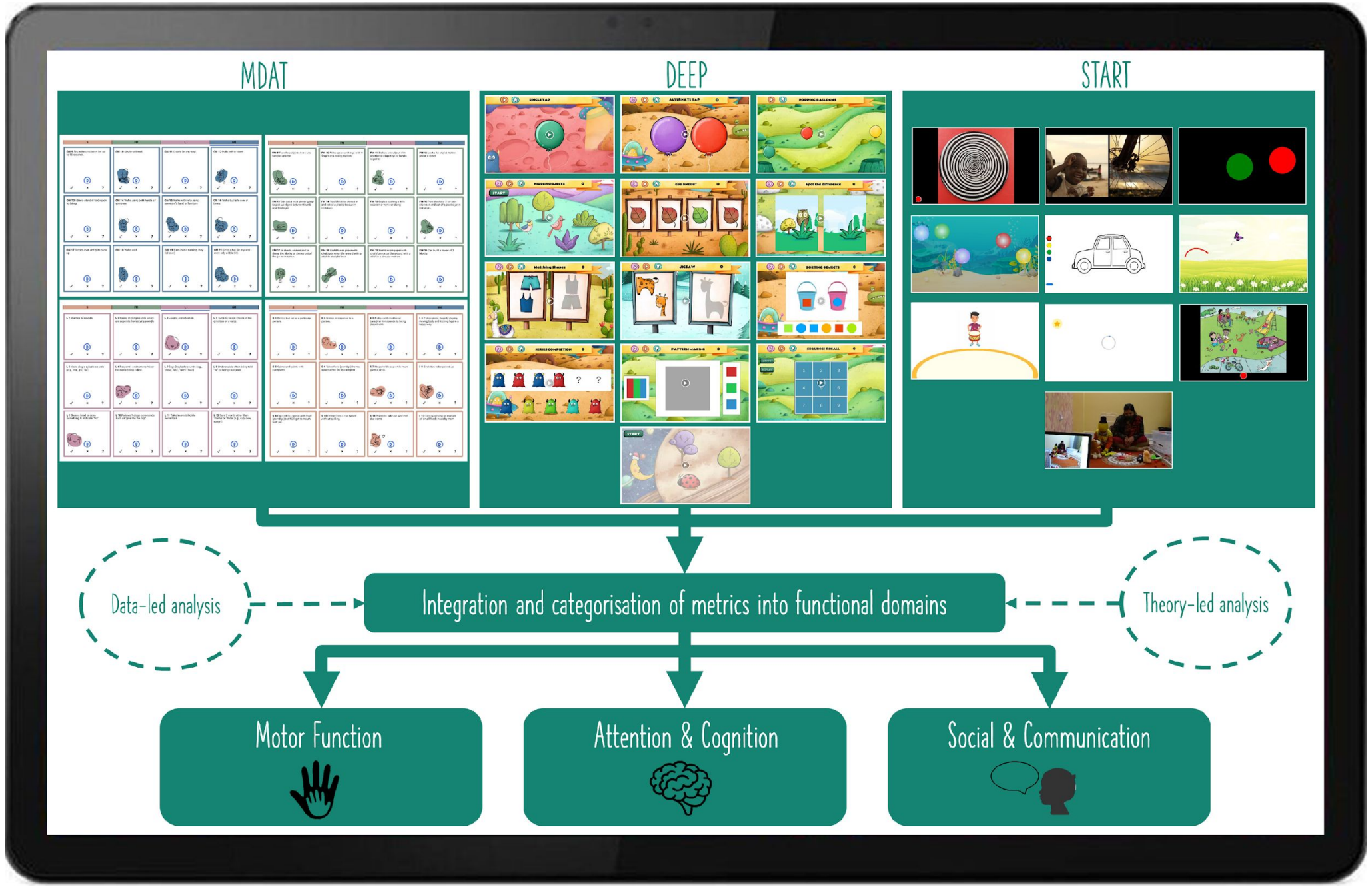
- The hope is that its implementation through public health initiatives in LMICs could facilitate the early detection of NDD risk.

Project Aims

- Develop a tablet-based application that assesses developmental abilities in children aged 0-6 years.
- Evaluate the validity and reliability of the app in measuring the developmental abilities of children and identifying children who may be at-risk of NDDs.
- Use the app to establish preliminary developmental reference data for both contexts.

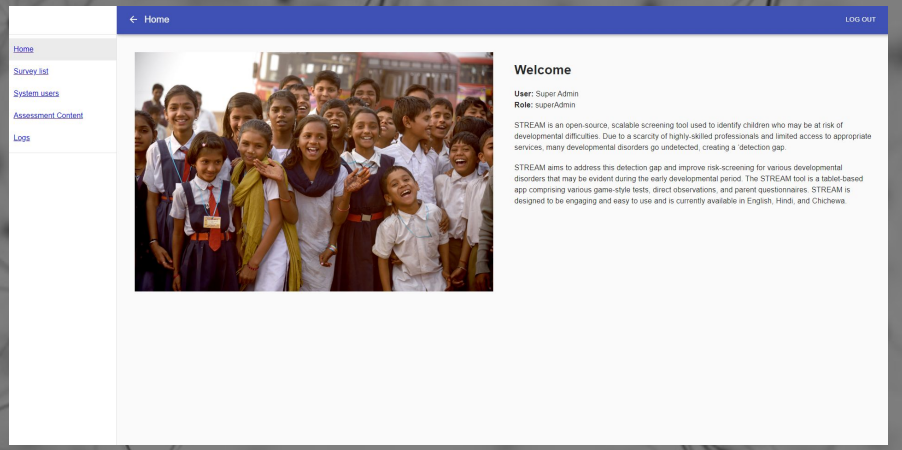
two

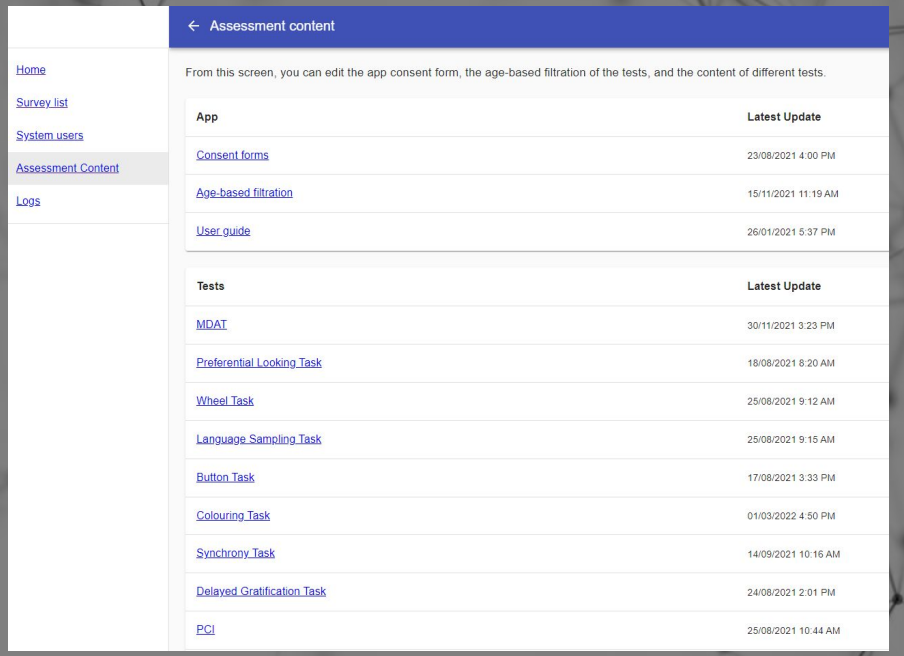
The STREAM App


MDAT
DEEP
START
Data-led analysis
Integration and categorisation of metrics into functional domains
Theory-led analysis
Motor Function
Attention & Cognition
Social & Communication

App Features

- The STREAM app includes:
  - Gamified tasks
  - Direct and passive observations
  - Parent-report
  - Audio and video recording functionality
  - Flexible survey engine
- The app works entirely offline, facilitating ease of use ‘in the field’.
- Dedicated backend server to securely store data and adapt app content (e.g., stimuli, language, age filters).





