

Research Report 2026

Inaugural Edition



Research Team

- Debjyoti Das Adhikary
- Shubhi Shukla
- Illa Sai Ravi Teja
- Abhishek Kumar
- Dr. Partha P Chakrabarti

Title : mySATHI Research Report 2026

Language : English

Author : mySATHI Foundation

Editor : K Ramachandran

Copyright © : 2026 mySATHI Foundation

No part of this book may be reproduced in a retrieval system or transmitted, in any form or by any means, electronics, mechanical, photocopying, recording, scanning and or without the written permission of the Author/Publisher.

Typeset & Published by :

GK Publications

A-45, Mohan Cooperative Industrial Area, Near Mohan Estate Metro Station,
New Delhi - 110044

Distributed by :

CL Educate Ltd (Publication Division)

Plot No. 63, Sector-27A, Near Sector - 28 Metro Station, Faridabad,
Haryana-121003

ISBN : 978-93-6914-478-5

Printer's Details : Printed in India, New Delhi.

For product information :

Visit www.gkpublications.com or email to gkp@gkpublications.com

CONTENTS

- Preface — Editor's Note

PART I

- **Chairman's Note**
Beyond the Elimination Game — Reimagining Entrance Exams in Indian Higher Education
R. Subrahmanyam, IAS
Former Secretary to Government of India, Department of Higher Education (i-iv)
- **Chairman Technical Committee's Note**
Rethinking Exams — A New Era of Assessments to Unlock Every Student's True Potential
Dr. Partha P Chakrabarti
Former Director and Professor of Computer Science and Engineering, IIT Kharagpur (v-viii)

PART II

1. **Global Viewpoints**
New Age Assessment as the Blueprint for Testing and Predicting Holistic Merit 1.1-1.2
2. **The Missing Metric — Why We Need Tests That Show How Likely Someone Is To Succeed**
Rahul Govind
Associate Professor, Business School, University of New South Wales 2.3-2.6
3. **Rethinking Assessment — Considering Business Value Psychology in Addition to Skills**
Yonatan Glaser and Raanan Haas
Yonatan Glaser is CEO and Co-founder of careernave, Israel
Raanan Haas is CEO and Chief Psychologist, youaco.com, Israel 3.7-3.10
4. **Rethinking Entrance Exams — My Vision for a Student-Centric Future**
Dr. Sadhana Prashar
Senior Director, National Testing Agency (NTA) 4.11-4.14

- 5. Lessening the Tyranny of One-Shot Exam**
Prof. Janat Shah
Former Director, IIM Udaipur 5.15-5.18
- 6. What If Talent Assessment Were Rethought for a New World?**
Bimal Rath
Founder & Chairman, Think Talent Services 6.19-6.22
- 7. A Principal's Perspective on Educational Transformation**
S. Sriram
Principal, The Mann School, New Delhi 7.23-7.26

PART III

- 8. The Report 2025** 8.27-8.28
- 9. Evaluating the 4Cs — A Psychometrically Based Framework for mySATHI Assessment** 9.29-9.32
- 10. Subject Matter Expert (SME) Empanelment for Assessment and Authoring of Assessment Questions** 10.33-10.40
- 11. Tech Infra Overview — mySATHI Application Connectivity** 11.41-11.46
- 12. AI Based Remote Proctoring v 3.0** 12.47-12.54
- 13. Nationwide Pilot of the mySATHI Assessment – Design, Execution, Findings and Implications** 13.55-13.60
- 14. mySATHI — The Computer Adaptive Testing (CAT) Ecosystem** 14.61-14.66

PART IV

- 15. A Summary of the Pilot Test Results** 15.67-15.68
- 16. Comprehensive Analysis Report of SATHI Pilot Test** 16.69-16.88

PART V

17. The Talent Observatory & Longitudinal Intelligence	17.89-17.94
18. Internationalisation of Higher Education in India — Aligning mySATHI Objectives with Study India Goals	18.95-18.98
19. Personality and Agency Assessment Suite — Completing the Holistic Readiness Lens	19.99-19.102
20. Looking into the Future — A Manifesto for Predictive and Holistic Assessment	20.103-20.106
• Acknowledgements and Contributions	107
• Glossary of Terms Used	109

PREFACE — EDITOR'S NOTE

For decades, high-stakes tests have been used to determine relative “merit” and admittance to higher education. The same method when it came to evaluating life-altering choices like certification, career advancement or job recruitments.

Today, most academics and well-meaning administrators know that these “one-time” or single-shot tests only capture a portion of a student’s aptitudes. According to such academics and administrators these exams don’t demonstrate how a pupil develops, faces a myriad challenges and changes over longer periods of time. Continuing this limited perspective in the era of Artificial Intelligence, is not only out of date, but be interpreted as not at all a responsible way to deal with young people.

Conversely and evidently, Institutions or organizations that embrace dashboards, feedback, and longitudinal insights in addition to more insightful assessments will have the advantage in future. If we give it some thought, we can see why high-stakes exams and tests are inadequate tools as they are:

- **One-dimensional:** Tests ignore teamwork, creativity, and resilience in favour of reducing skills to a numerical value.
- **Snapshot in time:** Years of training or potential may not be reflected in a single performance.
- **Opaque:** Pupils are unable to recognize their own strengths and shortcomings since they rarely receive, data-rich feedback or insightful criticism.
- **Unfairly decisive:** A future can be derailed by stress, illness, or a situation on test day. Universities that stick with “marks-only” systems run the risk of graduating students who are not equipped for the complexity of the real world.

Employers, parents, and students all deserve better. And for this reason, the mySATHI Foundation is working to create a new human capital paradigm. This research Report from mySATHI Foundation captures the essence of that work being done to create a more fair, balanced and transparent method of assessing students.

The report proposes a revolutionary change from “Assessment for Elimination” to “Assessment for Enablement.” By shifting from static, knowledge-based testing to an AI-driven, semantically adaptable framework, the mySATHI platform reimagines assessment as a corrective lens for what is now seen as a very distorted select / eliminate system.

The 4C Framework comprising Critical Thinking, Creativity, Collaboration, and Communication skills that are recognized as the fundamental skills for the workplace of the 21st century, forms the basis of this work.

In the 20th century model assessments, traditional sheets do not capture the multifaceted picture of a learner's merit. Whereas mySATHI assessment outcomes, quantify each of these '4C' attributes using computer adaptive testing and globally benchmarked psychometric methods.

This initiative's transformation into a longitudinal talent observatory platform is a key component. MySATHI seeks to reach an objective, that is, to function as a "Dynamic Passport of Potential," monitoring a student over three crucial life phases, in contrast to regular examinations that offer a single snapshot of performance:

- **University Entry:** It acts as a diagnostic tool for admissions to higher education, assisting establishments in finding "right-fit" individuals by considering their distinct innate skills rather than only competitive rankings.
- **Job Readiness:** To lower professional dissatisfaction and attrition, the platform offers talent and employability analytics as students enter the workforce, aligning particular 4C profiles to the real requirements of industry roles.
- **Mid-Career & Entrepreneurship:** An individual's capacity to develop into an entrepreneur, whose skill requirements are significantly more complicated, is measured by the engine, which acts as a barometer for higher-order skills.

This Research Report 2026 supports the conceptual strength and scalability of the mySATHI framework using data from a multi-State pilot that included over 3,200 students across cities in India. It provides a model for a system that respects the multifaceted potential of the human spirit, guaranteeing that talent is never lost in the gaps between institutions but is instead **fostered by a path illuminated by data. And in the long term play the role of a talent observatory** for pupils and professionals. It brings in an additional layer of behavioural and personality diagnostic into this integrated assessment science. The objective: help students, Universities, and organizations make better choices of talent and individuals on an ongoing basis and process. This is covered in the last part of the Report.

Our future effort would be directed towards strengthening all these aspects to make mySATHI be seen as a universally acceptable diagnostic, that is fair and transparent and supports all stakeholders for making better choices. For a country like India which is aspiring to reach a more sustainable and self supporting economy, mySATHI would seek to foray into more areas for assessments:

1. How can it become an assessor of crucial mental, cognitive traits of talent that needs to populate the millions of starts-ups and SMEs that drive employment and job creation
2. How mySATHI can add non-cognitive domains for skills in specialized areas for specific industries varying from human care to aviation, technology and drug development.

The work of mySATHI is a team effort: led from the front by Mr R Subrahmanyam, former Secretary to the Government of India in various departments including Higher Education department, and the technical aspects led by Prof. Partha P Chakrabarti, former Director of IIT-Kharagpur. While the list of people involved is long, our efforts have been strengthened by inputs from global academic leaders such as those from Educational Testing Services, psychologists Yonatan Glaser and Raanan Haas from Israel, academics Rahul Govind (from Australia), Prof Janat Shah (former director IIM Udaipur) among a host of others.

PART I

CHAIRMAN'S NOTE

Beyond the Elimination Game — Reimagining Entrance Exams in Indian Higher Education

R. Subrahmanyam, IAS

Former Secretary to Government of India, Department of Higher Education.

For decades, the narrative of Indian education has been dominated by a single, ruthless metric: the entrance score. It is a system where a fraction of a decimal point decides destinies, segregating the “winners” from the “failures” before they have even entered a classroom. Years of policymaking and observation show it is evident that we have built an ecosystem not of education, but of elimination. The “Kota Factory” phenomenon—where teenagers are sequestered in coaching hubs, stripped of leisure, social growth, and often their mental health, is not an aberration; it is the logical conclusion of a system that prioritizes rote recall over human potential.

The crisis is visible in two distinct but related tragedies. The first is the human toll: the rising curve of student burnout, anxiety, and suicide. The second is the systemic inefficiency: industries facing a talent crunch despite millions of graduates, simply because our testing mechanisms measure the ability to crack an exam, not the ability to thrive in a workspace. The National Education Policy (NEP) 2020 sounded the alarm, calling for a shift from high-stakes rote assessment to holistic evaluation. It is in this context that we introduce the SATHI (Scholastic Aptitude for Higher Ed Institutions) initiative—not merely as a new test, but as a corrective lens for a distorted system.

The Crisis of the “Dummy” Education

The current admission regimes, particularly for engineering and medicine, have inadvertently hollowed out the school system. The rise of “dummy schools, where institutions enroll students only for board eligibility while spending their days in coaching centers. Is severing the link between schooling and learning. The classroom, once a space for socialisation, debate, and holistic growth, has been rendered obsolete by the coaching center, which focuses strictly on hacking the exam pattern.

This model fundamentally fails the test of “Swa-dharma”—the ancient Indian concept that education should help an individual discover their inherent nature and calling. Instead of guiding students toward paths where their unique aptitudes lie, the system forces a homogenization of talent. A student with profound creative intelligence or high emotional quotient (EQ) is judged by the same rigid yardstick as one with high computational speed. The result: a workforce comprising of “successful” graduates who are often mismatched with their careers, leading to professional frustration and lower productivity.

The Paradigm Shift: From Filtering to Diagnostics

The core philosophy of SATHI is to move from “Assessment for Elimination” to “Assessment for Enablement”. If traditional exams are a gate that slams shut, SATHI is envisioned as a mirror. It is premised on the belief that every student has a unique area of strength, that traditional marks sheets fail to capture.

To operationalize this, we are moving away from static, knowledge-based testing toward an AI-driven, adaptive framework. Unlike a standard entrance exam where every student faces the same set of questions regardless of ability, an adaptive test adjusts in real-time. If a candidate answers correctly, the difficulty increases; if they falter, it adjusts to find their actual floor and ceiling. Thus, actually measuring the true extent of a student’s capability rather than just their ability to memorize standard answers.

This technological shift allows for a much broader evaluation matrix. While the Joint Entrance Examination (JEE) or the National Eligibility cum Entrance Test (NEET) or Common University Entrance Test (CUET) or the Common Admission Test (CAT) might test Physics, Chemistry, and Math, SATHI integrates the OCEAN model of personality and the “4Cs of 21st-century skills”.

Feature	Traditional Entrance Exams	SATHI Assessment Framework
Primary Goal	Elimination (Filtering out candidates)	Enablement (Diagnosing potential)
Methodology	Standardized, linear pen-and-paper/CBT	AI-driven, Computer Adaptive Testing
Scope	Academic Knowledge (PCM/PCB)	Holistic: Aptitude + Personality + Skills
Outcome	A single Rank/Score	A comprehensive “Swa-dharma” Report
Student Impact	High Stress, “Do or Die”	Self-awareness, Career Guidance Choice
Institutional Use	Filling Seats	Matching student to the right program

Decoding the Mechanism: What SATHI Measures

Our orientation with pilot institutions reveals that SATHI is not replacing subject knowledge but contextualizing it within a wider human framework. The assessment is concise—approximately 120 minutes—yet it gathers data points that usually take years of observation to compile.

1. The 4Cs of 21st Century Skills

The industry consistently complains that graduates lack “soft skills.” We quantify these explicitly through SATHI:

- Critical Thinking: Can the student analyze arguments and identify biases?
- Creativity: Can they generate novel solutions to open-ended problems?
- Collaboration: How do they function in team dynamics?
- Communication: Can they articulate complex ideas clearly and concisely?

By measuring these, SATHI provides a “Employability Quotient” before the student even begins their degree.

The Value Proposition for Institutions, Organisations

For the pilot institutions adopting this model, the benefits extend beyond just “better admissions.” The current admission process is a blind date; colleges know students only by their marks. SATHI transforms this into an informed partnership.

Strategic Cohort Analysis

Institutions and organisations gain access to an aggregated dashboard that maps the talent landscape of their incoming batch or candidate. A college might discover that their incoming Computer Science batch is high on analytical logic but low on resilience (Neuroticism). This insight allows the faculty to design specific interventions—such as bridge courses or mentorship programs—right in the first semester, significantly reducing dropout rates.

