What Works Hub for Global Education

BLAVATNIK SCHOOL OF GOVERNMENT

UNIVERSITY OF OXFORD

UK International Development

Partnership | Progress | Prosperity

POLICY & PRACTICE

From Measurement to Impact: Making Assessment Data Work for Student Learning

Using system assessment data for policy decisions in Indian states

Dr. Akashi Kaul • Dr. Geetika Bhandari • Shruti Bhadbhade

Central Square Foundation

From measurement to action: the policy challenge

PROBLEM

The bottleneck is not generating data. It is creating a reliable path from evidence to action.

How can FLN assessment evidence translate into specific decisions on curriculum, teacher development, governance and implementation?

EVIDENCE BASE

STATES	STUDENTS	GRADES
8	>300k	1–3

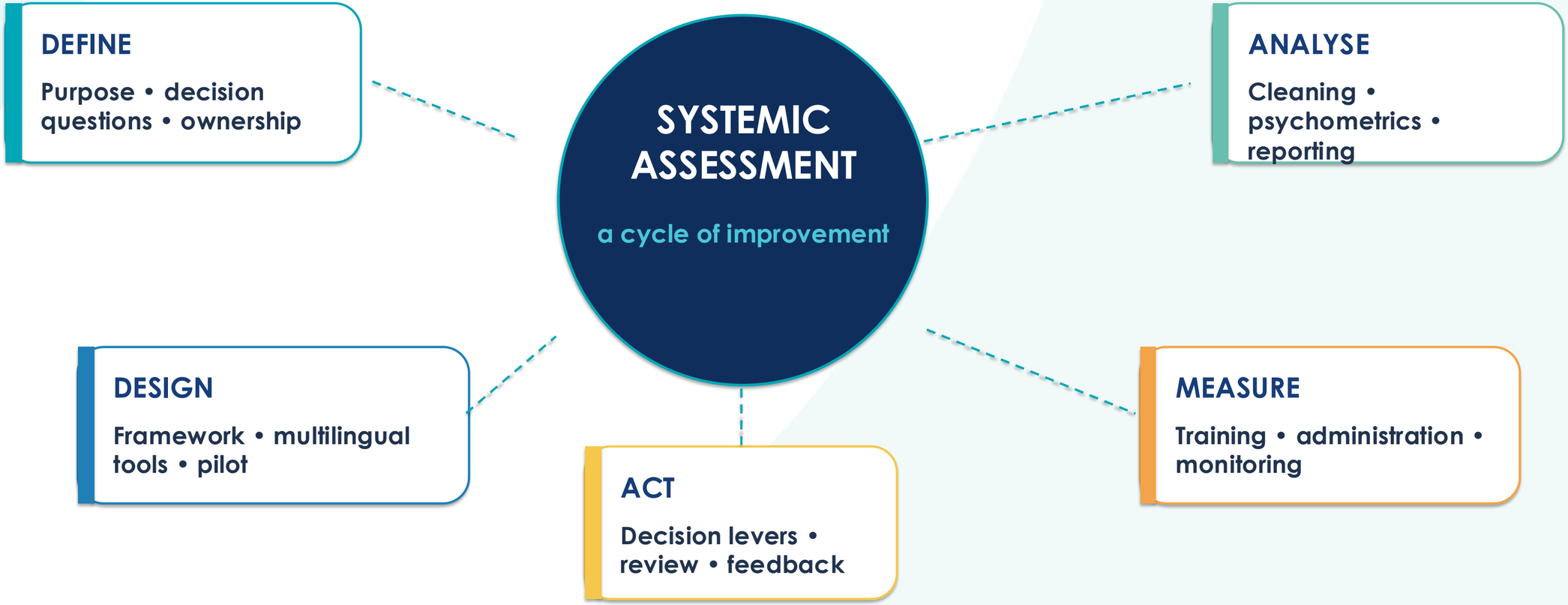
LANGUAGES	FOCUS
14+	FLN

MEASURE credible learning evidence › **INTERPRET** usable skill-level insights › **DECIDE** policy / programme lever › **ACT** implementation through existing systems

Risk: assessments become episodic or compliance-driven when findings are not linked to decisions, owners and review routines.

Treat assessment as a recurring decision cycle

OUR APPROACH



Applied across **8 states** • **14+ Indian languages** • **17+ assessment rounds since 2022**

WHY THIS MATTERS: **the same cycle addresses national alignment, assessment validity and reliability of state-level data before findings are used for policy or classroom action.**

Build credibility before data collection begins

BUILD CREDIBILITY

“Design errors cannot be fixed later.”