MDAT

- MDAT = Malawi Developmental Assessment Tool (Gladstone et al. 2010).

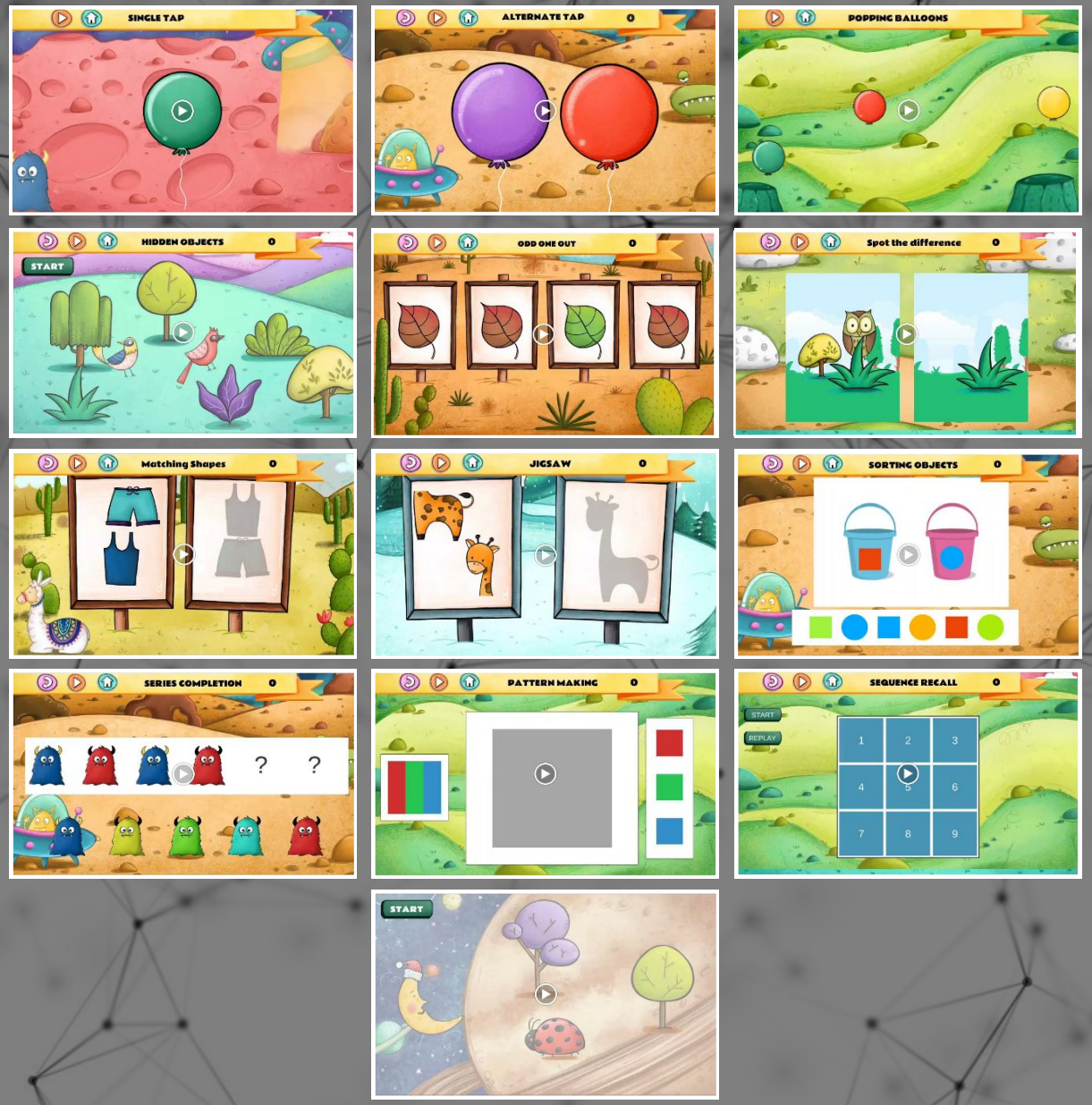
- MDAT is a milestone-based assessment tool which combines observational-performance items with parent-report items, covering four broad domains:

  - Gross motor
  - Fine motor
  - Language
  - Social

- Applicable across the full 0-6 years age range.

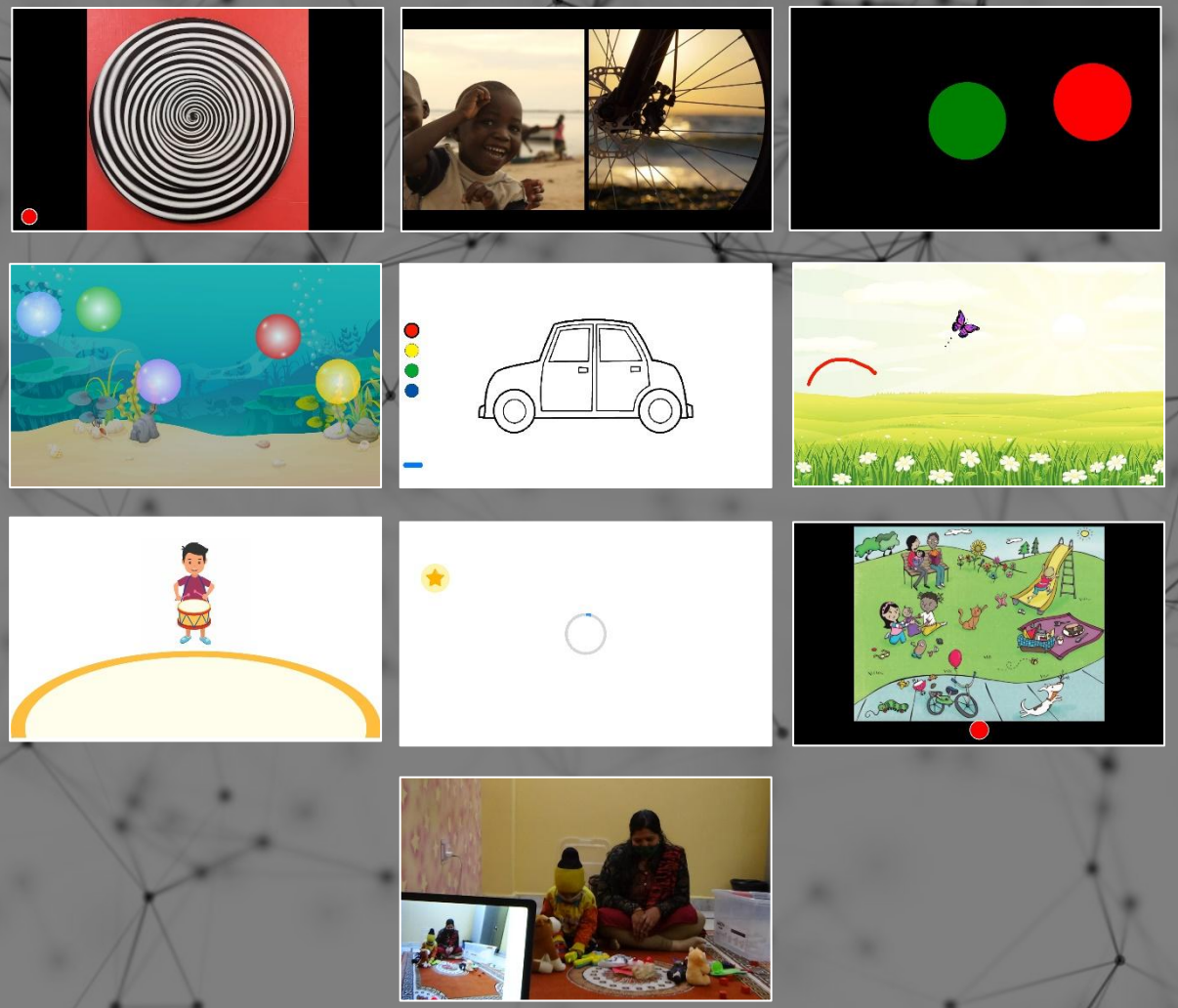
DEEP

- DEEP = DEvelopmental assessment on an E-Platform (Bhavnani et al., 2019).
- Comprises various games which are integrated into an overarching storyline to maximise engagement.
- Assesses cognitive abilities in children aged 2.5 to 6 years, including:
  - manual speed and coordination
  - inhibitory control
  - visual perception and integration
  - reasoning
  - categorisation
  - memory