Placement Alignment

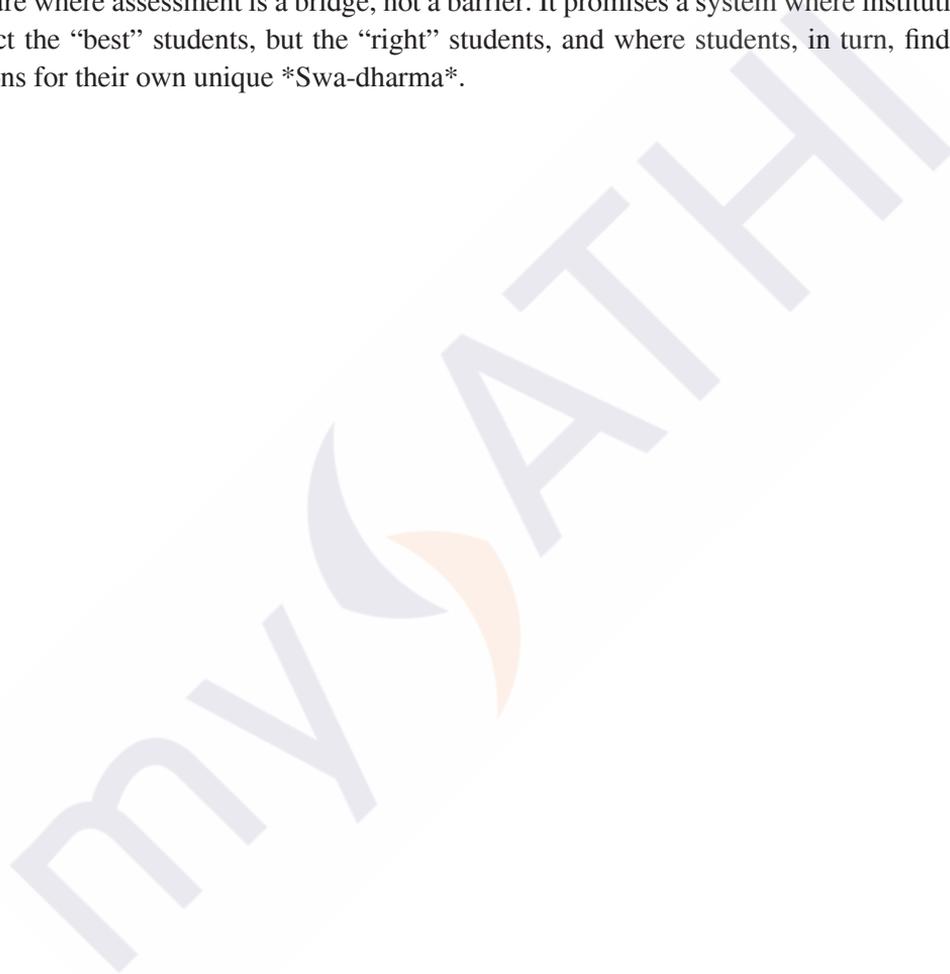
One of the most pragmatic applications of SATHI is in placement preparation. By having a diagnostic report of a student's aptitudes in the first year, career cells can guide students toward internships and roles that fit their profile. A student with high ‘Extroversion’ and ‘Communication’ skills can be steered toward sales or management consulting, while a student with high ‘Introversion’ and deep ‘Analytical focus’ might be guided toward backend development or R&D. This alignment increases the “stickiness” of placements—graduates stay in jobs longer because they are naturally suited for them.

The Path Forward: A Coalition of the Willing

The rollout of SATHI through our pilot institutions represents a “coalition of the willing”. We acknowledge that systemic change cannot be imposed from the top down overnight; it requires proof of concept. These pilot institutions are not just testing software; they are testing a new philosophy of education.

The ultimate goal is to dismantle the “fear of failure” that permeates Indian education. By providing a low-stakes, high-value diagnostic tool, we can shift the student mindset from “How do I beat the competition?” to “How do I understand and better myself?”.

The transition from a rank-based hierarchy to a skill-based taxonomy is difficult. It requires unlearning decades of reliance on convenient, singular numbers. However, the cost of the status quo—measured in student lives and economic stagnation—is far higher. SATHI offers a blueprint for a future where assessment is a bridge, not a barrier. It promises a system where institutions don't just select the “best” students, but the “right” students, and where students, in turn, find the right institutions for their own unique *Swa-dharma*.



CHAIRMAN TECHNICAL COMMITTEE'S NOTE

Rethinking Exams — A New Era of Assessments to Unlock Every Student's True Potential

Dr. Partha P Chakrabarti

*Former Director and Professor of
Computer Science and Engineering, IIT Kharagpur*

Every year in India, millions of students see their futures shaped by exam results. In the tension of the exam hall, under a proctor's sharp gaze and the stark glow of a computer screen, years of work collapse into a percentile, a rank, a score. For generations, this system has reliably sorted students for college, professional courses, and jobs. It is standardized and clear. Yet, it fundamentally rewards conformity rather than curiosity and values recall above innovation. It is designed to find the one right answer, while the world students now face is defined by ambiguity—by questions that have many possible answers, or even none at all.

As AI and automation revolutionize the workplace, the skills that ensure lifelong success—critical thinking, problem solving, communication, and collaboration—are precisely the ones that routine tests often ignore. In this transformed era, India must rethink not just what students learn, but how their abilities are assessed.

The Exam System: Strengths and Flaws

Indian exams have served a vital social purpose, making university admissions more transparent and merit-based. They have created generations of professionals able to tackle administrative hurdles and technical challenges with confidence. However, their drawbacks are clear: creativity is often stifled, divergent thinking is underappreciated, and students learn to predict what examiners want, not what complex problems need. The intense focus on cramming and last-minute preparation leaves little room to assess how learners tackle open-ended or unfamiliar challenges.

A Paradigm Shift: Semantically Adaptive Assessment

India now stands at the threshold of a genuine assessment revolution, driven by AI and computer-based testing. The semantically adaptive test isn't just a minor update—it is a wholesale rethinking of what exams can and should be. The shift begins with Computer Adaptive Testing (CAT), which sequences questions for each student in real time, based on their prior answers. If a student answers correctly, the next question is harder; if incorrectly, easier. This efficiently pinpoints a student's proficiency level with fewer questions.

The real leap comes with semantic adaptivity—where not only question difficulty, but context and cognitive demands, change dynamically. These advanced CATs don't just serve up incrementally tougher problems in the same area; they challenge students to apply central ideas across very different domains. For example, a student who understands fluid dynamics might be asked to use those principles to analyze traffic flow or network information—a true test of adaptive conceptual thinking. Like an athlete mastering new disciplines rather than just running faster sprints, students are challenged to transfer skills and knowledge flexibly. AI and the State Space Model: Mapping Knowledge

This leap relies on the “state-space paradigm” in AI, which sees problem-solving as the exploration of many possible states. In assessment, students journey through a series of scenarios, each probing their understanding and strategy. Each question is a point in a multidimensional constellation of knowledge and skill, and the test tracks how students move through this map. This approach measures both what students know and how flexibly and deeply they connect ideas, distinguishing those who memorize solutions from those who can invent new ones.

Beyond Cramming: The New Educational Mandate

Such change in assessment is about more than technology—it's a reflection of how education must evolve in the information age. Now, facts are easily accessible; what matters is filtering, synthesizing, and making sense of them. The main challenges for students are not ignorance but information overload and misinformation. Education must empower learners to judge credibility, relevance, and application, cultivating discernment and depth instead of rote memorization.

The 4Cs—Critical Thinking, Creativity, Collaboration, Communication

India's new exam paradigm targets the most vital skills for success today—the “4Cs”:

- **Critical Thinking:** Instead of seeking a single correct answer, future assessments will pose complex scenarios—like medical innovations with societal consequences—and track how a candidate reasons. AI evaluates not just the end result, but the assumptions considered, counterarguments weighed, and alternate paths explored, revealing the student's resilience and maturity of thought.

- **Creativity:** Creativity is not only artistry—it's making useful, original connections across fields. Students may face tasks requiring them to explain trends merging history, economics, and demographics, or to create solutions for neglected challenges. AI judges not just what the student produces, but also the originality, risk-taking, and thinking process involved.
- **Collaboration:** While conventional tests reward solo effort, innovation in the real world is collaborative. New CATs will simulate negotiations and team-based problem solving, with AI agents representing diverse stakeholders. The test records how students listen, persuade, share and build consensus—key components of leadership and teamwork.
- **Communication:** Success depends on more than just good ideas; communication must be nuanced, adaptable, and clear. Assessments use natural language processing to evaluate how students tailor messages for different audiences, articulate complex ideas, and persuade across varied contexts—from drafting executive memos, critical analysis to public announcements.

The Diagnostic Compass: Personalised Insights

A major advance is the move from bland rank or score to multidimensional diagnostic feedback. Students, teachers, and universities receive detailed maps of skills—much like a comprehensive medical profile—pinpointing areas of excellence and those that need improvement. This helps universities conduct more holistic admissions, enables teachers to target remediation more precisely, and allows parents to use actionable data for long-term development.

Implications for Education and Policy

Semantically adaptive assessments create a philosophical shift towards student-centric education, recognizing each learner's strengths and habits of thought. In an age of rapid change, these systems prepare young people not merely to pass exams, but to meet and create the challenges and opportunities ahead.

These changes will spread into curriculum design, teacher training, and education policy across India, emphasizing meaningful feedback, continual improvement, and lifelong learning. Institutions are incentivized to foster deep learning rather than shallow scanning of content. Ultimately, success will depend much more on adaptability and creativity, rather than just the ability to recall facts in a stress-filled exam hall.

Building Futures, Not Just Ranking Minds

India's adoption of semantically adaptive, AI-powered assessments is more than a technical fix—it is a reimagining of the purpose of education. It affirms that every student is a unique constellation of abilities and growth paths, and that exams should be tools to map, not flatten, this diversity. By emphasizing holistic skills and personalized learning journeys over rote recall, India is crafting an education system that genuinely equips its youth for the uncertainties and possibilities of the future—a true “Viksit Bharat.”

This reform offers hope that every student's potential will be discovered and developed: a new compact for Indian education in the age of intelligence, creativity, and imagination.

PART II

CHAPTER 1

GLOBAL VIEWPOINTS

New Age Assessment as the Blueprint for Testing and Predicting Holistic Merit

One of the goals of mySATHI Research Team is, on a continuous basis, to learn more about the architectural of the way Indian higher education tests students, in the backdrop of how to strengthen SATHI (Student Assessment and Testing for Holistic Inclusion) program.

We want to make a strong case for not just relying on scores from rote learning. Domain knowledge, like Physics or History, is necessary, but it is not enough to predict a learner's long-term success in their career or their ability to adapt to new academic situations in the modern world, especially now that machines do most of the work that people used to do before AI. The research team's effort was bolstered by contributions from leading experts in assessment sciences, highly accountable scholars, and individuals with practical experience yielding results. This part puts these kinds of information into two formats:

- One is a general section that shows how a truly reflective new age assessment should be used to show the student's overall, multi-dimensional merit. And this is not just for now; it is also a sign of how the same student's career will grow in the future.
- A part that puts together the individual reflections of the experts who graciously agreed to share their thoughts with nuance, responsibility, and deep thought.

The researchers of mySATHI continue to look at and work on making a strong framework for measuring a student's "invisible" qualities, like their intelligence, how well they communicate, and how mentally ready they are for the challenges of higher education. What are the structural elements and operational principles of this new assessment paradigm if this is the strategic goal?

A clear long-term goal is to ensure that the mySATHI platform transforms student assessment into a lifelong talent observatory, tracking skill evolution from university entrance through careers to entrepreneurial leadership. By measuring core competencies—critical thinking, creativity, collaboration, and communication—it predicts and shapes trajectories, enabling precise role matching and growth forecasting. Because a college entrance test need not be a valuable or accurate predictor engine for future success of an individual, nor would that person be able to nurture those skills that would be necessary for professional blooming and growth.

Undergraduate education marks the shift from rote learning to self-directed mastery, yet traditional exams overlook skills vital for real-world success. mySATHI fills this void with globally benchmarked evaluations, AI-proctored delivery, and actionable feedback. It begins at entrance, guiding admissions to skill-ready programs, then extends via continuous tools to align jobs, leadership paths, and ventures, turning assessment into a lifelong companion.

Another key differentiator of mySATHI is its detailed feedback dashboard that provides time accurate analysis and deep insightful analytics about the performance of a candidate in various parts of the question paper and how he or she has fared in each of the 4Cs, as also how the candidate has coped with the varied difficulty levels, besides the capability to think through each difficulty level.

Reasoning for the Argument

Current systems conflate knowledge recall with aptitude, favoring coaching over innate potential and yielding high dropout rates. mySATHI separates these via pillars like verbal ability, quantitative reasoning, abstract logic, and psychometrics, predicting trainability and resilience.

- Entrance: Benchmarks readiness for higher education.
- Careers: Matches skills to roles via gap analysis.
- Leadership: Identifies traits for business innovation. Complementary tools like TalentGuard for real-time tracking and O*NET for exploration ensure ongoing measurement, creating a seamless pipeline from student to entrepreneur.

Prognosis for the Near Future

Within 2-3 years, mySATHI-integrated platforms will strive to standardize skill transcripts for admissions and hiring in India, slashing mismatches by 30-40%. Edtech adoption will surge, with universities and firms using dashboards for scholarships, lateral entries, and upskilling, aligning with national goals for a Viksit Bharat workforce.

mySATHI as Talent Observatory

Long-term, mySATHI evolves into a predictive hub analyzing longitudinal data to forecast roles: analytical minds for data leadership, resilient creators for startups. Periodic reassessments track growth, recommend boosters, and signal shifts—like pivoting to entrepreneurship—empowering individuals and institutions to nurture India's diverse talent for decades ahead.

CHAPTER 2

The Missing Metric — Why We Need Tests That Show How Likely Someone Is To Succeed

Rahul Govind

Associate Professor, Business School, University of New South Wales

In university admissions, the manner in which students are assessed as an incoming cohort, is just as contested as the content and dissemination of education itself. Today, universities are grappling with how best to select their incoming cohorts. Across countries and institutions, this debate has produced markedly different selection regimes, ranging from standardized aptitude tests, to holistic evaluations emphasizing essays and extracurriculars, to curriculum-based credentials such as the International Baccalaureate, and rank-based indices such as Australia's ATAR. This fragmentation across and within universities towards an admissions strategy, points to a deeper and unresolved question: **are existing assessment systems actually measuring the attributes that matter for long-term success?**

As an Australian academic who has examined initiatives such as SATHI, it is clear that prevailing assessment frameworks are inadequate along a crucial dimension. The current systems are limited, not in their ability to rank students at a specific point in time, but their failure to **estimate the probability of sustained academic/professional success**. They are very good at sorting students based on short-term performance and factual recall. But they are far less effective at predicting outcomes even in the short run. **In other words, we are optimizing for selection convenience rather than predictive validity.**

In theory, university admissions metrics operate at the intersection of human capital formation and ability signaling. However, much of the existing admissions infrastructure is better aligned with signaling current efficiency than with human capital prediction. Aggregate test scores compress multidimensional abilities into a single nominal and/or ranks. While it helps administrative sorting but discards information relevant to field-specific learning trajectories and career outcomes.

When we have a large diverse pool of applicants, a single number/grade is a horrific measure! Thus, in the admissions metrics, we should discuss the use of sub-component scores than

relying exclusively on **aggregate scores**. Breaking scores into interpretable components will enable universities to evaluate how particular competencies translate into academic persistence, placement outcomes, and professional growth, **rather than merely exam performance**. More importantly, the variations in the metrics will allow them to choose the specific profile that fits the course requirements. These variations become more important for universities when one looks at cross-national standardization of metrics that has been long known as an issue in student selection.

From this perspective, subcomponent-based assessment is not an incremental refinement, but a theoretically grounded correction to a misalignment between what tests measure and what institutions ultimately care about.

In the Australian system the ATAR score functions as a universal ranking mechanism, allowing students to apply across disciplines even without subject-specific prerequisites. The process is streamlined, and additional screening is often minimal. However, even in this context, the core challenge remains unresolved: **whether the dominant metric is sufficiently robust, equitable, and predictive of downstream success**.

In India, the stakes are substantially higher. The sheer volume of applicants creates strong pressure for coarse screening tools, increasing the risk that high-potential candidates are excluded due to underperformance on standardized tests that fail to capture their full potential. **Students from resource-constrained institutions, smaller cities, or disrupted personal circumstances are disproportionately penalized, despite possessing the capacity to succeed once given access and support**. This is both an equity and an efficiency problem. For example, students transitioning from regional or domestic systems to international academic environments often face adjustment costs that are orthogonal to underlying ability, yet these costs are implicitly treated as signals of lower merit.

This is precisely where test design becomes consequential. Well-designed computer-based assessments can outperform traditional structured and non-malleable pen and paper tests in both fairness as well as informational richness. **Adaptive testing frameworks can dynamically adjust difficulty, reduce noise from test-taking conditions, and generate fine-grained diagnostics on cognitive and domain-specific skills**. Such systems are not only scalable but also capable of producing the subcomponent data required for meaningful prediction.