Multilingual ≠ translation. It is adaptation.

A common framework protects comparability; language- and context-specific development protects validity.

1

FRAMEWORK & BLUEPRINT

Clarify constructs, curriculum alignment and item proportions.

2

LANGUAGE + CONTEXT ADAPTATION

Develop tools separately where needed; use common constructs and structured review.

3

PILOT

Test administration, items and operational feasibility before scale-up.

4

RUN PSYCHOMETRICS

Use item-level evidence / IRT to revise weak items and test comparability.

Outcome: tools that are valid, contextually appropriate and comparable enough to support system decisions.

Protect fidelity from training to clean data

FIELD FIDELITY

RUNNING PSYCHOMETRICS spans pilot → live checks → post-collection validation

TRAIN & STANDARDISE

Standard package

Toolkits • administration guides
App practice • AV resources • FAQs
Practice sheets & field plans

Demonstrate readiness

Mock assessments • FI evaluation • inter-rater reliability

COLLECT & MONITOR

Standard protocols

Sampling • one-to-one administration • grading • app entry

Watch the signals

TIME | MISSING | FI VARIATION | PATTERNS

VALIDATE & CLEAN

Data forensics

Duplicates • inconsistent IDs
Outliers • implausible response patterns
Unusual completion times / scores

Act on flags

Field clarification / retraining / verification; exclude suspect records using predefined criteria where required.

Credibility is protected twice: by standardising administration and by continuously challenging the data.

Turn evidence into decisions, not just reports

DATA TO ACTION

FINDINGS
skill-level evidence

INSIGHT
simple + disaggregated

DECISION
time-bound policy lever

IMPLEMENT
state / district / block

REVIEW
feedback + next cycle

THE SAME EVIDENCE MUST WORK AT MULTIPLE LEVELS

STATE
Set policy direction and system priorities.

DISTRICT / BLOCK
Review, prioritise, troubleshoot and support.

SCHOOL / CLASSROOM
Translate findings into instructional guidance.

POLICY / PROGRAMME LEVERS

Instructional time | Teacher development | Curriculum & materials | Academic mentoring | Governance & monitoring

What determines whether assessment data gets used?

FINDINGS

CLEAR DECISION PATHWAYS

Tie findings to a specific policy lever and a time-bound response.

USABLE DATA

Quick, disaggregated, competency-level reporting makes decisions more precise.

DATA → ACTION

is mediated by the system around it

INSTITUTIONAL REVIEW ROUTINES

Monthly or quarterly forums create follow-up, tracking and course correction.

CAPACITY + TRUST

District/block ownership and confidence in data quality shape whether evidence is acted on.

Core finding: the relationship between assessment and policy is not linear; governance, leadership and system capacity determine the last mile.

Data shaping policy and programme design

What Works Hub for Global Education

Blavatnik School of Government, University of Oxford

UK International Development
Partnership | Progress | Prosperity

STATE STORIES

BIHAR

2022 baseline
SLO: 2025, 2026

SIGNAL

Weak foundational skills and need to strengthen early-grade readiness.

DECISION

3-month school readiness + Balvatika; large-scale teacher training and TLM rollout.

MECHANISM

Quarterly review systems for continuous monitoring.

UTTAR PRADESH

District rounds
2025, 2026

SIGNAL

Implementation gaps and pressure on protected instructional time.

DECISION

120-minute FLN block; time-and-motion guidelines; earlier material rollout.

MECHANISM

"Valid visit" mentoring norms + technology-enabled monitoring through VSK; district monitoring rounds covered 6 districts (Sep 2025) and 10 districts (Mar 2026).

PUNJAB

Annual SLO
2025, 2026

SIGNAL

Baseline and ongoing data exposed variation in competency and remedial needs.

DECISION

Competency grouping + Teaching at the Right Level + exit thresholds.

MECHANISM

Steering Committees and Academic Support Groups institutionalised data use; district data also informed resource allocation.

Across all three, data mattered because it was connected to a concrete decision — not simply because it was reported more often.

Data embedded in governance and implementation (1/2)

STATE STORIES

MADHYA PRADESH

Annual SLO assessment: 2025

DATA USE

SLO assessments set state and district learning targets.

SYSTEM RESPONSE

Targets were operationalised through CRC/BRC mentoring and district-led implementation; recurring reviews strengthened follow-through.

ODISHA

Annual SLO: 2025; MLE study 2026

DATA USE

The state embedded data into existing governance rather than creating a parallel review structure.