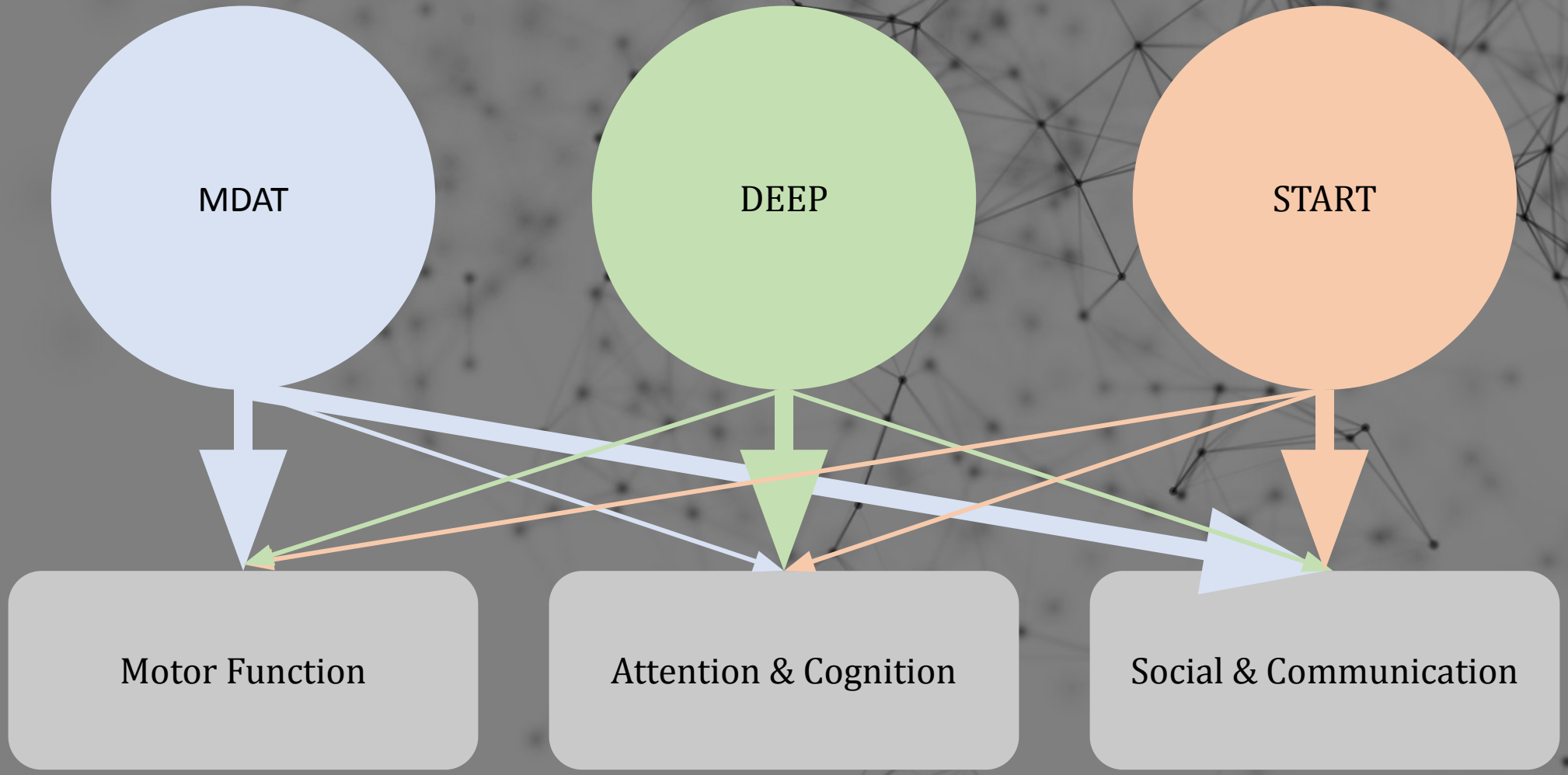


START

- START = Screening Tool for Autism Risk using Technology (Dubey et al., 2023).
- Mobile, open-source platform for early detection of autism risk in children aged 2.5–6 years.
- Assesses several domains associated with the autistic phenotype, including social functioning, sensory processing, motor skills, expressive language, as well as parent-child interaction.
- It has shown high classification accuracy (78%) in identifying children at-risk of autism (Dubey et al., 2023).



Domains Assessed



three

Study Design

Study Overview

- Children Aged 0-6 years in India (N=2000) and Malawi (N=2000)
- Primarily semi-urban catchment areas
- Two recruitment samples: community and enriched
- Community sample aims to be representative not ‘ideal’
- Not epidemiological study – only representative of where we recruit

Study Groups

Community

India: Children (N=1850) aged 0 to 6 years will be recruited from the South East District of New Delhi, covering urban and rural settings.

Malawi: Children (N=1850) aged 0 to 6 years will be recruited in Blantyre, Malawi. Recruitment efforts will focus on Health Centres.

Enriched

India: children at-risk of, or diagnosed with, a NDD (N=150) will be recruited at tertiary clinical centre specialising in the assessment and care of children with developmental difficulties.

Malawi: In Malawi, children at-risk of, or diagnosed with, a NDD (N=150) will be recruited from paediatric wards, paediatric neurology clinics, physiotherapy clinics, occupational therapy clinics, psychiatric clinics, and malaria follow-up clinics.

Inclusion and Exclusion Criteria

Inclusion Criteria

1) Aged 0-6 years

2) Resident within the catchment area of the study sites

3) Mother/caregiver willing and able to provide written informed consent

Exclusion Criteria

1) Acutely unwell at the time of testing

2) Severe vision/hearing/motor problems (child is unable to interact with tablet device)*