The logic underlying improved assessment design can be broken into three linked stages.

- First, assessment structure determines the information content of admissions signals. Items that test the performance into interpretable subcomponents, generate richer data on applicant strengths and weaknesses than aggregate scores.
- Second, richer information enables better matching among students, programs, and institutional support structures. When universities observe not only overall rank but also domain-specific aptitude and learning potential, they can allocate students more efficiently across disciplines, tailor interventions, and reduce mismatch-driven attrition.

- Third, improved matching and targeted support can lead to superior long-term outcomes, including academic persistence, skill acquisition, and labor market performance. Importantly, this mechanism does not rely on changing student ability itself, but on reducing informational friction at the point of selection. So, gains in equity and efficiency can occur simultaneously, particularly for applicants whose potential is poorly captured by conventional standardized tests.

This framework implies that the success of any new testing system should be evaluated not by short-run correlations with existing exams, but by its ability to predict downstream outcomes.

SATHI can become transformative, not as merely another entrance examination, but its ability to demonstrate clear and incremental value. Make top tier Indian universities adopt it., as this empirical **validation is essential**. SATHI outcomes should be benchmarked against established international standards such as the GMAT, while more critically, they must be linked to longitudinal indicators including placement quality, career progression, and sustained academic performance – something the other assessment don't do.

Ultimately, universities require metrics that enable comparability, particularly for international mobility. However, educators require metrics that do more than just rank. They require metrics that predict. **A system that maps student capabilities to future pathways allows institutions not merely to allocate seats, but to allocate opportunity.** By explicitly accounting for heterogeneity in strengths across domains, such a framework would move higher education closer to recognizing and cultivating diverse forms of talent, rather than filtering them out prematurely.

CHAPTER 3

Rethinking Assessment — Considering Business Value Psychology in Addition to Skills

Yonatan Glaser and Raanan Haas

*Yonatan Glaser is CEO and Co-founder of careernave, Israel
Raanan Haas is CEO and Chief Psychologist, youaco.com, Israel*

As assessment specialists, HR leaders, and systems thinkers deeply embedded in the worlds of education and policy, we believe the current discussion regarding student assessments needs to extend far beyond the traditional, limited realm of college entrance tests. We view the SATHI initiative not merely as another exam to be administered, but as a potential comprehensive, multi-stakeholder platform capable of addressing some of the most critical issues currently facing employment and higher education in India.

Our shared goal is to move the industry away from conventional skill-based tests and toward a future-proof model. We envision a system that acts as a gateway for students, a useful signal for businesses, and an astute strategic tool for colleges. Below, we outline our vision for making SATHI a genuinely distinctive and competitive endeavor.

Beyond Skills: The Business Value Psychology Approach

We contest the basic assumption of the majority of standardized exams, including the SAT. We believe these assessments serve largely as “gatekeepers” and are based on a “skills-based psychology” methodology that is rapidly falling short in a world that is changing fast. Particularly with the disruptive power of Artificial Intelligence, today’s skills can become outdated tomorrow.

Hence, we propose adding a “second layer” to skills evaluation - “business value-based psychology” whose novel paradigm seeks to assess more profound, durable human qualities essential for sustained success in the the classroom and at workplace. We dissect this idea into several essential parts.

First, we emphasize motivation. We strongly believe that without the drive to use a skill, simply possessing it is useless. Therefore, we propose that assessments to determine a student’s intrinsic motivators and drivers should be included in SATHI. This may assist young people in finding a meaningful path by identifying their specific triggers.

Second, we distinguish between being a good employee and being a good student. The expectations, abilities, and mindsets needed for these two positions differ significantly. We contend that it is the duty of a university to prepare students for both, and that SATHI can serve as a diagnostic instrument to determine a candidate's preparedness for both situations.

Third, we focus on alignment with employer expectations. The demands of the sector are directly incorporated into our suggested approach. Employers are looking for graduates who can instantly add value to the company. SATHI would benefit students, colleges, and the companies that would eventually hire these graduates by evaluating qualities that meet actual employer standards.

Innovative Assessment: The Projective Technique

Expanding upon this novel psychological model, we propose a fundamental change in the design of the test items themselves. - a projective technique evaluates a person's comprehension of an idea without ever putting the explicit question out there—as opposed to asking a clear query like “what is 1+1?”.

This approach decreases maintenance and cheating, as a test based on this methodology is much harder to prepare for in a traditional way. This lessens the incentive to cheat and cuts down on the long-term expenses of test revision. Furthermore, it encourages equity and inclusion. In a large and diverse nation like India, the projective techniques can uncover untapped potential and guarantee that everyone has an equal opportunity to showcase their skills. Our techniques are intended to uncover the underlying motivations and strengths in these people.

SATHI as an Institutional Intelligence Platform

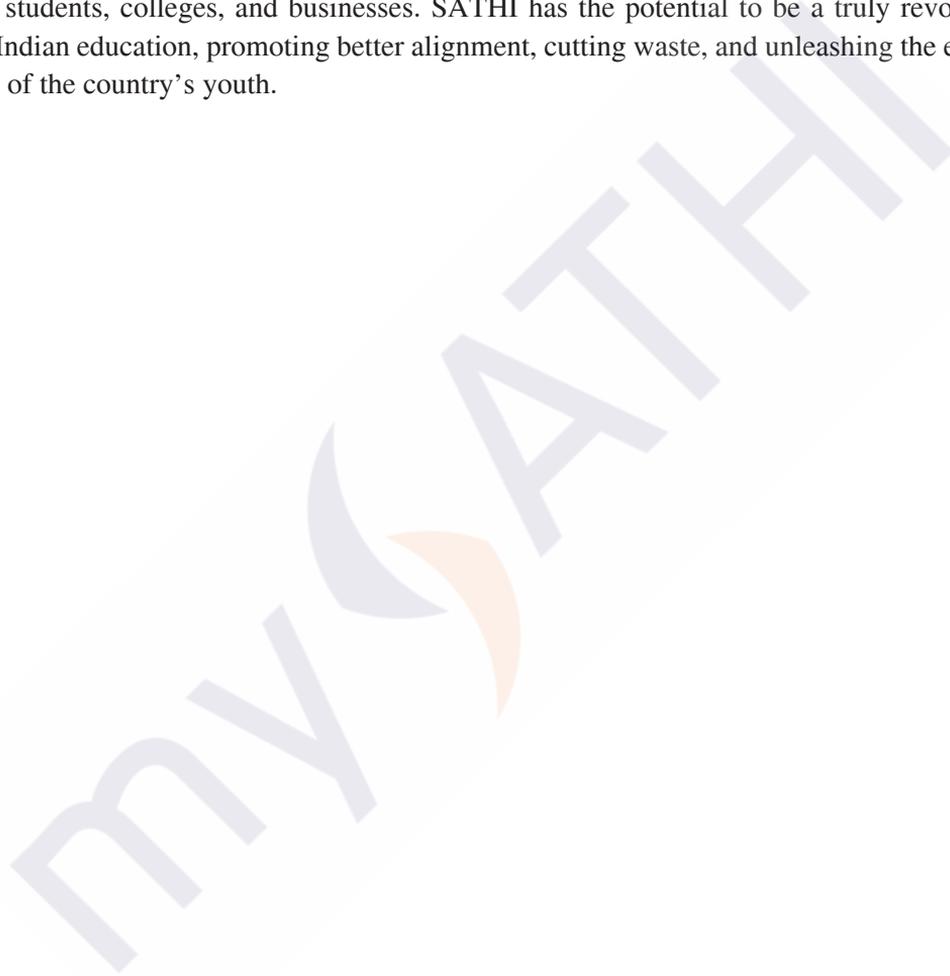
We provide a more comprehensive, systems-level viewpoint to supplement this psychological paradigm.

We feel that SATHI should be developed to solve “Institutional Blindness.” Many Tier II and III universities, might not have the resources necessary to comprehend the requirements of their own programs or the actual career routes they lead to.

Our most bold recommendation is that SATHI should develop into a potent Institutional Intelligence Platform rather than just a test administered to students. This would make SATHI a “must-have” rather than just a “nice-to-have,” turning it into a vital B2B tool for colleges. This platform would provide analytics for program-job fit, giving educational institutions data-driven insights into how their different curricula relate to particular job outcomes. It would also align with economic models, assisting organizations in managing operational and financial concerns—such as lowering dropout rates and increasing on-time completion—by better matching students to courses using diagnostic insights.

Finally, looking at the specific Indian situation, we contend that a focus on motivation and appropriate career alignment is a national necessity, directly connecting this to the startling statistic of 30 million unemployed university graduates in India. We strongly advise incorporating the employer's viewpoint directly into SATHI's governance structure—perhaps through an advisory council with participation from industry giants—to guarantee that it remains relevant.

The goal is to develop a sophisticated, multi-layered platform that transforms the interaction between students, colleges, and businesses. SATHI has the potential to be a truly revolutionary force in Indian education, promoting better alignment, cutting waste, and unleashing the enormous potential of the country's youth.



CHAPTER 4

Rethinking Entrance Exams — My Vision for a Student-Centric Future

Dr. Sadhana Prashar

Senior Director, National Testing Agency (NTA)

In my role at National Testing Agency, I was able to observe the immense scope of aspiration and concern that characterizes the Indian education system today because of my position as Senior Director of the National Testing Agency (NTA). Millions of young Indians take important admission exams like the CUET, JEE, and NEET each year. Administratively speaking, these are just called admission tests, but I now view them as something far more ephemeral—emotional minefields.

What bothers me is that a student’s entire future is frequently decided by one day’s work. I have personally seen students and their families fall under the weight of these expectations because of the tremendous pressure this current system places on them. I have personally heard tragic stories of our young people suffering from insomnia, panic attacks, and burnout. I am truly concerned about the fact that we have turned educational assessment into a “survival game,” which is a startling and disturbing conclusion.

Earlier than ever before, the effects of this high-pressure environment are apparent. Test preparation workshops are being offered to kids as early as sixth grade. Thus, entire ecosystems of “mock schools” and coaching facilities have developed. Their goal is to “crack” the system rather than educate in the traditional sense. The regrettable result of this commercialization is the widening, noticeable divide between the rich and the poor.

The harsh reality is that not everyone is treated properly by our existing system. Students from non-native English speakers, state board backgrounds, and rural places are clearly at a disadvantage. People who were raised in urban, English-medium surroundings still have a disproportionate advantage in achievement, even though we provide assessments in other languages.

Here is a broader systemic issue that goes beyond these socioeconomic differences. Instead of assessing cognitive talents, we are increasingly focusing on memorizing skills. We are not assessing introspection, creativity, or potential. This strategy, in my opinion, wastes the human potential of our country in addition to being unfair.

This is the exact reason I think the SATHI program is so important. SATHI is simple but revolutionary: I want to fundamentally shift the paradigm from rejecting kids to selecting them, starting with “pulling them forward” instead of “filtering out” them.

Our tests should provide pupils the encouraging message, “This is what you are good at, and here is where you can develop.” A good test should really aim to open doors rather than close them. It is crucial that this system feels uniquely Indian even though we are developing it to match international standards like the SAT or GRE. It must be equal to a student from a small Haryana village, relevant to a learner in Telangana, and accessible to a student in a remote Uttarakhand village.

What, then, ought a decent test to measure? Instead of asking students to repeat facts they have learned from textbooks, I think we should be examining how they think. Aptitude—problem-solving, abstract reasoning, language and spatial skills, and yes, even creativity and emotional intelligence—is the main focus of SATHI.

This construction is made to be flexible. For example, the system may indicate that a student is headed for a profession in architecture or design if they demonstrate exceptional spatial intelligence. They might excel in research or the pure sciences if they have the capacity for abstract thought. Instead of forcing a way, we are helping people find it.

Inclusion must be incorporated by design in order for this system to be genuinely equitable. We are making certain that all 22 of the scheduled Indian languages will be allowed to take the SATHI exam.

Integrate biometric authentication, video proctoring, and a dynamic item bank with continuously changing questions to ensure security. This implies that each test will be unique.

Most importantly, SATHI is made to let students unwind. Students should be able to bank their greatest score and take the test more than once a year; it shouldn’t be a one-time affair. Just that flexibility lowers the pressure somewhat. Each test form is algorithmically built for uniqueness, and the infrastructure is being created to work flawlessly on PCs, tablets, and smartphones—even in locations with intermittent access.

The kind of report that students will receive is one of the things that most excites me. The object will be a mirror, not just a score. The scorecard will include percentile-based skill profiles, diagnostic remarks on areas of strength and weakness, recommended career or educational pathways, and even a brief psychometric personality test. In addition to students, this thorough input will help parents, teachers, and counselors provide better guidance to young people.

According to early conversations with vice-chancellors and admissions directors, many of them are ecstatic about the potential of SATHI to streamline the admissions process, offer consistency, and support the National Education Policy 2020, which encourages comprehensive and multilingual assessments.

This endeavor satisfies industry demands and employability standards in addition to being intellectual. Using SATHI’s descriptive profiles, employers may eventually choose candidates based on their

capacity for ethical reasoning or problem-solving. By comparing these scores with job performance, HR teams might be able to map leadership potential or domain fit.

SATHI is more than merely an exam system, in my view. Our youth need a system that sees them as whole beings, not only as statistics and titles, but as bright minds and kind hearts. That is my goal for the future.



CHAPTER 5

Lessening the Tyranny of One-Shot Exam

Prof. Janat Shah

Former Director, IIM Udaipur

There are significant gaps in the current admission exam landscape, in my opinion. Tests such as the SAT and GRE are essentially creations of the United States. They are expensive, heavily reliant on coaching, and make it difficult for many students, especially those from outside North America, to get in.

Why I Think the Exams We Take Now Aren't Working

Nearer to home, Till probably 2005, the CAT excelled in identifying the top 1% or 2%, which was its intended purpose. In the mid-percentile range, however, it was far less effective – Current CAT does not sharply differentiate at the top. The “bad day” effect also affects it; if a student has one bad day, their odds of succeeding plummet. Unfortunately, there isn't enough solid validity and reliability data available to the public to reassure stakeholders of its scientific rigor, and the coaching arms race surrounding it has produced a culture of repetitious tactics.

IPMAT and other institution-specific exams are frequently merely miniature CATs. They seldom ever delve further into more general characteristics and don't map the multifaceted skills we ought to be looking for in young students.

How a Better Evaluation Might Appear

The next-generation assessment, in my opinion, needs to be human-centered, adaptive, and digital-first. It should permit repeated attempts and provide several approved windows. This improves our understanding of a candidate's actual capability and lessens the tyranny of the one-shot exam.

Higher-order critical thinking and argument analysis are added to the cognitive areas of verbal and mathematical reasoning, data interpretation, and more. However, we must do more. Strong measures of aptitude include tasks that assess spatial intelligence, gamified modules that track curiosity and risk-taking, and even a financial literacy component for graduates in commerce.

Making room for a “fifth answer” is one concept that really excites me. That is, letting students come up with a solution that goes beyond the possibilities that are provided. Divergent thinking is rewarded, the power dynamic is altered, and it is communicated that we favor originality rather than conformity.

The Essential Clusters of Skills That We Need to Assess

Any reliable evaluation needs to cut across discipline boundaries. We must assess critical thinking, data and visual interpretation skills, numeric literacy, reading comprehension and synthesis, and the capacity to spot trends and abnormalities.