SYSTEM RESPONSE

CRCC visits, cluster meetings and FLN registers informed iterative changes to academic plans, teacher training and material use.

MADHYA PRADESH — SPECIFIC SYSTEM LEVERS

- Annual SLO assessment completed in September 2025.
- State and district learning targets were linked to structured CRC/BRC mentoring and district-led implementation.
- Recurring reviews also reinforced use of the multigrade structured-pedagogy teacher guide.

ODISHA — GOVERNANCE + MULTILINGUAL EXTENSION

- Annual SLO assessment in 2025; the 2026 round evolved into an MLE-focused study.
- Continuous monitoring through CRCC visits, cluster meetings and FLN registers fed into academic plans, training and material use.
- The 2026 MLE study extends language-wise FLN evidence across tribal languages and Odia, with findings intended to inform intervention redesign.

Both states show how assessment becomes useful when it is embedded in mentoring, review and existing academic routines.

Data embedded in governance and implementation (2/2)

STATE STORIES

HARYANA

SLO 2025
SPOT 2026

DATA USE

FLN evidence informed curriculum decisions; SPOT adds quick, real-time checks of student learning.

SYSTEM RESPONSE

Math curriculum was revamped using FLN data. SPOT training focuses on early misconceptions and timely teacher/mentor support.

TELANGANA

Annual SLO
2025 • 2026

DATA USE

Annual state-level assessment data were analysed to inform intervention and monitoring plans.

SYSTEM RESPONSE

Across three state-level assessments, evidence fed into intervention/monitoring planning; local radio was used to reach remote communities.

HARYANA — FROM SYSTEM DATA TO CLASSROOM FEEDBACK

- Annual SLO assessment completed in August 2025.
- FLN findings were used to revamp the mathematics curriculum.
- The 2026 SPOT approach focuses on quick checks that surface misconceptions in real time, so mentors/teachers can respond with targeted support.

TELANGANA — REPEATED MEASUREMENT TO INTERVENTION

- Annual state assessments were conducted in February 2025 and February 2026.
- The systemic-assessment materials document three state-level assessments feeding into intervention and monitoring plans.
- Assessment use extended beyond reporting: local radio was used as an implementation channel for remote communities.

The common thread is a feedback loop: system evidence sets direction; shorter-cycle checks and implementation channels help translate it into action.

Implementation considerations at scale

CONSIDERATIONS

These are not reasons to avoid assessment - they are design constraints that need to be planned for.

APPROVALS + LEADERSHIP CONTINUITY

Repeated permissions can extend timelines. Leadership changes may require re-alignment on purpose, roles and decision pathways.

CALENDAR ALIGNMENT

PRS years, school calendars and changing priorities need to be aligned so rounds are purposeful and do not duplicate measurement.

FINANCING + OPERATIONS

Tool development, language adaptation, training, fieldwork, technology and reporting need recurring budget commitments.

YEAR-ON-YEAR CREDIBILITY

Tool updates need a stable core framework and documented alignment so trends remain interpretable across rounds.

MULTILINGUAL IMPLEMENTATION

Language-appropriate tools and assessors are essential. Standardisation comes from common constructs and protocols, combined with language-specific adaptation and training.

Institutionalise the cycle — so it survives the next round

WAY FORWARD

The goal is to move from project-by-project execution to a standing assessment capability inside the system.

1

ALIGN

Multi-year roadmap aligned with national assessment requirements (PARAKH Rashtriya Sarvekshan / NAS) and state cycles; define the purpose of each round in advance.

2

INSTITUTIONALISE

Standing assessment cell / steering mechanism with documented handovers across leadership transitions.

3

STABILISE

Stable core framework / item bank, anchor elements and change logs so validity and year-on-year comparability are protected.

4

BUILD CAPACITY

Standing multilingual MT/FI pools; common materials, mock practice, IRR exercises and refreshers.

5

EMBED REVIEW

Recurring budgets + result review in existing state/district/block forums with named owners and timelines.

Multilingual / MLE lesson: adaptation is not translation - build consistency through common constructs, protocols and assessor preparation.

Thank you

RESOURCES

From Measurement to Impact

Making assessment data work for student learning through credible evidence, clear decision pathways and sustained review.

Central Square Foundation

Presenter: akashi@centralsquarefoundation.org

CSF systemic assessment / microsite link can be added here once the final URL is confirmed.

FLN Assessment Playbook

Practical guidance on designing and implementing robust FLN assessments at scale.

Scan to open the playbook

Open the playbook