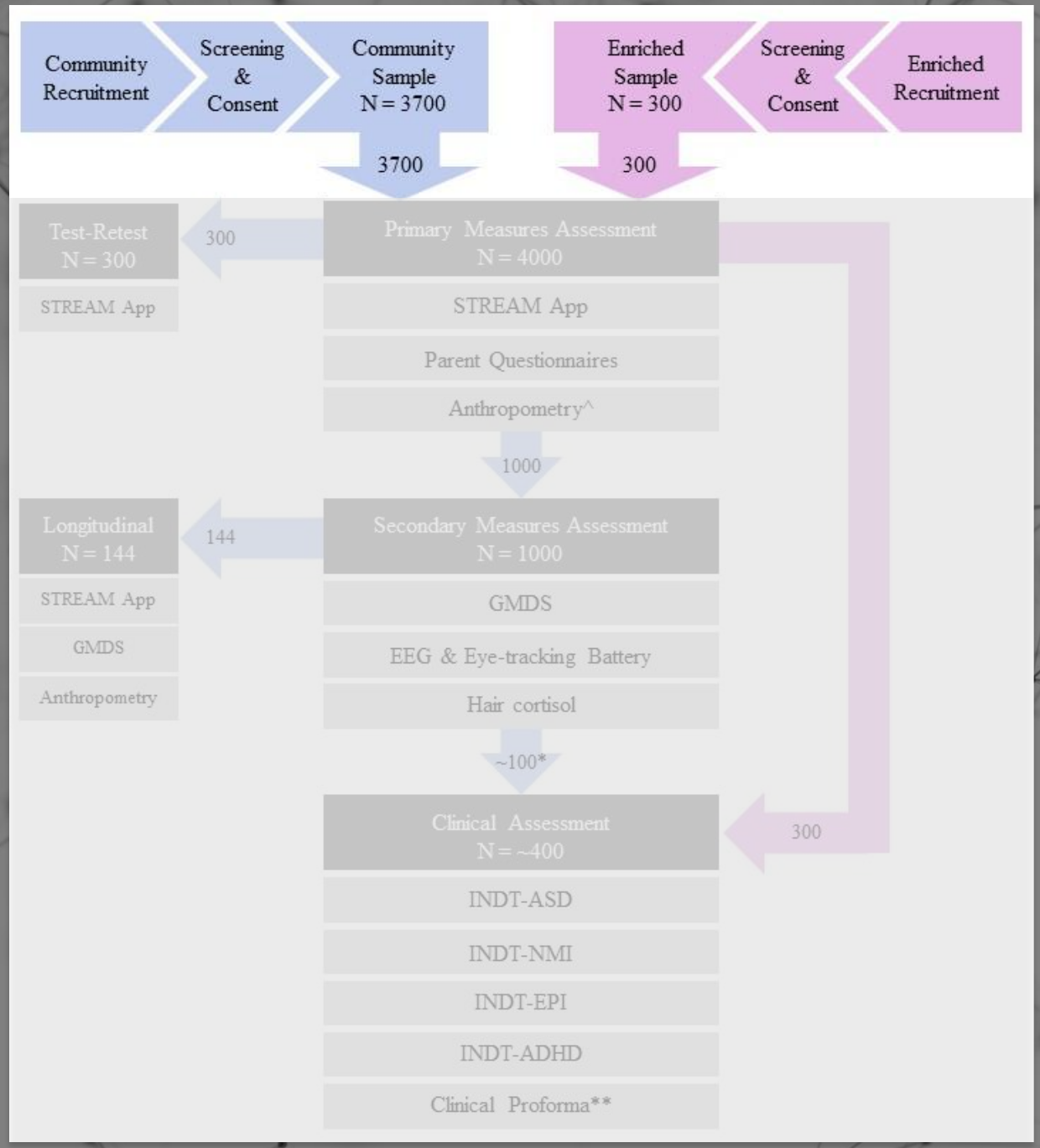
3) Uncontrolled seizure in the past 2 weeks

4) Sibling has already participated in the study

5) Parent has current severe learning disability or severe psychiatric disorder

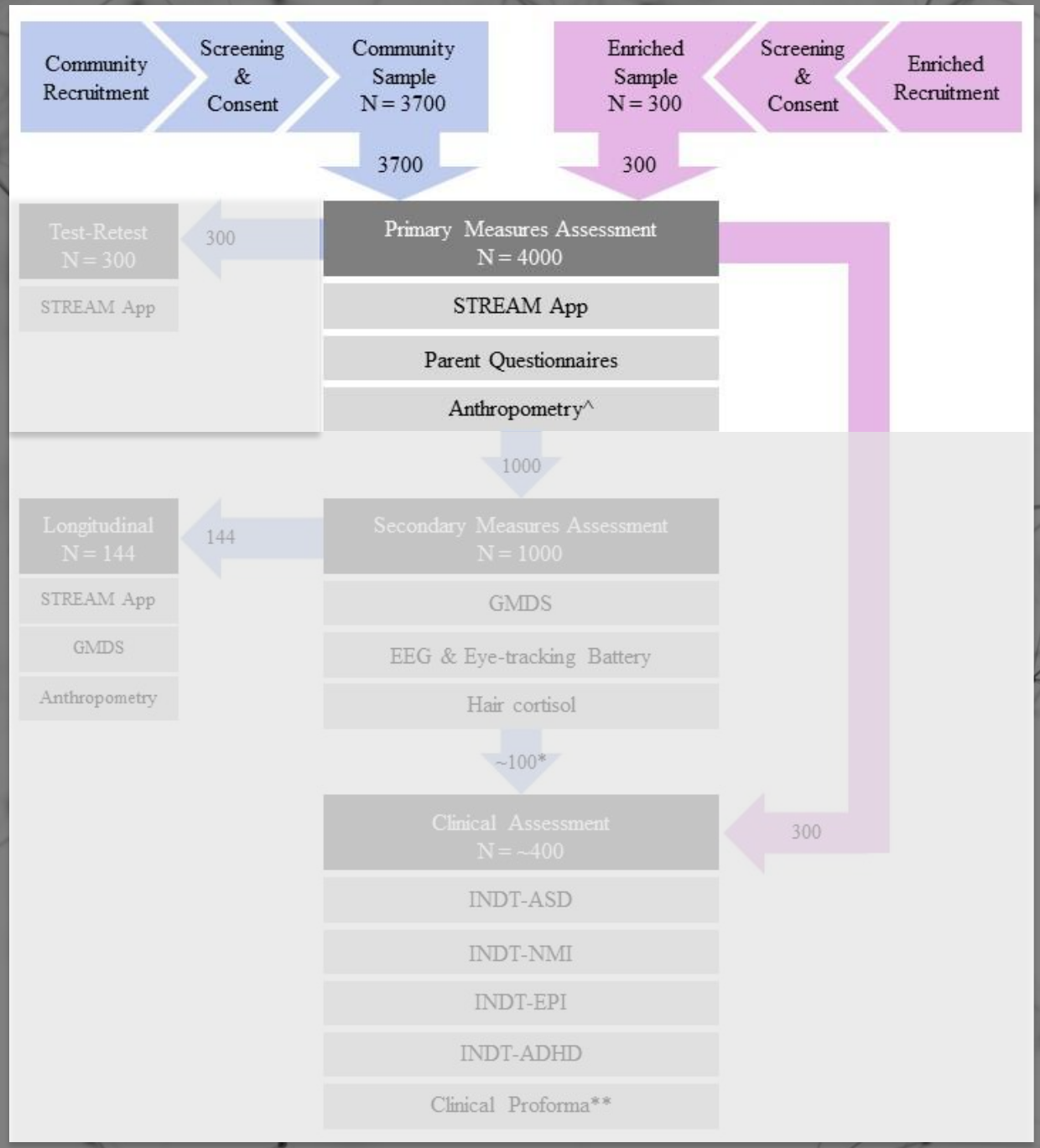
Recruitment Flow (1)

- Community sample Children are recruited from the community locations
- Enriched sample children are recruited from the enriched locations
- Recruitment is currently ongoing.