Meta-traits like curiosity, risk tolerance, and a desire to learn are equally significant. These are more difficult to quantify, but they are crucial in a world where flexibility is just as vital as knowledge.

Developing Trust via Openness and Strictness

Without confidence, no exam can be successful. For this reason, I fervently support total transparency, including the sharing of studies on score stability and fairness, the publication of blueprints, and the release of reliability indices.

It is reasonable for universities to request proof that test results indicate academic achievement. In order for universities to analyze longitudinal results and correlate admission data with GPA, I envision us providing analytics toolkits. The required signaling effect will be produced by early adoption by “gold standard” institutions, such as leading Indian privates and up-and-coming ASEAN leaders.

Making It Ethical, Inclusive, and Accessible

Inclusion must be at the heart of a test created for a varied, international student body. Localized test centers, sliding fee rates, and scholarships are non-negotiable. Access and security must be balanced; home-based testing increases reach but increases fraud risks, thus technology and human auditing must work together.

I also pay close attention to avoiding the coaching spiral. Exam designers must therefore avoid using cryptic techniques and openly share sample forms so that students are aware of what to expect.

Where I Think This Will Go

In my opinion, the first phase of adoption involves undergraduate and a few postgraduate admissions to Indian private universities. From there, we may go into MENA, Africa, and South Asia, where there is a significant need due to the youth bulge. In the end, reciprocity with US and EU colleges would be fantastic, turning this into a genuinely worldwide standard.

The value case for schools is equally compelling: lower costs than administering institution-specific exams, a larger and more diverse applicant pool, and diagnostic feedback that can be used for academic guidance.

Restoring the Learner's Power

Fundamentally, this goes beyond simply developing a new test. Rebalancing the equation is the goal. The learner is at the mercy of the system, which has all the power these days. I want an evaluation that reverses that dynamic, where pupils are seen as complete individuals rather than as a set of numerical values.

If we do this correctly, we will not only create a better test but also a more fair, perceptive, and compassionate road for millions of youth to find their identity and potential.

CHAPTER 6

What If Talent Assessment Were Rethought for a New World?

Bimal Rath

Founder & Chairman, Think Talent Services

This is what I have learnt from my decades in HR and leadership positions: we need a drastic rethink of how we evaluate talent in India, both students and working professionals. Our current approach to assessing potential is out of step with the direction we are taking. Pass/fail binary systems are still in use today, when systems that take into account the full person are what we actually need.

Approach to assessment should be simple: stop excluding, being understanding. Not in assessments that merely declare someone unsuccessful. After completing an evaluation, each person should have new knowledge, a profile of their areas of strength and growth, and, ideally, a development plan that will help them in the future.

Consider it akin to a SWOT analysis for individuals—not merely a numerical value, but a sophisticated comprehension of your current situation and potential future directions.

What Is Important in the AI Age?

AI will perform the majority of the routine, left-brained tasks in the future. What does that leave us humans with, then? the capacity for cross-disciplinary thought, creation, reflection, and meaning-building. The idea of multiple intelligences is strong because of this. IQ is only one aspect of intelligence. It's physical expression, emotional sensitivity, musical inventiveness, spatial aptitude, and even existential awareness. Consider tests that measure traits that will be more important than merely solving for x , such as self-reflection, resilience, and purpose.

A two-tier architecture is what I see:

- A core battery that assesses fundamental behavioral and cognitive characteristics.

- Additionally, there are optional, modular tests that correspond to particular fields, such as kinesthetic for athletics, musical intelligence for the arts, or spatial reasoning for design.

This would make perfect sense. Employability in 2030 and beyond will be determined by these exact skills.

Why Today's Exams Are Not Up to Par

Exams like the JEE, SAT, and IELTS have a place, and I respect them. Let's acknowledge the obvious, though:

- They just have one dimension.
- They just pay attention to language and math skills.
- They provide little feedback for improvement and are elimination-centric.
- They also make you complete strict multiple-choice questions that don't allow for any originality or subtlety.

The majority of employers will agree that academic performance is not a reliable indicator of success in the real world. The things we're testing are incorrect.

A Test That Allows Individuals to Express Their True Selves

What, therefore, would my perfect evaluation entail?

- Initially, it would encompass non-cognitive qualities like tenacity, moral discernment, collaboration, and emotional intelligence.
- Second, the test-taker would regain agency. Allow them to select modules based on their areas of strength. Allow them to react as they see fit, whether it is with an essay, a video, a simulation, or even an original answer to an MCQ (multiple choice questions_ that wasn't one of the possibilities provided.
- Thirdly, it would facilitate ongoing diagnosis. Not a one-time event. Over time, allow people to monitor their progress, hone their approach, and develop self-awareness.

By using this approach, a student may improve their JEE score from 100,000 to 15,000—not just by working harder, but also by recognizing and filling in particular behavioral or cognitive weaknesses.

What Needs to Be Measured?

We must look well beyond academic intelligence. Our evaluation ought to include:

- cognitive skills such as problem-solving, logic, and pattern recognition.
- creative abilities such as design thinking, diverse thinking, and brainstorming.
- Critical thinking includes assessing claims and analyzing information.

- Behavior characteristics include resilience, emotional intelligence, and teamwork.
- Existential awareness includes inner balance and purpose clarity.

Additionally, there is domain-specific intelligence, such as bodily-kinesthetic, musical, and spatial.

Scientific Rigor Must Be The Foundation of This

The significance of testing across India's languages, regions, and socioeconomic categories cannot be emphasized enough. Robustness in psychometrics cannot be compromised. We will examine test-retest validity, item-response curves, alpha coefficients, and other related metrics.

Most importantly, though, we'll connect it to practical results like employer feedback, college GPA, and internship success. We establish credibility in this way. It's also an ethical obligation.

This assessment model needs to:

- Be culturally neutral.
- Reduce test anxiety by using a developmental rather than punishing framework.

Building a framework for age-specific employability, helping young people find fulfilling careers, and providing mid-career professionals with a means of reskilling and reimagining themselves are all opportunities that I see as being available across the country.

I'll leave you with a few examples: Consider a young lady in rural Bihar who pays a small amount to take the test in Hindi at a nearby center. According to her study, she had great kinesthetic and spatial intelligence. She is directed toward an industrial design vocational program.

Or consider a candidate responding to an MCQ with a fifth original answer. A company seeking innovation in AI ethics has shortlisted her after the scoring engine acknowledges her uniqueness.

It's more than simply an exam. It serves as a forum for career guidance, self-discovery, and showcasing national talent, changing the focus from gatekeeping to mentoring and from universal growth to customized development.

CHAPTER 7

A Principal's Perspective on Educational Transformation

S. Sriram

Principal, The Mann School, New Delhi

Recently, I engaged in a profound discussion with members of SATHI (Scholastic aptitude test for Higher education Institutions - an initiative designed to move beyond mere domain skills and assess the crucial 21st-century competencies of collaboration, communication, critical thinking, and creativity of students.

The conversation prompted me to reflect on the systemic challenges and transformative opportunities facing school education in India today.

The Great Disconnect: Our Fragmented System

The current educational landscape operates in three distinct, disconnected silos: the school, the higher education institution, and the industry. Each works with its own patterns, addresses its own problems, and maintains very little meaningful connection with the others. This fragmentation is, in my view, the primary source of the systemic failure we witness. Schools are engrossed in their own world, higher education institutions lament the quality of incoming students without actively engaging with schools, and the industry constantly points to the deficit between what universities produce and what the modern workplace demands.

This disconnect is obvious in a young student's transition from school to college. The school system, by its very nature, is a teacher-dependent, fully taught system focused on syllabi and examinations. Many students fall through the higher education system, not due to a lack of passion or intelligence, but because they were never equipped with the aptitude for self-directed learning. This represents a colossal waste of time, potential, and institutional resources. Furthermore, our higher education system, with a few exceptions, lacks a robust, integrated research component at the master's level, which is standard in most foreign universities. This absence of deep-dive research leaves our graduates with superficial knowledge, ill-prepared for the in-depth problem-solving required by the industry.

Beyond Compliance: A Call for Curiosity and Connection

I feel the core problem lies in the very premise of our schooling system. We operate on a foundation of what I call the two 'C's: **Compliance** and **Completion**. Teachers chase the syllabus, and the entire ecosystem is geared towards exam scores. This leaves precious little room for exploration, for wonder, for the two 'C's that truly define education: **Curiosity** and **Connection**. We are producing a generation that knows much but understands little; a generation that is informed but not truly educated. In essence, we are practicing literacy—the acquisition of information—not education, which is the ability to use that information with wisdom and creativity.

When a child's natural curiosity is not nurtured, they cannot form meaningful connections—between different subjects, between classroom theory and real-world application, or between their own learning and their personal passions. In such a rigid, compliance-driven environment, we inadvertently value conformity over creativity, turning potential innovators into dutiful implementers - is a profound loss for both, the individual and the nation.

Empowering the Educator as an Innovator

While all our reforms rightly claim to be student-centric, we neglect the most critical agent of change in the classroom: the teacher. Despite countless policy updates, classroom practices remain largely static. This is because our teachers are overwhelmingly trained to *deliver* content, not to *design* learning experiences. They are handed readymade lesson plans and expected to execute them.

I believe teachers must be empowered as innovators. At The Mann School, we consistently encourage faculty to create their own lesson plans and design their own teaching materials. When teachers become designers, they curate lessons that cater to the diverse capabilities within a classroom; craft experiences that are not just informative but also transformative. This requires a fundamental overhaul of our teacher training and professional development systems: move beyond outdated pedagogical theories and equip our educators with the skills to facilitate inquiry-based learning, foster collaboration, and truly mentor the next generation.

Integrating Life Skills and Real-World Exposure

Another area where we fall short is in career guidance. Treating this as a one-day career fair is a complete eyewash. It must be a structured, multi-year program woven into the school journey, starting from Grade 8 or 9. Life skills must be embedded within the curriculum. We run a program called the Mann School Finishing School (MSFS), where etiquette, manners, and a vast range of practical life skills are integrated into daily subject teaching. For instance, a chemistry lesson on concentrations can be connected to understanding the difference between a jam and a marmalade. A history lesson on diplomacy can segue into a discussion on modern professional communication. This approach ensures students *live* these competencies rather than just learning about them. Furthermore, every teacher should be trained as a career counselor for their subject. Who better to discuss the future of physics than a physics teacher?

A Vision for a Transformed Future

Looking ahead, our strategic priorities for transformation are clear. First, we need a complete pedagogical redesign, shifting from monologue-based delivery to an inquiry-based, experiential, and collaborative learning. Second, build a national ecosystem for teacher excellence, creating professional learning communities where educators can share innovations and best practices.

Finally, technology and AI, as envisioned by platforms like SATHI, hold immense promise for personalizing learning and providing real-time feedback. However, deploy it ethically, with robust data privacy and a commitment to ensuring equitable access for all, including the untapped potential in Tier-2 and Tier-3 cities. Build a system that nurtures not just intelligent, capable minds, but also empathetic, ethical, and emotionally intelligent human beings who thrive in and contribute to a complex world.

PART III

CHAPTER 8

The Report 2025

This section deals with the mechanics of mySATHI assessment – starting with the clear cut definitions and explanations on the 4C framework, its philosophy and its various dimensions. It seeks to explain various aspects of Critical Thinking, Creativity, Collaboration and Communication – as a skills matrix.

The research team also provides insights into how the team empanels subject matter experts (SMEs) for creating the test architecture, the actual questions, its authoring, the human and technology aspects that work behind the actual assessment. We also have compiled individual articles on the tech infra behind mySATHI and how such infrastructure allows learners / candidates to take examination on demand (whenever, wherever they can and want to)

The second part of the section is the central piece of the entire Research Report 2026 – providing insights into the pilot that was done for over 3500 school students who are preparing themselves to enter their first degree program in universities.

The Core researchers and statistical experts analyse the individual learners and also provide a deep analysis of the questions, the concept and difficult levels of the mySATHI pilot test. The Report also seeks to show how institutions and Corporates can move beyond uni-dimensional tests or entrance exams and embrace dashboards, feedback, and longitudinal insights. These tools not only enrich learning but also give institutions a competitive edge. Leaders who ignore this shift that higher education and employability is moving to, risk being left behind in an increasingly data-driven academic landscape.

The report also explains how students rarely receive meaningful feedback, leaving them blind to their strengths and weaknesses. Especially in the normal ecosystem where stress, illness, or circumstance on one exam day can derail a future.

mySATHI team seeks to showcase how a well-designed assessment dashboard transforms raw scores into actionable intelligence, offering:

- **4C Profiles:** Insights into critical thinking, creativity, collaboration, and communication.
- **Time-Accuracy Analysis:** Measuring efficiency under pressure—vital for professions like consulting, law, or medicine.

- **Trend Analysis:** Comparing performance over multiple attempts to track growth.
- **Personalized Guidance:** “You are strong in evidence-based reasoning, but need to improve communication clarity.” Such insights empower students to self-correct and faculty to personalize learning. They also allow universities to demonstrate accountability and transparency to regulators, employers, and society.



CHAPTER 9

Evaluating the 4Cs — A Psychometrically Based Framework for mySATHI Assessment

Today the world, communities and job markets are becoming more complicated. The need for skills that go beyond traditional academic success is also growing exponentially.

The 4Cs—critical thinking, creativity, cooperation, and communication—are now commonly seen as the basic skills needed to do well in school, at work, and as a citizen. The SATHI assessment system incorporates these four competencies as fundamental construct, with the specific objective of delivering a relevant, equitable, and future-oriented evaluation of learners.

International organizations and national education systems are increasingly aligned on the significance of these competencies, integrating them into curriculum objectives and evaluation frameworks. In this global setting, SATHI sees the 4Cs not as “soft skills” of secondary importance, but as key outcomes that change how students think, interact with others, and act in real-life circumstances. By putting the 4Cs at the center of its work, SATHI aims to help students grow in all areas, encourage fair evaluation processes, and make sure that its design is in line with the needs of a world that is changing quickly. This paper introduces the fundamental principles that govern the integration of the 4Cs into SATHI, the conceptual framework for each competency, and the psychometric methodology employed for their measurement.

1. The 4Cs’ Main Ideas in the SATHI Assessment

- 1.1 Holistic Development above Memorization and Repetition:** A key design element of the SATHI framework is to prioritize holistic development over rote performance. Assessment focuses on cognitive processes, attitudes, and competencies instead of superficial information recall. The 4Cs encompass cognitive, social, and expressive aspects of learning, facilitating a holistic perspective on learner development. This approach lets SATHI focus on how students think, create, work together, and talk to each other, rather than just what they can produce on a timed test.
- 1.2 Real-World Application and Transferability:** Tasks based on the 4Cs are designed to mirror real-life situations. Students are tested on how well they can use what they know, solve issues, cooperate with others, and share their thoughts in situations similar to those

they would encounter in school, work, or social settings. This emphasis on authenticity frames assessment results as measures of preparedness for real-world situations, rather than merely as isolated exam performance.

- 1.3 Process-Based Assessment:** The SATHI framework clearly values the steps that learners take to reach their goals. The 4Cs stress reasoning, idea generation, interaction, and self-expression. This makes process data an important part of understanding a learner's growth. SATHI promotes introspective thinking, metacognition, and ongoing progress by focusing on the journey rather than just the destination. Assessment is a way for students to practice and think about higher-order thinking in a systematic way.
- 1.4 Fairness and Inclusion:** Another guiding idea is the promise to be fair and open to everyone. The 4Cs offer various avenues for learners to exhibit proficiency, catering to a range of backgrounds, learning modalities, and capabilities. The framework aims to reduce prejudice and provide all learners more chances to exhibit what they know by cutting down on language-heavy or recall-based tasks. This inclusive design helps make better decisions and ensures a more fair distribution of recognition and opportunity.
- 1.5 Learner Agency and Ownership:** The SATHI assessment sees students as active participants in their own learning. The 4Cs promote agency by encouraging students to make choices, share their thoughts, and work together. They do this by focusing on inquiry, creativity, teamwork, and expression. Assessment is viewed as an instrument that can enhance learners' sense of ownership over their education, rather than simply a means of external appraisal.
- 1.6 Compatibility with Future Preparedness:** The 4Cs are clearly connected to being ready for the future. The framework meets the needs of today's schools and workplaces by focusing on critical thinking, innovative problem solving, and clear communication. Students get help not only with reaching academic goals, but also with building the flexibility, social skills, and resilience they need to deal with changing life and job situations. These principles make the SATHI assessment developmental, learner-centered, and future-focused.

2. The 4Cs' Conceptual Context and Definitions

The idea behind the 4Cs in SATHI is that they are high-level, transferable skills. Each C stands for a different but connected area that helps learners deal with complicated situations.

- 2.1 Critical Thinking:** The SATHI framework defines critical thinking as the ability to look at information, weigh options, and make fair, well-reasoned choices when there are opposing claims or insufficient data. Students should challenge what they think they know, find gaps or biases, and use what they know in new situations. The evaluation of critical thinking is related to the goal of developing independent and reflective learners.

- 2.2 Creativity:** Creativity is the capacity to produce novel, functional, and contextually relevant ideas through the integration of knowledge and imagination. It encompasses not just artistic expression but also scientific, technical, social, and entrepreneurial fields. In SATHI, being creative means coming up with a lot of different ideas, trying out new ways of doing things, and finding workable solutions.
- 2.3 Collaboration:** Collaboration means being able to work well with others, share responsibility, and come up with solutions together by using everyone’s abilities. SATHI stresses the social and interpersonal aspects of learning because it knows that real-world problems need people to work together, respect each other, and be responsible for each other’s actions. Assessment looks at interpersonal skills, empathy, and leadership.
- 2.4 Communication:** Communication is the ability to clearly share, understand, and exchange ideas with others in different situations and media. It includes speaking, writing, and digital literacy. To communicate well, one needs to be clear, logical, and listen carefully. SATHI wants to ensure learners can express their understanding and participate in academic and professional discourse with confidence.

3. Psychometric Design and Dimensional Structure

To make the 4Cs useful for assessment, SATHI uses a dimensional structure that shows both the breadth and depth of each capability while remaining reliable and easy to understand. Each C is broken down into specific dimensions to prevent fragmentation into micro-skills.

Psychometric Accuracy

The psychometric design based on Item Response Theory (IRT) supports measurement within SATHI. Test items are chosen from a wide range of subjects and are made to fit the unique skills needed for each of the 4Cs. Before being put into use, items go through pilot testing to make sure they meet high criteria of validity and reliability. This ensures every response gives an accurate and fair evaluation of the skill being tested.

The SATHI assessment framework is built around these 4Cs to redefine evaluation as a developmental process. By using precise conceptual concepts and an IRT-based psychometric methodology, SATHI facilitates the measurement of these competences in a manner that is reliable and pedagogically significant. These dimensions provide the basis for a global standard. Instead of taking the place of subject knowledge, the 4Cs improve the use of disciplinary knowledge, helping students connect with content on a deeper level and apply what they know to complex, real-world situations.

CHAPTER 10

Subject Matter Expert (SME) Empanelment for Assessment and Authoring of Assessment Questions

Introduction

Perhaps one of the most crucial part of any assessment process is the rigor and comprehensiveness that is used for the identification, shortlisting, empanelment, and engagement of Subject Matter Experts (SMEs) for question authoring and review activities. The singular objective is to ensure the development of high-quality, valid, reliable, and standards-compliant assessment items (questios) across diverse subject domains for high-stakes examinations.

mySATHI's framework is clear that this objective has to be met each and everytime and this is relevant and demonstrated across every cohort, every session and every question paper.

In this section, the research report explain the process for:

- Identification and shortlisting of SMEs
- SME empanelment methodology
- Question authoring guidelines and workflows
- Multi-level review and quality assurance mechanisms
- Governance of a large and diverse SME panel spanning more than 90 subject domains

Guiding Principles

The SME engagement and item development framework is aligned to globally accepted assessment and measurement standards, ensuring validity, reliability, fairness, and defensibility of test content. The guiding principles include:

- Alignment to Bloom's Revised Taxonomy (Cognitive Process & Knowledge Dimensions)
- Adherence to psychometric principles of validity, reliability, and objectivity

- Academic credibility and institutional integrity
- Domain depth and pedagogical expertise
- Standardization and fairness in assessment
- Confidentiality, ethics, and compliance
- Continuous quality improvement

Alignment to Assessment Standards and Psychometric Framework

Bloom's Revised Taxonomy Alignment

All items authored and reviewed under this framework are explicitly mapped to Bloom's Revised Taxonomy, ensuring appropriate cognitive coverage:

- **Remember** – Recall of facts, definitions, and basic concepts
- **Understand** – Interpretation, explanation, and summarization
- **Apply** – Use of knowledge in new but structured situations
- **Analyze** – Differentiation, organization, and problem decomposition
- **Evaluate** – Judgement based on criteria and standards
- **Create** – Synthesis and generation of new constructs (where applicable)

SMEs are instructed to design questions beyond mere recall, with a mandated proportion of items targeting **application and higher-order thinking skills**.

Knowledge Dimension Mapping

Items are also categorized across knowledge dimensions:

- Factual Knowledge
- Conceptual Knowledge
- Procedural Knowledge
- Metacognitive Knowledge (limited to advanced assessments)

This dual-axis mapping (cognitive level × knowledge type) ensures balanced test blueprints.

Psychometric Design Principles

The item development lifecycle incorporates the following psychometric rigor:

- **Content Validity:** Items mapped to syllabus, learning outcomes, and test blueprint
- **Construct Validity:** Items measure intended competencies without construct-irrelevant variance

- **Reliability:** Consistency ensured through standardized authoring and multi-level reviews
- **Objectivity:** Clearly defined correct responses and scoring rules
- **Fairness and Bias Review:** Language, cultural neutrality, and accessibility checks

Difficulty and Discrimination Considerations

The subject matter experts (SMEs) are guided to design items across calibrated difficulty levels:

- Easy (baseline competency)
- Moderate (application and reasoning)
- Difficult (complex reasoning, multi-concept integration)

Senior reviewers evaluate potential item discrimination power, ensuring high-quality distractors and avoidance of cueing.

Institutional Sources of SMEs

The subject matter experts (SMEs) are primarily drawn from government-funded and government-recognized institutions of high repute, including:

- Central Universities
- State Universities
- National Institutes and Deemed Universities under statutory bodies

Preference is given to faculty members and academic experts with demonstrated teaching, research, and assessments experience.

Stage 1: SME Identification and Shortlisting

Eligibility Criteria

Potential SMEs are shortlisted based on:

- Relevant postgraduate or doctoral qualifications
- Minimum prescribed years of teaching/research experience
- Subject specialization aligned with examination syllabus
- Prior experience in assessment, evaluation, or question setting (where available)

Initial Screening

Applications and nominations are screened for:

- Academic credentials
- Institutional affiliation
- Subject domain relevance

Stage 2: Orientation and Item Generation Guidelines

SME Orientation

Shortlisted SMEs are provided with a structured orientation covering:

- Examination objectives and assessment purpose (selection, ranking, or qualification)
- Overview of Bloom’s taxonomy and cognitive targeting
- Psychometric quality indicators and common item-writing flaws
- Expectations regarding quality, confidentiality, and timelines

Item Generation Guidelines

Detailed item authoring guidelines are issued, explicitly aligned to assessment standards:

- Cognitive level tagging (Bloom’s taxonomy)
- Knowledge dimension classification
- Avoidance of construct-irrelevant difficulty
- Language clarity, simplicity, and bias avoidance
- Distractor functionality and plausibility
- Negative marking and guessing minimization strategies

Supported Question Types and Assessment Alignment

Question Type	Primary Cognitive Levels	Psychometric Notes
MCQ	Remember – Analyze	High reliability, strong distractor design required
MRQ	Understand – Evaluate	Controls guessing, higher discrimination
True/False	Remember – Understand	Used sparingly due to guessing probability
Match the Following	Understand – Apply	Assesses association and classification
Numerical Response	Apply – Analyze	Eliminates cueing, high construct validity

Stage 3: Sample Question Development

Sample Item Submission

Each shortlisted SME is required to submit a predefined number of sample questions strictly adhering to the guidelines.

Documentation Requirements

Sample submissions must include:

- Correct answer(s)

- Detailed solution or rationale
- Source/reference where applicable

Stage 4: Expert Review and Evaluation

Review by Senior Experts

Submitted sample items are evaluated by senior subject experts against defined quality benchmarks:

- Conceptual accuracy
- Cognitive appropriateness
- Language and structure
- Absence of ambiguity or bias

Evaluation Outcomes

Each SME is assessed as:

- Recommended for empanelment
- Recommended with conditions (minor improvements)
- Not recommended

Stage 5: SME Empanelment

Empanelment Approval

SMEs whose sample items meet prescribed standards are formally empanelled.

Empanelled SME Database

Empanelled SMEs are onboarded into a centralized repository capturing:

- Domain and sub-domain mapping
- Question type proficiency
- Past performance and review scores

Currently, the empanelled panel comprises **800+ SMEs**.

Subject Domain Coverage

The empanelled SME pool spans more than **100+ subject domains**, broadly categorized under:

- Science
- Engineering and Technology
- Management
- Medical and Paramedical Sciences
- Law
- Arts and Humanities

This ensures depth, scalability, and redundancy across disciplines.

Stage 6: Live Question Authoring

Assignment of Authoring Tasks

Empanelled SMEs are assigned question authoring tasks based on:

- Domain expertise
- Prior performance
- Examination requirements

Quality Controls During Authoring

- Structured templates
- Mandatory solution and justification fields
- Version control and traceability

Stage 7: Multi-Level Question Review

First-Level Review (Subject Accuracy and Alignment)

Conducted by senior domain experts focusing on:

- Conceptual correctness
- Bloom's level appropriateness
- Knowledge dimension accuracy
- Compliance with item-writing guidelines

Second-Level Review (Psychometric and Quality Assurance)

A dedicated quality assurance layer evaluates:

- Alignment to test blueprint
- Difficulty balance across the paper
- Distractor efficiency and option symmetry
- Redundancy and overlap detection
- Bias, sensitivity, and accessibility considerations

Pre-Operational Readiness Checks

Before final approval, items undergo:

- Consistency and formatting validation
- Metadata completeness (domain, sub-domain, cognitive level)
- Version control and audit trail verification

Governance, Confidentiality, and Ethics

- All SMEs adhere to strict confidentiality and non-disclosure norms
- Conflict of interest declarations are mandatory
- Intellectual Property Rights (IPR) transfer
- Secure platforms are used for content exchange

Continuous Improvement and Panel Management

- Periodic performance evaluation of SMEs
- Refresher guidelines and updated best practices
- Expansion and diversification of the SME panel based on emerging domains

Conclusion

The structured and multi-layered SME shortlisting, empanelment, question authoring, and review process ensures the creation of robust, fair, and defensible assessment content. The scale of the empanelled panel, depth of domain coverage, and rigor of quality controls collectively reinforce the credibility and reliability of the examination ecosystem.

CHAPTER 11

Tech Infra Overview — mySATHI Application Connectivity

1. Executive Summary

Here is a comprehensive overview of the network architecture supporting the **mySATHI application**. The architecture is designed to deliver secure, reliable, and scalable connectivity for online examinations conducted PAN India across DOTC (DEXIT Own Test Center) Centres, with centralized application hosting at the Primary Data Centre in Mumbai and a Disaster Recovery (DR) site in Noida.

The solution emphasizes strong security controls, high availability, disaster recovery readiness, and predictable performance. Connectivity from candidate endpoints is securely routed over redundant internet links to centralized application and database infrastructure. Real-time data replication between the Primary and DR sites ensures minimal data loss and rapid recovery in the event of failures.

The architecture has been validated through live deployments across multiple DOTC Centres, demonstrating low latency, stable performance, and operational resilience. This design aligns with organizational objectives for scalability, continuity, and examination integrity.

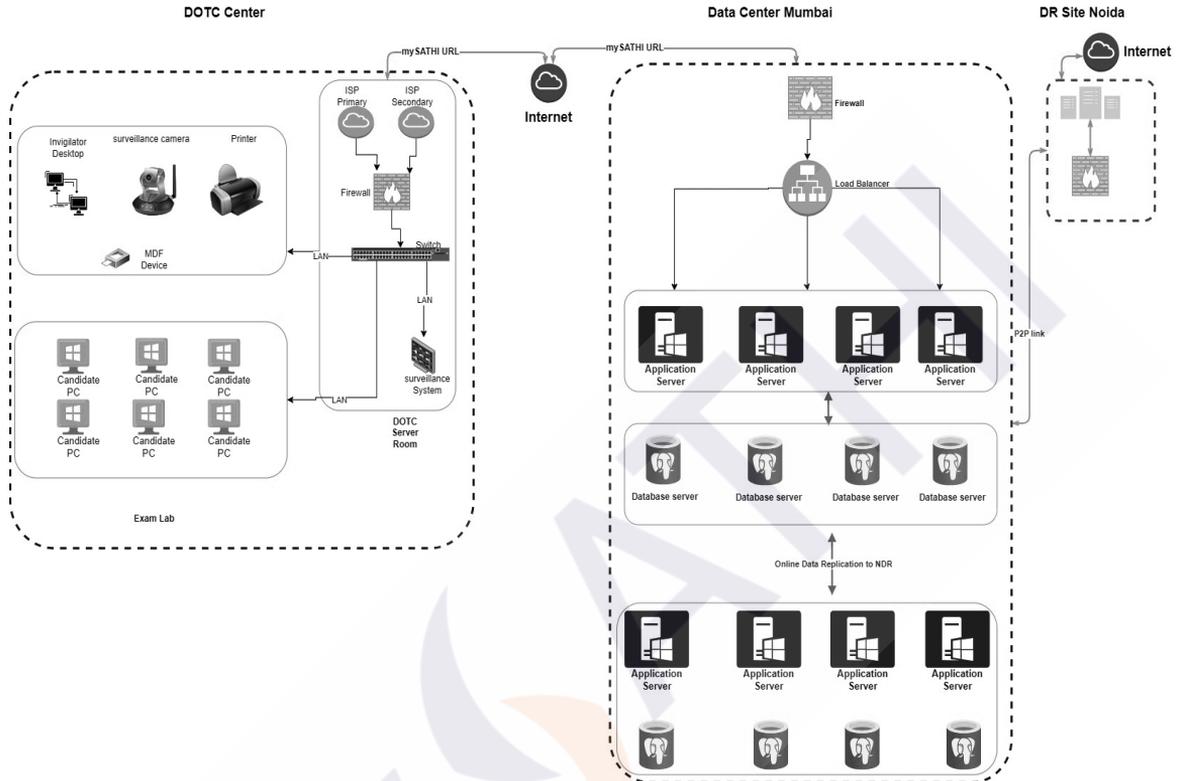
Note: mySATHI application is currently hosted in cloud infrastructure and will be used for phase 1 implementation. Infrastructure will be moved to on-premises as mentioned in document from phase-2 onwards.