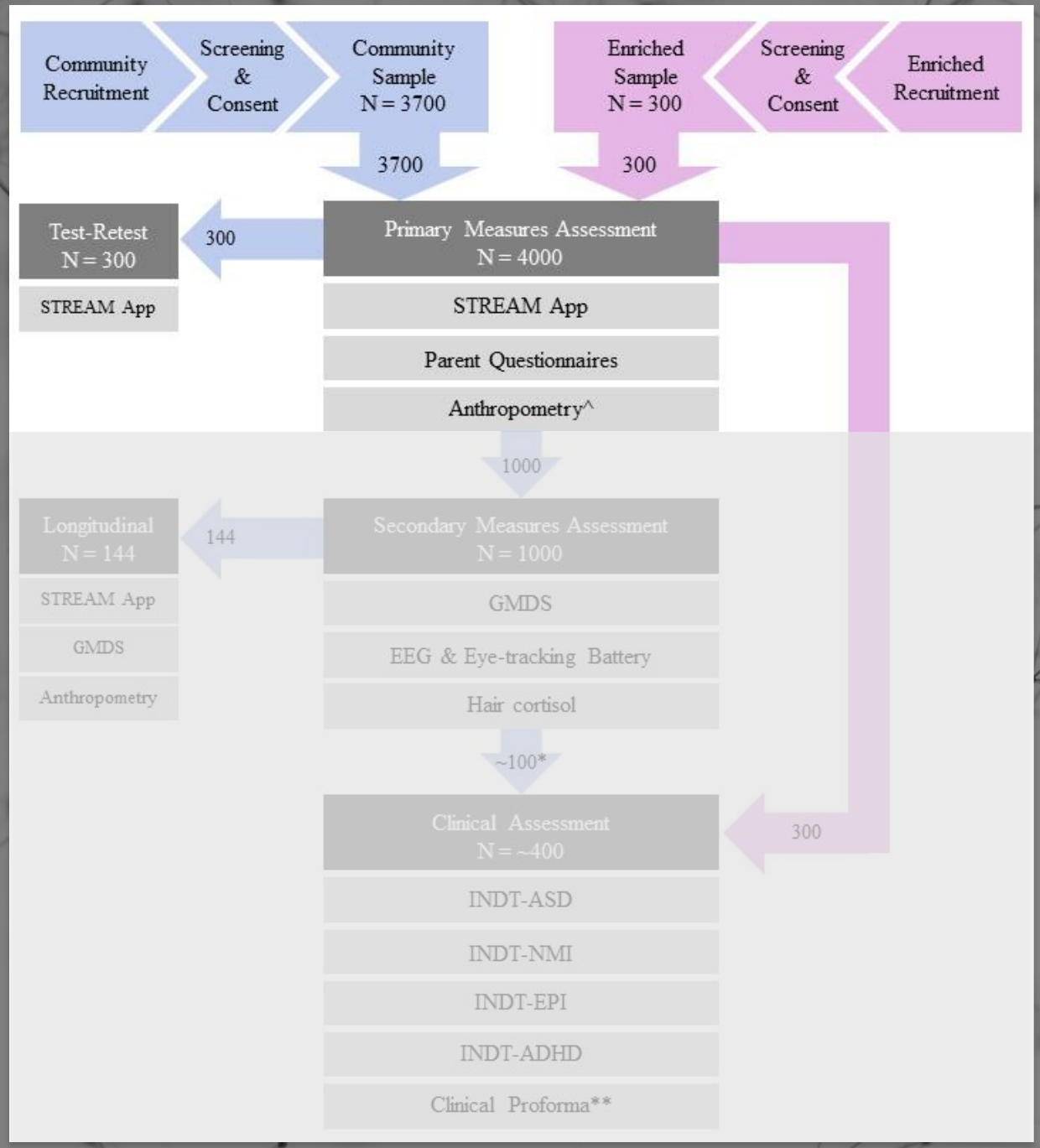
Recruitment Flow (2)

- All 4000 children do the Primary battery of Measures:
  - STREAM app
  - Anthropometry (height/length, weight, MUAC, head circumference)
  - Security and safety - Childhood psychosocial adversity scale (CPAS)
  - Responsive caregiving (0-3yrs) - Mothers object relations scale-SF (MORS-SF)
  - Responsive caregiving (3-6yrs) - Child parent relationship scale-SF (CPRS-SF)
  - Home Stimulation - Family care indicators (FCI)
  - Caregiver depression - Patient health questionnaire 9 (PHQ-9)
  - SES - DHS Wealth Index



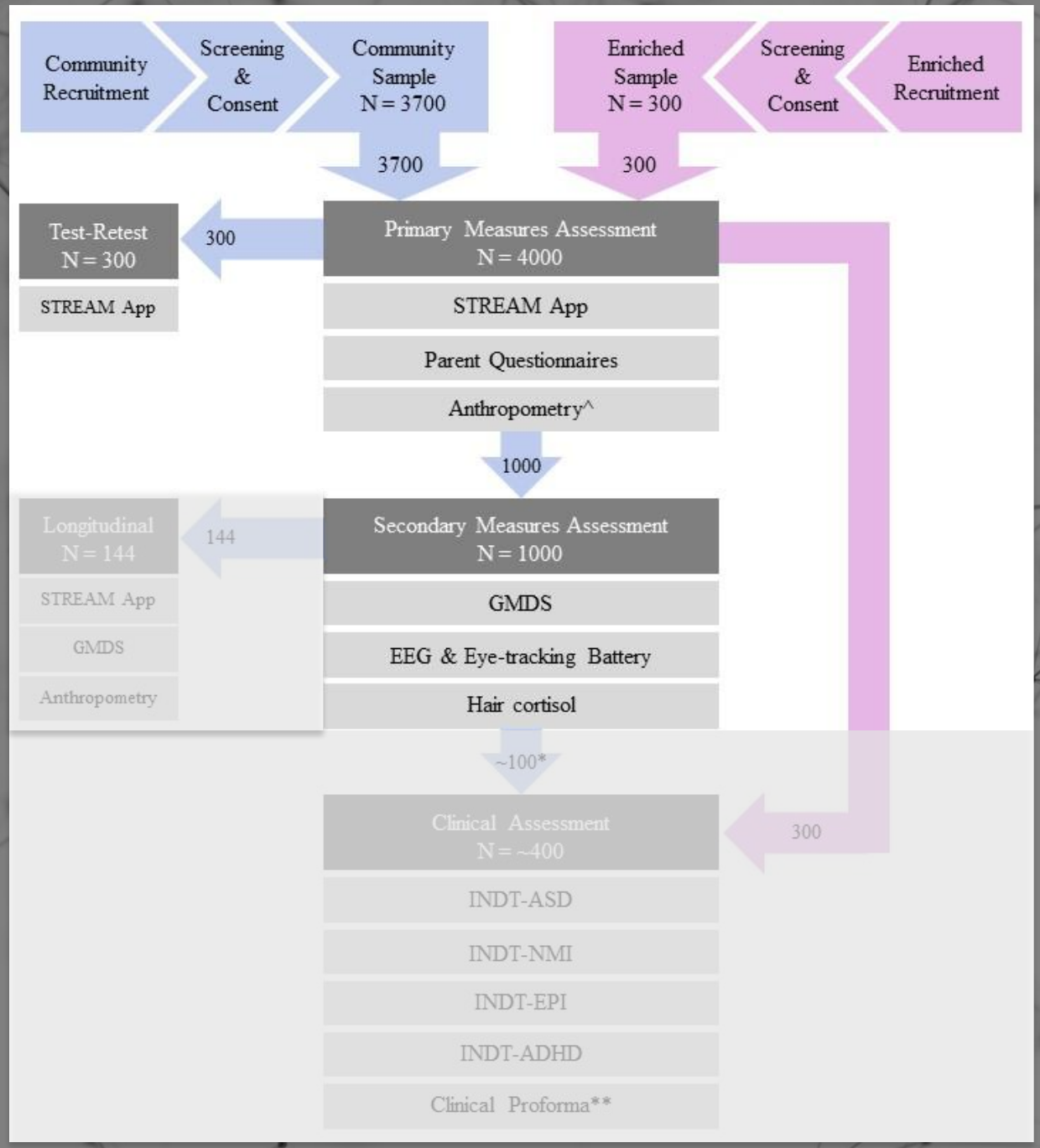
Recruitment Flow (3)