2. Network Architecture Diagram

The following diagram illustrates:

- DOTC Centres hosting candidate examination environments.
- Secure connectivity over the Internet using dual ISPs.
- Centralized application and database hosting at the Primary Data Centre (Mumbai)

- Real-time data replication to the Disaster Recovery Site (Delhi)
- Firewalls, load balancers, application servers, database servers, and redundancy mechanisms



3. High-Level Architecture Overview

3.1 DOTC Centres

Each DOTC Centre hosts the local examination environment comprising:

- Candidate PCs
- Invigilator PC
- Surveillance cameras
- Printers
- A centralized MDF (Main Distribution Frame) device and network switch manage internal connectivity.

All devices are connected through a secure Local Area Network (LAN). The LAN is protected by a firewall and routed through **dual Internet Service Providers (Primary and Secondary)** to ensure uninterrupted connectivity. Candidates access the mySATHI application through designated URLs, with traffic securely routed to the centralized data centre.

3.2 Primary Data Centre – Mumbai

The Mumbai Data Centre acts as the central processing hub for the mySATHI application. Incoming traffic from DOTC Centres passes through a firewall and is distributed via a load balancer across multiple application servers.

The backend consists of multiple database servers responsible for persistent data storage. **Online data replication** is enabled to synchronize all critical data with the Disaster Recovery site. This centralized model ensures consistent performance, simplified management, and scalability.

3.3 Disaster Recovery Site – Delhi

The Disaster Recovery site mirrors the application and database infrastructure of the Primary Data Centre. It is connected via a secure **point-to-point (P2P) link**, enabling continuous data replication. In the event of a failure at the primary site, the DR site can assume operations with minimal downtime and data loss, ensuring business continuity.

4. Security, Availability, and Resilience

Security

- Firewalls at DOTC Centres and the Primary Data Centre control inbound and outbound traffic
- Network segmentation isolates examination devices from external threats.
- Secure communication channels protect inter-site data transfers.
- Controlled access to application and database layers

High Availability & Disaster Recovery

- Dual ISP connectivity at DOTC Centers eliminates single points of failure.
- Load balancing across multiple application servers.
- Redundant database servers
- Real-time data replication to the DR site

These mechanisms collectively support high uptime and uninterrupted examination operations, even during infrastructure or connectivity failures.

5. Bandwidth Requirements (Estimates)

DOTC Centre

- **Per Candidate PC:**
1–2 Mbps (download/upload) for application access, question loading, and answer submission

- **50 Candidates (Aggregate):**
50–100 Mbps
- **Surveillance Camera:**
2–5 Mbps per camera (HD streaming)
- **Total LAN Bandwidth:**
100–500 Mbps
(Gigabit Ethernet switches recommended)
- **Internet Uplink:**
Primary and Secondary ISPs at 100–200 Mbps each with burst capability

Data Centre – Mumbai

- **Internet Ingress:**
1–10 Gbps (based on total concurrent users)
- **Internal Bandwidth:**
10–40 Gbps between load balancer, application servers, and database servers
- **P2P Link to DR Site:**
1–5 Gbps for real-time data replication with low latency

Disaster Recovery Site – Delhi

- **Internet Uplink:**
1–10 Gbps
- **Internal Bandwidth:**
Sized to handle full production load during failover.

6. Performance Validation

Response time testing conducted across **15 DOTC Centres** showed application response times ranging between **10 ms and 100 ms** for opening the mySATHI URL and downloading question papers.

Response times may vary depending on question paper size and local network conditions.

7. Device Compatibility & System Requirements

(Executive Summary Table)

7.1 Minimum System Requirements (All Devices)

Parameter	Requirement
Desktop RAM	Minimum 4 GB (Recommended: 8 GB+)
Mobile RAM	Minimum 2 GB
Internet	Stable broadband or mobile data
Screen Resolution	Minimum 1280 × 720
Input Devices	Keyboard & mouse (Desktop), Touch (Mobile)

7.2 Desktop Compatibility – Windows

OS	Browser	Supported Versions
Windows 10 (22H2)	Chrome	108.0.5359.125 (64-bit)
	Firefox	108.0.1 (64-bit)
	Edge	108.0.1462.76 (64-bit)
	Opera	94.0.4606.61
	Brave	1.45.131 (Chromium 108)
Windows 11 (22H2)	Same as above	TPM 2.0 required

Hardware Requirements

Component	Requirement
Processor	Intel Core i3 / AMD Ryzen 3 or higher
Architecture	64-bit recommended
Security (Win 11)	TPM 2.0 enabled

7.3 Mobile Compatibility – Android

OS Version	Browser	Supported Version
Android 9 / 10	Chrome	108.0.5359.128
	Firefox	108.0.2
	Samsung Internet	19.0.4107.49
	Edge	108.0.1462.48

Hardware

- Minimum RAM: 2 GB
- Screen size: 5 inches or larger (recommended)

7.4 Mobile Compatibility – iOS

Device	OS	Browser	Version
iPhone	iOS 12 / 13	Safari	12.x / 13.x
iPad	iPadOS 13	Safari	13.x
iPhone / iPad	iOS 12 / 13	Chrome	108.0.5359.124
iPhone / iPad	iOS 12 / 13	Edge	108.0.1462.47

Hardware

- Minimum RAM: 4 GB
- Apple-supported devices only

CHAPTER 12

AI Based Remote Proctoring v 3.0

1. Proctoring AI with Remote Proctoring?

Layer	Description	Core Software Functions
Remote Proctoring Model	Allows candidates to sit an exam outside a physical test centre. Supervision is performed by a combination of software agents and (optional) human reviewers.	Secure browser/VM launch, device integrity checks, encrypted streaming.
AI Intelligence Layer	Continuously analyses multimodal signals (video, audio, screen, device telemetry, behavioural biometrics) to flag cheating in realtime and generate audit ready evidence.	Face ID & liveness, eyegaze & headpose, keystroke timing, ambient audio classification, process list monitoring, anomaly fusion, deepfake detection.

2. Existing Approaches

2.1 Traditional and Early Remote Proctoring

The transition from in-person invigilation to remote examinations has exposed a critical trust gap, as early remote proctoring solutions remain costly, hard to scale, and vulnerable to modern, AI-assisted fraud. With remote assessments now foundational to education, licensure, and recruitment, the core problem is not isolated cheating, but the absence of a continuous, auditable trust framework for remote human-machine interactions.

2.2 Current AI-Based Proctoring

Recent systems incorporate computer vision, audio analysis, and basic behavioural monitoring. These solutions represent meaningful progress; however, most remain event-driven and reactive, rely on limited signal sets, and struggle against adversarial techniques such as synthetic media, AI-generated answers, and coordinated collusion.

Key limitation: Existing solutions largely treat proctoring as a *monitoring problem*, whereas it is fundamentally a *digital trust and assurance problem*.

3. Vision: Remote Proctoring as a Digital Trust System

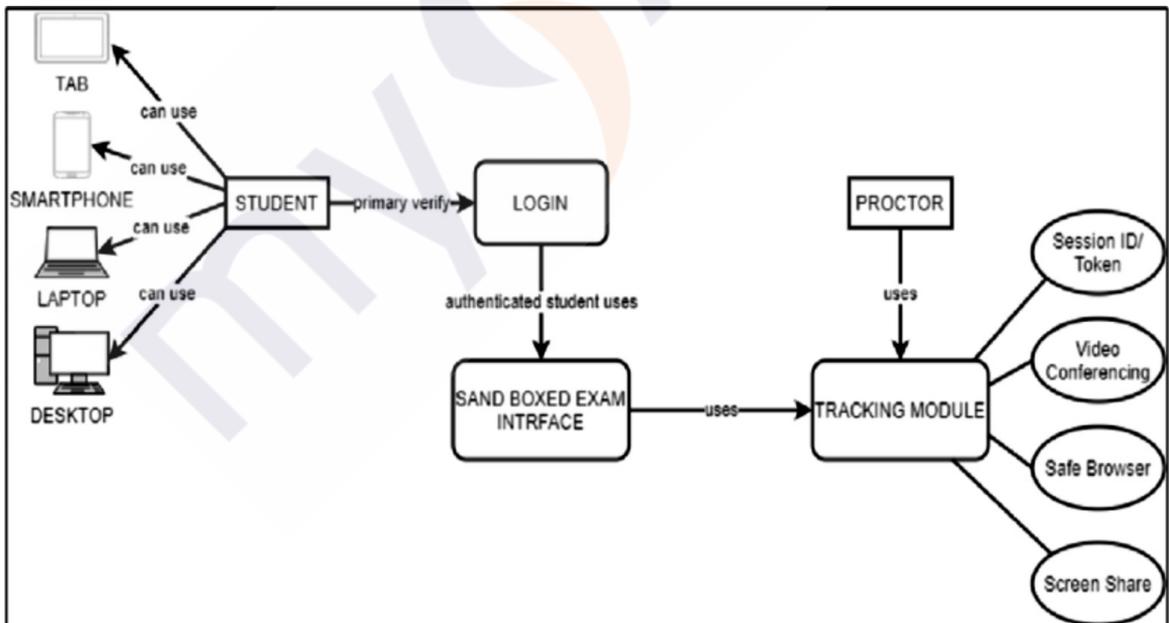
The future of remote proctoring lies in reframing it as a continuous AI-driven trust framework rather than a surveillance mechanism. The objective is not constant observation, but the ongoing inference of identity, intent, and environmental integrity, supported by auditable evidence.

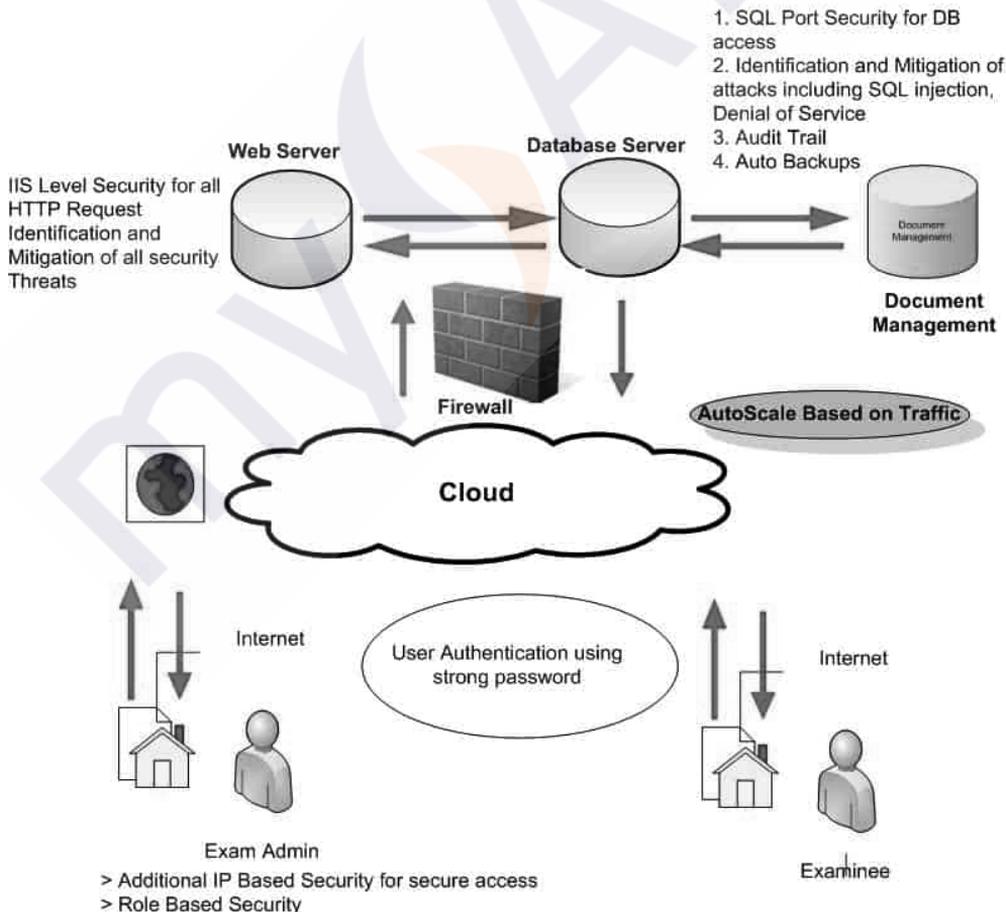
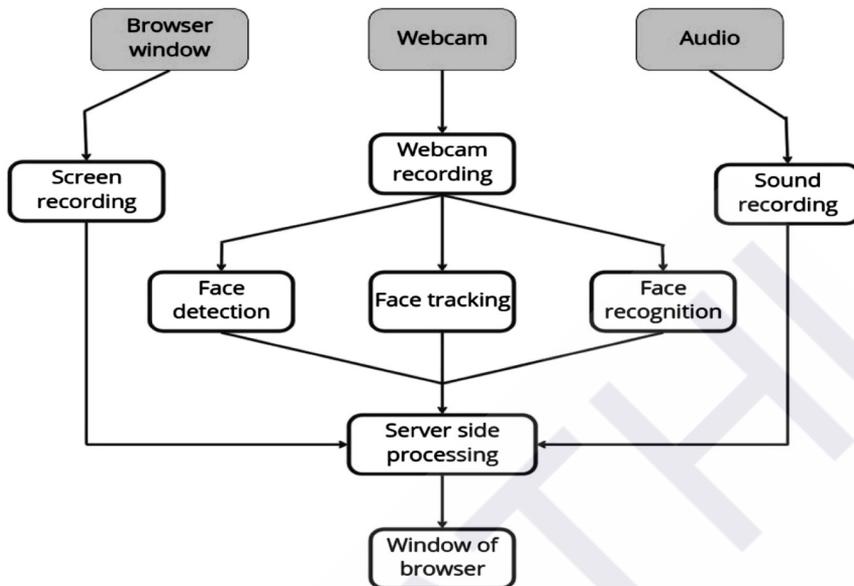
This vision shifts the paradigm:

- From single-signal alerts to multimodal behavioural inference
- From static rules to contextual risk modelling
- From post-fact detection to real-time deterrence
- From human-only judgment to human-AI collaborative governance

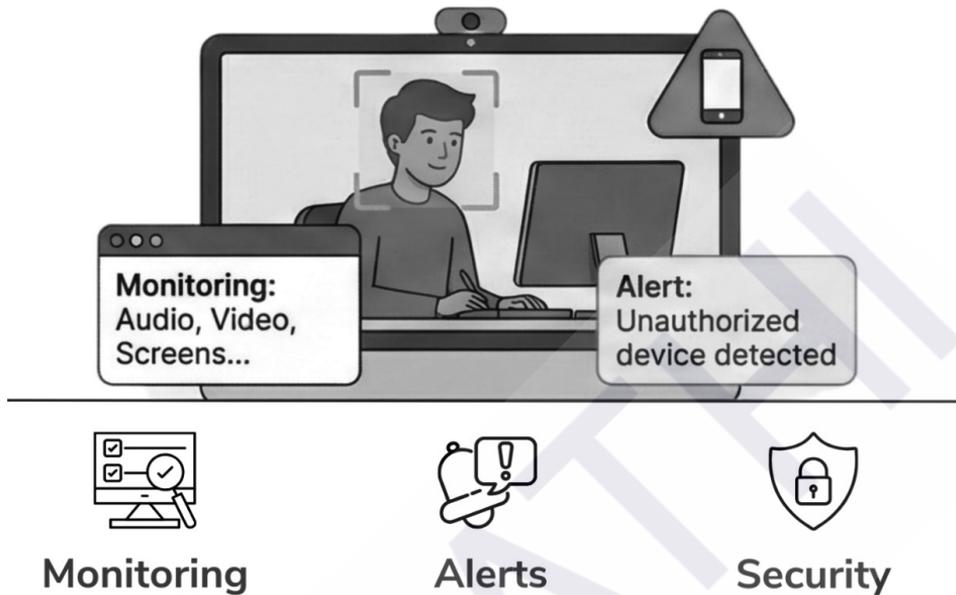
4. End-to-End Remote Proctoring Architecture

Here is a visualisation of the remote proctoring architecture from the exam hosting sever to candidate home terminal (see the accompanying three diagrams)





AI Proctored Assessments



Given below is a description of the end-to-end architecture for AI-Enabled Remote Proctoring

The examination is hosted on a secure central infrastructure, deployed either on cloud or on-premises data centres, and delivered to the candidate's home terminal over the internet using encrypted communication channels.

Hosting & Control Layer (Cloud / On-Prem)

Exam content, question banks, and session logic are hosted within a secure data centre.

Security controls include identity management, encryption, access control, logging, and compliance enforcement.

A central orchestration service manages candidate sessions, policies, and escalation workflows.

AI & Analytics Layer

Multimodal AI engines analyse real-time inputs such as video, audio, screen activity, device telemetry, and behavioural signals.

AI correlates these signals over time to generate risk scores and explainable alerts, rather than relying on single-event detection.

Alerts are prioritized and routed to human proctors for review.

Human Proctor & Alert Management Layer

Proctors access a secure dashboard displaying live feeds, AI-highlighted anomalies, and event timelines.

AI acts as a decision-support system, enabling proctors to focus only on high-risk candidates.

Proctors can issue warnings, initiate live challenges, pause exams, or terminate sessions when required.

Candidate Home Terminal

The candidate accesses the exam through a secure browser or native proctoring client.

The client enforces local security controls (browser lockdown, device checks, screen protection).

Webcam, microphone, screen activity, and input behaviour are captured, encrypted, and streamed securely to the hosting platform.

Secure Communication & Evidence Layer

All data exchanged between the candidate terminal and hosting environment is encrypted end-to-end.

Time-stamped logs, recordings, and AI decisions are stored in a tamper-evident audit repository for post-exam review and compliance.

Overall, this architecture demonstrates how a centrally hosted examination platform securely extends into a candidate's home environment, while AI continuously assists human proctors by detecting risk, reducing manual monitoring effort, and preserving exam integrity at scale.

5. Future Evolution for RP Solutions: Near-, Mid-, and Long-Term Evolution

5.1 Near Term: Multimodal AI Assurance

The immediate focus is the deployment of multimodal AI systems that fuse synchronized video, audio, screen activity, behavioural biometrics, and system telemetry. Temporal models analyse patterns over time rather than isolated events, producing explainable risk scores that guide human intervention. Post-exam forensic analytics—including answer similarity graphs and stylometric analysis—provide audit-ready validation.

Outcome: Improved accuracy reduced false positives, and scalable human oversight.

5.2 Mid Term: Zero-Trust Exam Execution

The next phase introduces hardware-rooted trust and zero-trust execution models. Exam runtimes are delivered through containerized or micro-VM environments verified using TPM or secure enclave attestation. Continuous device health checks and federated learning allow models to improve without centralizing sensitive raw data.

Outcome: Exam environments that are *provably untampered*, rather than assumed secure.

5.3 Long Term: Continuous Human Authentication

In the long term, remote proctoring evolves toward continuous human authentication and intent inference. Behavioural biometrics, optional wearables, and population-level anomaly detection enable systems to reason about identity consistency, stress patterns, and coordinated behaviour across cohorts.

Outcome: Assessments transition from discrete supervision events to continuously trusted digital interactions.

6. Futuristic Enhancements Enabled by Additional Hardware

Hardware Technology	Forward-Looking Capability
TPM / Secure Enclave	OS integrity attestation, VM and kernel tamper prevention
Secondary Camera Devices	360° environmental awareness, blind-spot elimination
Eye-Tracking Hardware	High-precision attention and gaze validation
RF / Signal Sensors	Detection of covert communication devices (research phase)
Biometric Wearables	Continuous identity assurance and stress correlation

These enhancements move the system toward ambient, low-friction trust enforcement, reducing reliance on intrusive controls.

7. Constraints and Open Challenges

Despite its promise, AI-driven remote proctoring faces unresolved challenges:

Privacy and Ethics: Balancing integrity with individual privacy, informed consent, and data minimization.

Bias and Fairness: Addressing demographic variability, disabilities, and cultural differences in behavioural modelling.

Accessibility: Supporting low-bandwidth environments and avoiding mandatory hardware dependencies.

Adversarial Evolution: Sustaining effectiveness amid rapidly advancing AI-based cheating techniques.

These constraints necessitate continuous research, transparent governance, and human oversight.

8. Conclusion

Remote proctoring via AI is evolving from a tactical monitoring tool into a foundational digital trust infrastructure. By combining multimodal intelligence, zero-trust execution, and ethical human governance, future systems can preserve exam integrity while enabling truly global, scalable, and equitable assessments.

The challenge ahead is not technological feasibility, but responsible, forward-looking design. Institutions that invest early in this evolution will shape the standards for trustworthy digital credentials in the decade ahead.

References (Selected)

IEEE Transactions on Information Forensics and Security, *Multimodal Wearable Biometrics for Online Exam Integrity*, 2024

Nature Communications, *DeepFake Detection for Secure Video-Based Proctoring*, 2024

McKinsey & Company, *The Economics of Remote Exam Proctoring*, 2024

EU Digital Exam Integrity Directive, Official Journal of the European Union, 2025

Microsoft Research, *Zero-Trust Secure Execution for Online Assessments*, 2025

Journal of Security & Privacy, *Dynamic UI Perturbation to Detect Script-Based Cheating*, 2025

CHAPTER 13

Nationwide Pilot of the mySATHI Assessment – Design, Execution, Findings and Implications

The pilot test of mySATHI assessment exercise was conceptualised as a critical validation step prior to the nationwide rollout of the assessment in 2026. The pilot’s primary objective was to evaluate the robustness, relevance, and operational feasibility of the mySATHI assessment framework — anchored in the 4Cs model—across diverse student populations and geographies in India.

For this purpose, 27 cities across the country were selected to ensure regional and institutional diversity. Team mySATHI collaborated with Career Launcher (CL), one of India’s leading edtech organisations with an extensive physical presence and deep engagement in the entrance preparation ecosystem. Given CL’s strong footprint across these cities and its experience in conducting large-scale, standardised assessments, CL centres were entrusted with executing the pilot tests. The outcomes of these pilots were intended to feed directly into the validation study conducted by DEXIT Global and to inform decisions related to scale, delivery mode, and stakeholder communication for the full rollout.

The choice of Career Launcher as the implementation partner was driven by both strategic and practical considerations. CL centres primarily work with students of Classes XI and XII preparing for undergraduate admissions—precisely the cohort for whom mySATHI is most relevant as an early aptitude discovery and guidance tool. Since mySATHI is positioned in the context of university admissions and long-term academic decision-making, this alignment ensured access to an appropriate and engaged student base.

Additionally, CL’s operational maturity in conducting physical tests, managing student logistics, and coordinating with schools and parents made it a natural choice for a multi-city pilot. The collaboration also enabled SATHI to test its assessment in real-world, high-stakes academic environments rather than in controlled or artificial settings.

Pilot Design and Test Formats

The pilot testing was structured across three distinct test formats—Retest, Parallel, and Split—to allow for psychometric validation and reliability analysis:

Retest Format: To assess score stability and consistency over time.

Parallel Form: To evaluate equivalence between different versions of the test measuring the same constructs.

Split Form: To examine internal consistency and section-level robustness.

In total, 3,257 students participated across these three formats. The sample size was sufficient to generate meaningful insights into test behaviour, reliability, and student response patterns, while still being manageable within a pilot framework.

Mode of Delivery and Rollout Constraints

While the final mySATHI assessment is designed to be computer-based and adaptive, the pilot was deliberately conducted in physical mode, using printed question papers and answer response sheets. This decision was influenced by several contextual constraints:

The pilot window (October–December 2025) coincided with school examinations, academic schedules, and entrance preparation cycles, making extended digital lab access difficult in many centres.

Physical testing allowed for tighter invigilation control and uniformity during early-stage validation. It enabled participation from centres and schools with varying levels of digital infrastructure.

All response data was subsequently digitised, centralised, and analysed by DEXIT Global, which carried out the validation study.

Demographic Profile of Participants

Geographic and Regional Spread

Students participating in the pilot came from Tier 1, Tier 2, and Tier 3 cities, ensuring exposure to varied academic and socio-economic contexts. The selected cities were carefully distributed across North, South, East, West, and Central India, allowing the pilot to reflect regional diversity rather than being metro-centric.

However, analysis of the data indicates a higher concentration of students from North and NCR regions, which mirrors the stronger operational presence of CL centres in these areas. While this skew does not invalidate the findings, it is an important contextual factor when interpreting nationwide applicability.

School and Institutional Profile

The majority of students were from private, English-medium schools, predominantly CBSE-affiliated.

Participation from government and aided schools was limited, though notably, one CSR-supported school (indicating students are from the lower economic strata) was included, providing valuable insight into how students from under-resourced backgrounds engage with aptitude-based assessments.

Most students were currently enrolled in Classes XI and XII, consistent with CL's core student demographic and the university-admissions focus of mySATHI.

Academic Streams

Data from the pilot indicates stronger participation from Commerce and Science streams, with Commerce students forming a marginally larger share in several centres. This reflects existing enrolment trends in UG entrance preparation ecosystems.

Student Experience and Behavioural Insights

Encountering a fundamentally new assessment format was both a challenge and a learning opportunity for students. Accustomed to tests aligned with national-level entrance examinations, many students were initially surprised by the nature of the mySATHI test.

Student reactions broadly fell into three categories:

Curiosity and Excitement: A segment of students showed keen interest in understanding how the test measures aptitude and potential rather than rote knowledge.

Cautious Engagement: Some students attempted to approach the test strategically, treating it like a competitive exam despite repeated messaging that it was diagnostic in nature.

Apprehension and Anxiety: Time limits, unfamiliar question types, and fear of “performance judgement” contributed to anxiety among a subset of participants.

These behavioural patterns corroborate findings from the Excel data, which show variability in completion times and section-wise performance consistency—often linked to familiarity rather than ability.

Operational Challenges Observed

Operational execution across 27 cities surfaced several challenges:

(This was a Paper based pilot. The 4C test is planned to be a digital test and in DEXIT centres in the actual roll out))

Scheduling Conflicts: Aligning test dates with school calendars and coaching schedules proved difficult, directly impacting attendance in some centres.

Attendance Variability: Despite registrations, actual turnout varied due to academic workload and exam pressure.

Purpose Clarification: Persistent questions around “Why this test?” and “How does it help?” highlighted the need for stronger pre-test orientation for students and parents.

These challenges underline that even well-designed assessments require significant on-ground communication and institutional buy-in to achieve optimal participation and engagement.

Assumptions vs. Ground Reality

Key Assumptions at Rollout

- Students would treat the assessment as low-stakes and exploratory.
- Physical mode would reduce digital friction and anxiety.
- CL centres could seamlessly integrate the pilot into their schedules.

Observed Reality

- Many students defaulted to exam-oriented behaviour.
- Anxiety persisted despite the diagnostic positioning.
- Operational integration required more handholding than anticipated.
- This contrast between assumption and reality provided crucial learning inputs for future design and communication strategies.

Limitations and Reasonable Restrictions of the Pilot

- The findings of the pilot must be interpreted within its inherent limitations:
- Geographic skew towards private, CBSE, urban institutions.
- Limited representation from government and rural schools.
- Physical-mode testing, whereas the final product is digital and adaptive.
- Short pilot duration, restricting longitudinal outcome tracking.

These constraints were consciously accepted as part of a controlled pilot aimed at learning and validation rather than representation.

Implications for Nationwide Rollout in 2026

The pilot conclusively validates the conceptual strength of mySATHI while offering clear direction for scale:

- Robust **orientation and expectation-setting** will be critical.
- Digital adaptive delivery must be accompanied by readiness support.
- Regional customisation and school-type sensitivity are essential.
- Nationwide rollout should be phased, with continuous feedback loops.
- Use mySATHI assessment for UG/PG admissions, and entry level job recruitment.

In sum, the pilot establishes mySATHI as a credible, scalable assessment framework, as long as its rollout strategy incorporates the contextual and operational learnings surfaced during this study.

CHAPTER 14

mySATHI — The Computer Adaptive Testing (CAT) Ecosystem

Abstract

mySATHI is more than just an online assessment platform; it's a sophisticated measurement system that aims to improve the integrity, robustness, accuracy, and speed of high-stakes tests. It offers a fully “On-Demand” testing experience by integrating DEXIT’s strong operational infrastructure with its adaptive assessment engine. This means that results are always comparable, fair, and secure throughout different testing windows and groups of candidates. MySATHI assessment employs Computer Adaptive Testing (CAT) to constantly guess how good each candidate is and then choose questions that are just the correct level of difficulty for them. This lets institutions and employers get rid of one-day, high-pressure exams and fixed paper sets and move toward a continuous, scalable, defensible assessment paradigm that is hard to cheat on, memorize by heart, or leak questions.

The Testing Architecture: A Journey Through Four Phases

The mySATHI Computer Adaptive Test (CAT) is a multi-stage process that gradually improves the accuracy of the candidate’s skill level. Each phase has a different strategic goal, making sure the exam is engaging, psychometrically solid, and works well.

Phase 01—Warmup

- The first questions are meant to help applicants get used to the exam environment, lower their stress levels, and give the adaptive engine some baseline response data.
- This early data gives the system a rough idea of each person’s skill level so that it doesn’t start everyone at the same level of difficulty and can swiftly move on to items that are just right for them.

Phase 02—The Main Part

- The core exam is split into “nodes” (which are groups of up to five questions). Each node is linked to a certain level of difficulty and a certain set of topics.

- As candidates answer questions, the engine changes the complexity of the next nodes in real time, pushing them up or down to find the “adequate” level of difficulty where questions are not too easy or too hard.

Phase 03: Carry Over

- If they have time, applicants can go back to questions they didn’t answer or marked questions, or they can work on specific items that assist them get a better idea of their ultimate ability.
- This step makes the test more fair and accurate by letting the engine add more data points without making the test last longer.

Phase 04—Challenge Round

- Only applicants who meet certain performance standards in the main section can answer high-difficulty questions – a stage that identifies top notch candidates with fewer questions – useful for providing scholarships, advanced programs or competitive job roles.

Deep Dive: The “Main Part” That Changes

The adaptive engine accomplishes much of its work in the “Main Part” by evaluating ability and making sure that there is a lot of coverage over the whole syllabus.