- An age-stratified subsample of (n=300) children will be randomly selected from the <u>community</u> group for a **test-retest reliability** follow-up.
- This follow-up session will be administered the STREAM app only.
- It will take place no later than 2 weeks after the first assessment.
- *The Test-retest time gap is particularly challenging in younger children.*
- *Why didn't we do Inter rater-reliability*



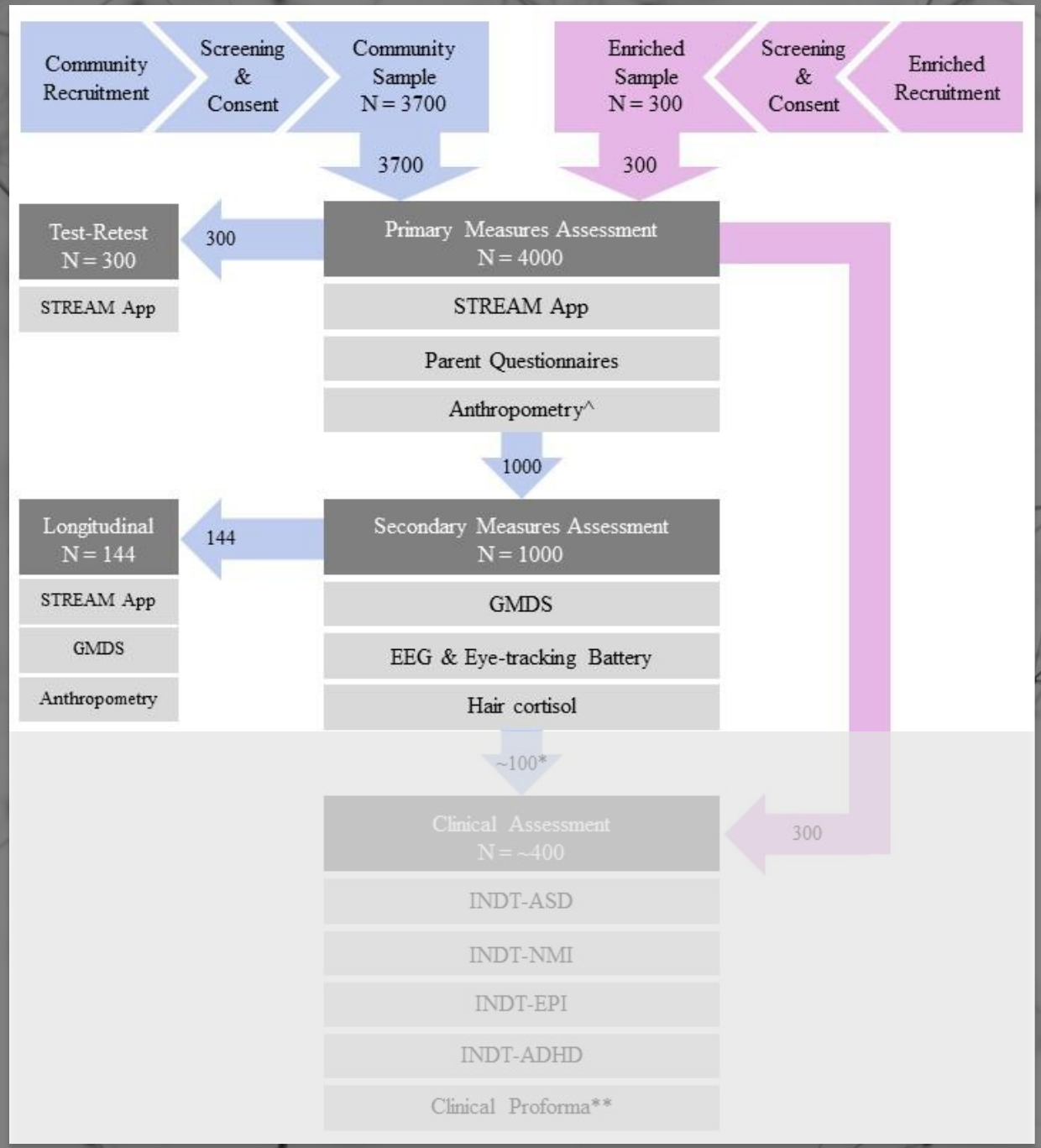
Recruitment Flow (4)

- An age-stratified subsample of (n=1000) children will be randomly selected from the community group.
- These children will :
  - Be administered the Griffiths Mental Development Scales (GMDS)(Concurrent Validity)
  - Be administered the Braintools task battery, (EEG and eye-tracking*)(Convergent Validity)
  - Provide a hair sample for cortisol analysis(Convergent Validity)
- *Why did we not give all children this secondary battery?*



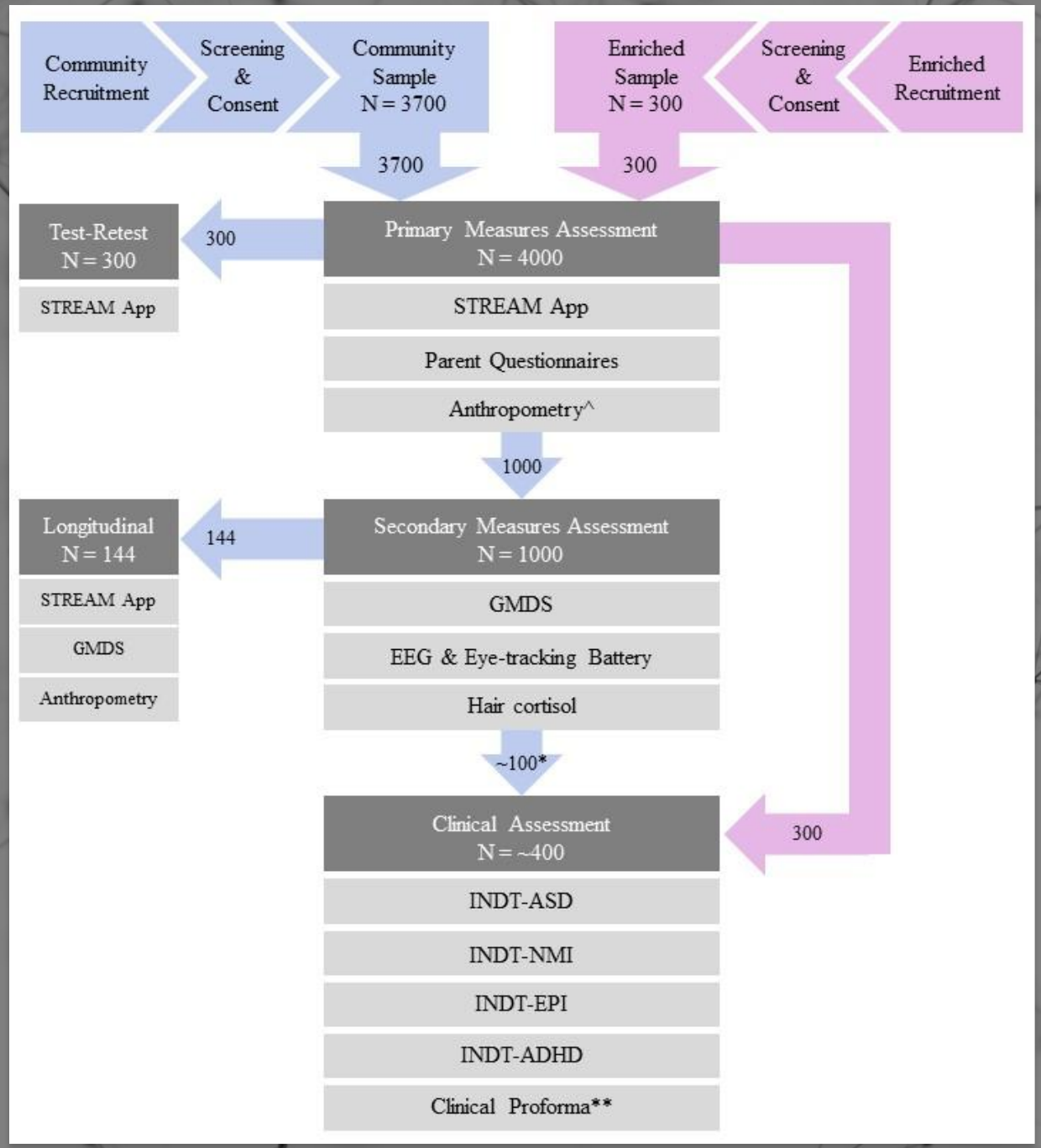
Recruitment Flow (5)

- A subsample of (n=334) children will be randomly selected from the secondary measures group for a longitudinal follow-up .
- This follow-up will take place 18mths after the first assessment and will involve:
  - the STREAM app,
  - anthropometry,
  - Griffiths Mental Development Scales
- The objective of this work will be to assess the responsiveness (i.e., sensitivity to change over time) of the STREAM.



Recruitment Flow (6)

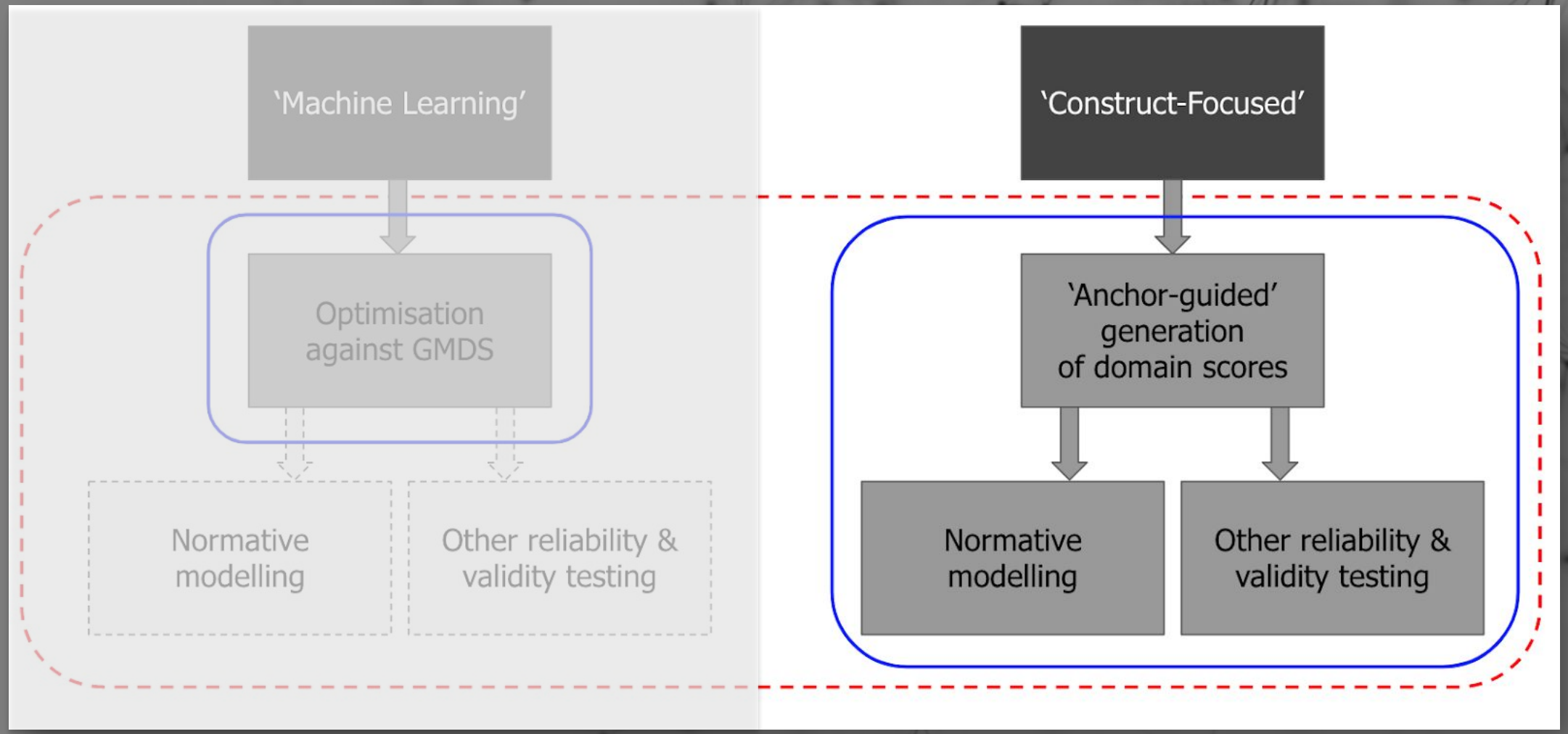
- A subsample of (n=~400) children will be recruited from the enriched group and the secondary measures group for a full clinical workup.
- This will include all (n=300) children from the enriched sample in addition to any children from the (n=1000) secondary measures group who are identified by the RBSK as NDD-risk.
- The clinical workup will comprise: a measure of intellectual disability (ID) and a clinical proforma for all children (0-6yrs); the INCLEN tools for children aged 2-6yrs (ASD, ADHD, EPI, NMI).
- The aim of this clinical assessment is to provide a better characterisation of the NDD/at-risk children within our sample, and to ensure some degree of consistency in our reporting of this information across sites.



four

Data Analysis

Analysis Streams



The “Construct Focused” Method

- The objective of the construct-focused score generation methodology is to take the output metrics for each task/section of each of these tools and create a valid, reliable and responsive, theoretically grounded set of final scores for each of the constructs we aim to measure in the STREAM platform.

- The construct-focused method will rely on collating expert judgements about which of our three target domains each task in the STREAM platform is measuring and using these judgements to inform the design of a series of confirmatory factor analyses.

PLATFORM	Suggested Primary domain being assessed	DOMAIN (Specific)	Task Name
MDAT (Grid A)	Motor	Gross motor control	Grid A (39 items in total)
MDAT (Grid B)	Motor	Fine motor control & performance	Grid B (42 items in total)
MDAT (Grid C)	Social & Communication	Language (receptive/expressive)	Grid C (40 items in total)
MDAT (Grid D)	Social & Communication	Social participation	Grid D (36 items in total)
DEEP	Attention & Cognition	Manual processing speed	Single Tap (balloon pop)
DEEP	Attention & Cognition	Manual processing speed & coordination	Alternate Tap (two balloon pop)
…	…	…	…
START	Motor	Motor planning and coordination; executive function	Bubble Popping task
START	Motor	Visuo-motor integration	Motor Following Task (Bee game)
START	Motor	Fine motor control	Colouring Task
START	Social & Communication	Social motivation (implicit)	Preferential Looking Task
START	Social & Communication	Social motivation (explicit)	Choice Task (Button game)
…	…	…	…

SEM Modelling (1)

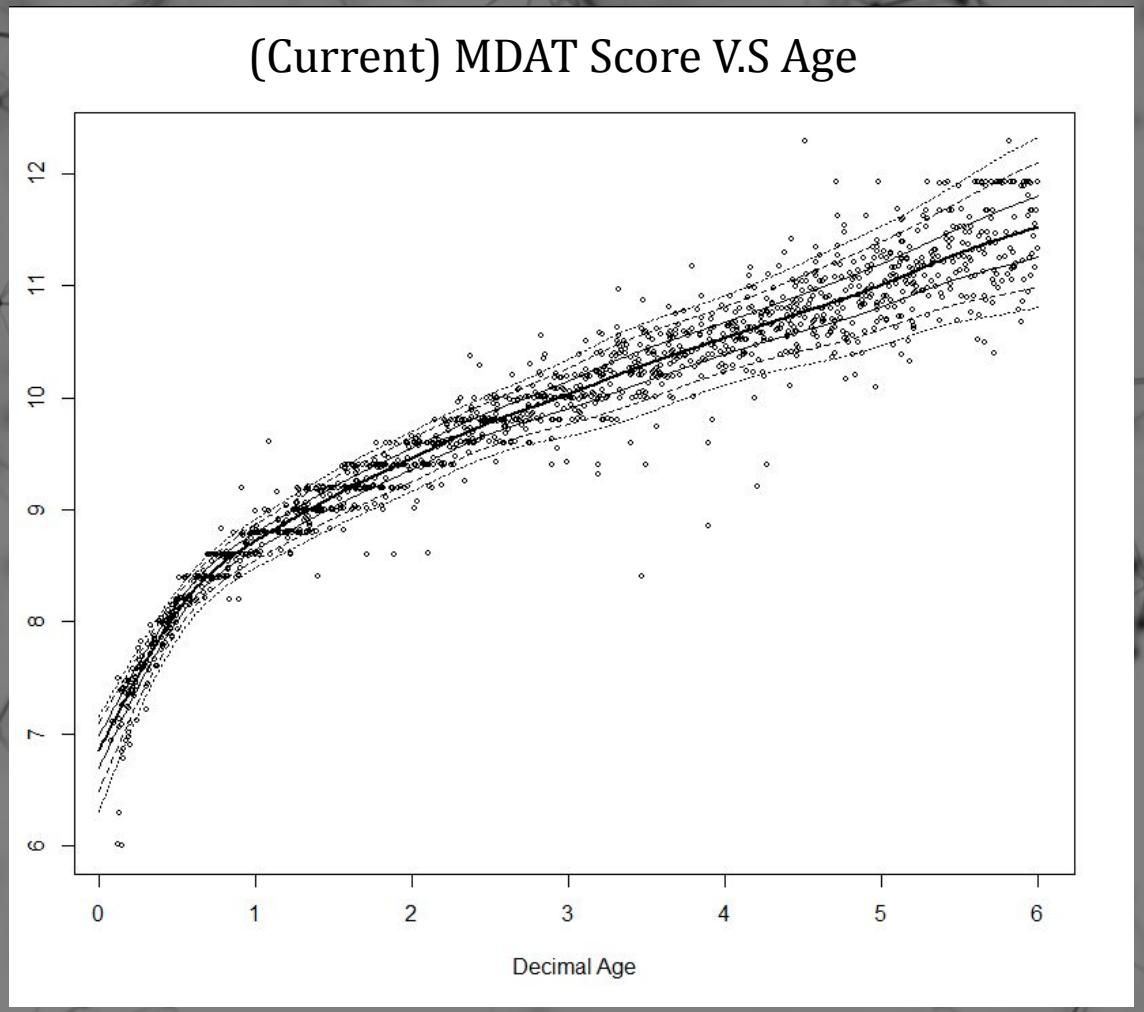
- Five factor analysis models will be created and compared on the data from n=1000 randomly selected children, with n=2700 to be a hold-out sample for final score creation after the measurement model has been selected
- **Model 1** WILL be a unidimensional model with all items loading on the same factor. (Baseline)
- **Model 2** WILL use the judgment data to ascribe one and only one task to each domain, i.e., no cross-loading between tasks.
- **Model 3** WILL allow cross loadings in cases where the judgment data indicate that a task is measuring more than one domain.
- **Model 4** WILL NOT allow cross loadings and will also include an additional second order factor representing a “total score”.
- **Model 5** WILL allow cross loadings and will also include an additional second order factor representing a “total score”.

SEM Modelling (2)

- The models will be fitted in R using Robust Maximum Likelihood estimation, known to be robust to violations of multivariate normality.
- The inferences about the final scoring models will be based on interpretation of absolute (Chisq, RMSEA, CFI) and comparative (AIC, BIC, ADJ-BIC) model fit, using standard cut values as well as qualitative judgements about final model utility.
- Decisions to be made, based on the analyses will be whether: 1) a total score is to be modelled, and 2) cross-loading of tasks be allowed.
- The final selected model will be fitted on the n=2700 children hold out sample and the fit statistics (both absolute and comparative against the unidimensional models) from this would form the final evidence for structural validity.
- The final model parameters would provide the scoring algorithm for the construct-focused methods of STREAM scoring.

Controlling for age

- We will construct our age adjusted scores for our three domains using the Generalized Additive Models for Location scale and Shape (GAMLSS)
- From these we will extract the *contextually normed* development scores (NDS), analogous to Height for Age Z-scores (HAZ) and or Weight for Age (WAZ) anthropometric scores.
- The NDSs are expressed on a standard normal distribution based on the age contingent distribution of abilities from the model.
- It represents a measure, which is not dependent on age, of a child's development on a domain, relative to their peers in the study.



Validity Analyses

Property	Variable	Method	Minimal statistic
Criterion Validity			
Concurrent validity; Reference standard	Griffiths Mental Development Scales (GMDS)	Correlation	Motor > 0.3, A&C > 0.3, S&L > 0.3, Total > 0.3
Discriminative validity	Binary diagnostic status (RBSK)	Area Under Curve (AUC)	Motor > 0.7, A&C > 0.7, S&L > 0.7, Total > 0.7
Convergent Validity			
Participation	Child and Adolescent Scale of Participation (CASP)	Correlation	0
Stunting/malnutrition	Weight for age z-score (WAZ) Height for age z-score (HAZ)	Correlation	0
Household SES	Country specific tools (see protocol)	Correlation	0
Caregiver Education	Country specific tools (see protocol)	Correlation	0
Caregiver depression	Patient Health Questionnaire 9 (PHQ9)	Correlation	0
Exposure to violence/ conflict; neglect/ abuse	Childhood Psychosocial Adversity Scale (CPAS)	Correlation	0
Responsive caregiving	Mothers object relations scale-SF (MORS-SF); Child parent relationship scale-SF (CPRS-SF)	Correlation	0
Home Environment and play / stimulation	Family Care Indicators (FCI)	Correlation	0
Neural/eye-tracking measures of development	Braintools combined EEG & eye-tracking	Correlation	0

Reliability and Responsiveness

Reliability			
Internal consistency	STREAM domain scores; item vs item total correlations	Cronbach's alpha	lower 95% CI > 0.7 for each domain
Test-retest	STREAM scores at T1 and T2 (2 weeks after T1)	ICC	lower 95% CI > 0.7 for each domain.
Other analyses			
Responsiveness / sensitivity to change	STREAM scores at T1 and T2 (18mths after T1), benchmarked against RBSK or GMDS (TBC)	Mixed effects multilevel modelling with covariates, Guyatt's Statistic using repeated measures, ANCOVA	

Post-project Plans

- Further streamlining of the app
- STREAM app to be open-source, released under a free license
- Shiny apps to support data processing and analysis
- Potential follow-ups in new contexts

The STREAM Team

Bhismadev Chakrabarti, Melissa Gladstone, Gauri Divan, Vikram Patel, Emmie Mbale, Sharat Chandran, Supriya Bhavnani, Debarati Mukherjee, Matthew Belmonte, Emily Jones, Mark Johnson, Elin Williams, Maria Crespo-Llado, Teresa Del Bianco, Georgia Lockwood-Estrin, Diksha Gajria, Vukiwe Ngoma, Chisomo Namathanga, Ananyapam De, Alok Ranjan, Ulemu Kawelama, Innocent Mpakiza, Richard Nkhata, Anindita Singh, Naina Midha, and all the amazing STREAM community health workers!

https://research.reading.ac.uk/stream/

University of Reading

Keele University

Sangath

University of Cambridge

University of Liverpool

Birkbeck University of London

University of Malawi

Indian Institute of Technology Bombay

Harvard University