- Node-based design: Each node (of 5 questions) are labeled by topic, skill, and level of difficulty. This lets the engine choose the next node depending on both performance and the needs of the blueprint.
- Dynamic routing: The system adjusts the candidate’s estimated ability after each node and sends them to an easier, similar, or harder node, looking for a balance where the questions are just right.

Smart Progression

- A candidate passes on to the next topic or skill cluster when they meet particular conditions, such as getting a certain number of correct answers, reaching a certain number of attempts, or being close to the time limit.
- This makes sure that the test doesn’t focus too much on a small number of items and instead generates a full profile across several areas while still being flexible.

Complete Coverage with Personalization

Content mapping makes sure that all important topics are covered on the exam form, even though each candidate sees a different collection and order of questions. Hence, each candidate takes a test that is just right for him/her; institutions nevertheless get scores that are comparable and based on the same ability scale.

Why mySATHI? What Makes It Stand Out?

The mySATHI CAT engine seeks to stand apart from others in the assessment business, by focusing on psychometric rigor, operational security, and actionable analytics.

Real adaptability, not randomization

- The choice of questions is based on a real-time estimation of ability that takes into account accuracy, response speed, and new signs of proficiency, rather than just shuffling fixed papers.
- This means shorter exams that nevertheless give very accurate results, which cuts down on candidate fatigue while still giving the decision-maker a lot of authority.

Depth, variety, and validity of items

- Each test item has many parallel variants (based on the same skill and difficulty level) but use various numbers, contexts, and distractors.
- So memorization is hard, permits testing in multiple windows at once, and keeps the test focused on understanding concepts instead of remembering them.

The layer of trust and defensibility

- Operational controls like device and IP logging, candidate identification verification, and extensive audit trails make sure that each score can be justified if there is a dispute.
- Policies for item rotation and exposure control prevents big leak or organized wrongdoing.

Useful information and detailed reports

- Reporting extends beyond just a single total score to show competence profiles by skill, topic, and level of difficulty. This lets you focus your remedial efforts and make smarter decisions about who to admit or hire.
- Dashboards let institutions look more closely at certain groups, centers, and time periods, which makes problems clear and helps them keep getting better.

The mySATHI assessment ecosystem is built on a combination of infrastructure and intelligence, bringing together the best parts of DEXIT and mySATHI's adaptive engine that can control dynamic difficulty, routing and ability estimate. It makes sure that each test is based on psychometric data and is adapted to every candidate; the engine has a big, verified bank of objectives with mapped skills and calibrated difficulty levels – all of which make CAT work. The full-stack solution makes sure that the operations are safe and does ID checks, uses item variants, exposure limitations and rotation algorithms to prevent cheating

- Adaptive difficulty lets you tell the difference between people with higher abilities more clearly, which is important for competitive programs with few seats.

- Hiring a lot of people for their first jobs
- Employers can keep hiring people without having to make fresh paper sets every time by using mySATHI as a high-throughput, always-on filter.

Anti-cheating methods and item variation lower the possibility of cooperation and assist keep reliability high across locations and batches.

- High-stakes hiring for postgraduate and lateral positions
- You may evaluate specialized tasks with fewer but smarter questions that focus on advanced talents. This saves time for both candidates and hiring managers.
- Reporting on skill levels in detail helps with data-driven shortlisting, preparing interviews, and training new hires when they start working.

How AI Helps Make The Questions:

The mySATHI's AI-driven question production pipeline is meant to be organized, in line with the curriculum, and quality-controlled, making sure the adaptive engine always has access to a rich, new, and reliable item bank.

Step 1: Begin with the curriculum and the goals

- The process starts with the client's grade or level, chapter list, and objective learning outcomes. This makes sure that the client's goals are in line with the institution's syllabi or competency frameworks.
- These results are broken into smaller "skills," like "add fractions" or "area word problems." These skills are the building blocks for making and analyzing questions.

Step 2: Use AI to make groups of items

For every ability that AI finds, it makes:

- various versions of the same question that test the same idea but with various numbers, situations, or words.
- A right response with a step-by-step solution that lets you give explanations, answer keys, and reasons for wrong answers.
- A range of difficulty levels (Easy → Medium → Hard) so that CAT can choose objects that are right for each candidate's present ability level.

Quality Checks Built-In Before Human Review

The automated quality checks seek to get rid of any questions that are wrong or not aligned before any human expert looks at them.

- **Check for the right answer:** Makes sure that the solution matches the final answer and that everything is consistent inside the solution.

- **No-ambiguity check:** For questions with more than one answer, it makes sure that there is only one correct answer and that the wrong answers are believable but plainly wrong.
- **Skill match check:** Makes sure that the question really tests the skill it was meant to test and doesn't accidentally test something different.
- **Language check:** Looks for language that is appropriate for the age group, is clear, and doesn't use confusing or culturally inappropriate phrases.

This automated pre-filtering lets human specialists focus on improving and validating instead of just finding basic errors, which makes the process faster and more consistent.

Human Validation and Calibration Based on Data

- The validation phase adds human judgment and real-world facts to the material that AI created, making the connection between generation and real-world performance.
- Review by a subject expert
- Experts in the subject check each item to make sure it is conceptually correct, fits with the syllabus, and has the right processes for solving the problem.
- Experts can accept, change, or reject things, making sure that only content that fulfills academic and institutional criteria goes into the live bank.

Pilot data and psychometric feedback (optional)

- When possible, products are tested with a small group to provide information on how hard they are and how well they work.
- Items that do well are kept and given empirical parameters, which feed directly into more accurate adaptive algorithms.

Making Levels Easy, Medium, and Hard

mySATHI's difficulty levels are not allotted arbitrarily, but planned carefully:

- **Easy:** Questions that are direct and only require one step, using numbers or situations that are easy to understand, check basic memory or habitual use.
- **Medium:** Two-step issues that demand some thought, such putting together ideas or going from a verbal problem to a numerical statement.
- **Hard:** Questions that need more than one step or a lot of reasoning, have more complicated situations, hidden traps, or hard distractors that test a deeper comprehension.

A.I suggests questions at all three levels, and professional reviewers check or change these labels ensuring that difficulty tags are based on both theoretical design and actual behaviour.

Use mySATHI to Make Assessments More Powerful

With its AI-powered CAT engine and content pipeline, mySATHI is THE platform that is more valuable for schools, businesses, and applicants, because it converts a generic test delivery system to a decision-grade measurement platform where every score is underpinned by adaptive logic, validated content, and data that is ready for an audit. Institutions can execute frequent, safe tests without sacrificing fairness, comparability, or test quality thanks to on-demand scheduling, strong operations (DEXIT), and adaptive intelligence.

In real life, this signifies that:

Admissions and hiring choices can be made on exact, skill-level information instead of merely total scores on preset exam sets; without coaching system or fear of questions leaking. Stakeholders obtain more detailed data to plan bridging courses, remedial actions, and focused training, which closes the gap between student performance and objectives.

In fact, mySATHI offers a future-ready assessment infrastructure that can grow with the demands of universities, test owners, and employers by intimately linking AI-powered content development, strict validation, and computer adaptive delivery inside an operationally safe ecosystem.

PART IV

CHAPTER 15

A Summary of the Pilot Test Results

This section deals with the most important part of the Research work – namely to look at the results of the assessment and a deeply insightful analysis of the data obtained from the assessments. Initially, the SATHI pilot test report analyzes responses from 3,257 students on 100 questions, assessing skills in analytical reasoning (ALR), verbal ability (VARC), quantitative aptitude (QADI), and situational awareness (SIA). It breaks down performance by difficulty levels (easy, medium, hard), domains, and cognitive categories (C1-C4), revealing clear patterns in student strengths and weaknesses. Advanced clustering shows students fall into high- and low-performing groups based on overall accuracy, with test types influencing results.

Key Performance Trends

Overall accuracy drops predictably with difficulty: 53.73% on easy questions, 48.32% on medium, and 43.60% on hard, showing questions get tougher as intended. SIA leads all domains at 61.66% accuracy, while ALR lags at 39.42%, highlighting students' stronger grasp of real-world scenarios over logical puzzles. Category C4 tops at 61.66%, but C1 struggles most at 40.84%, indicating some thinking skills challenge students more uniformly.

Domain and Category Breakdown

Domain/Category	Overall Accuracy	Notes
SIA	61.66%	Strongest across easy (68.8%), medium (60.3%), hard (59.6%); high variability shows mixed question difficulty
QADI	49.23%	Good on easy/medium (57%/52.1%), drops on hard (38.1%).
VARC	45.62%	Medium questions best (50.7%), unusual dip on easy ones.
ALR	39.42%	Weakest; sharpest drop on medium (33.4%)

Communication	61.66%	Best category, consistent strength even on hard items
Creativity / Collaboration	47.29%/45.62%	Mid-range, balanced but below top performers.
Critical Thinking	40.84%	Toughest, steadily declines to 34.3% on hard.

Test Variations

Test (2,291 students) shows solid progression: 57.64% easy to 46.24% hard. Re-test (622) and split-test (344) mirror this closely in patterns, though split-test accuracies are oddly low (e.g., 19-25%), possibly due to alternate questions. Across six datasets (468 to 3,257 students), SIA stays strongest and ALR weakest, proving test reliability.

The student grouping insights from the report use t-SNE visualizations and K-means clustering to reveal consistent patterns in how 3,257 students approach the SATHI test. Test type (Test, Re-test, Split-test) creates the main divide, with sub-groups emerging based on performance style—high attempters with better accuracy versus cautious low attempters. These patterns hold steady across datasets up to 3,212 students, showing reliable behavioral archetypes rather than random variation.

t-SNE Visualization Details

t-SNE maps student behaviors like attempt rate, blanks, accuracy by category/section/difficulty into 2D space. Test types form distinct clusters: Split-test is tightest due to unique questions, while Test and Re-test spread more. Inside each, two sub-clusters appear—higher performers (many attempts, high accuracy) separate from lower ones (fewer attempts, more errors)—with topic labels (C1-C4, domains) mixing inside, not driving splits.

Using just mean correctness (accuracy on attempted questions), K-means (k=2) splits students cleanly into high- and low-performers. Profiles across categories/sections/difficulties are parallel: high group always ~10-20% better, no selective strengths (e.g., high group leads in SIA 61-69% vs. low's 40-50%). Dataset 4 (2,953 students) and 5 (3,212) match perfectly, confirming scale doesn't change the two-group structure.

Practical Interpretation

High-performers tackle more questions confidently, boosting scores; low-performers skip or guess poorly, dragging results. Content metadata weakly predicts clusters—behavior trumps topic familiarity—suggesting tests measure engagement as much as knowledge. For educators, this flags need for strategies boosting attempt rates in weak domains like ALR.

CHAPTER 16

Comprehensive Analysis Report of SATHI Pilot Test

Introduction

This report presents a comprehensive psychometric analysis of the SATHI pilot test dataset comprising 3257 student responses across 100 items. Also, it presents a comparison of all the datasets shared so far. The dataset includes metadata rows and candidate response data:

Table I: Structure and Description of the SATHI Pilot Test Dataset

Row	Description	Example
Row 0	4C Category (C1, C2, C3, C4)	Cognitive or concept classification per question
Row 1	Section Name	ALR (Analytical Reasoning), VARC (Verbal Ability & Reading Comprehension), QADI (Quantitative Aptitude & Data Interpretation), SIA (Situational Awareness)
Row 2	Difficulty Level (1–3)	1 = Easy, 2 = Medium, 3 = Hard
Row 3	Answer Key	Correct option per question (A/B/C/D)
Row 4	Column Labels for Student Data	Enrolment No., Correct, Incorrect, Incomplete, etc.
Rows 5–3264	Candidate Responses	Each student's answers for Q1–Q100

I. Analysis on the Current Dataset

A. Accuracy by Difficulty Level

Figure 1 presents the aggregated accuracy across three difficulty tiers.

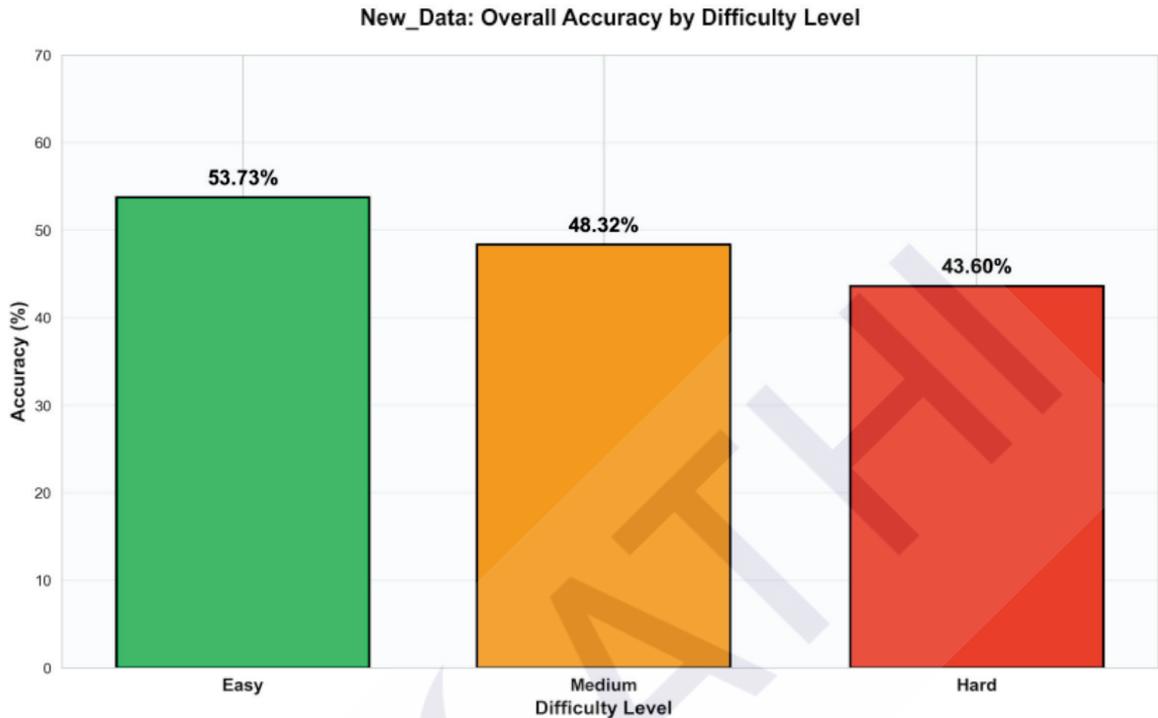


Figure 1: Overall Accuracy by Difficulty Level

- **Easy** questions achieve the highest score at **53.73%**, showing that the students handles simpler items reasonably well.
- **Medium** questions drop to **48.32%** accuracy, reflecting increased cognitive or interpretive demands.
- **Hard** questions result in the lowest accuracy, **43.60%**, marking a clear decline in performance with item complexity.

B. Accuracy by Domain

Figure 2 reports the accuracy across four domains after aggregating all three evaluation settings: Test, Re-test, and Split-test. The results reveal substantial variation in domain-wise performance.

- **ALR** (Analytical Logical Reasoning) yields the lowest accuracy at **39.42%** with a standard deviation of $\pm 13.14\%$, indicating persistent difficulty for students in legality- and regulation-oriented tasks, along with relatively lower variability compared to other domains.
- **QADI** (Quantitative Ability) achieves an accuracy of **49.23%** with a higher standard deviation of $\pm 20.01\%$, suggesting moderate overall performance but substantial variability in student responses across ques-tions.

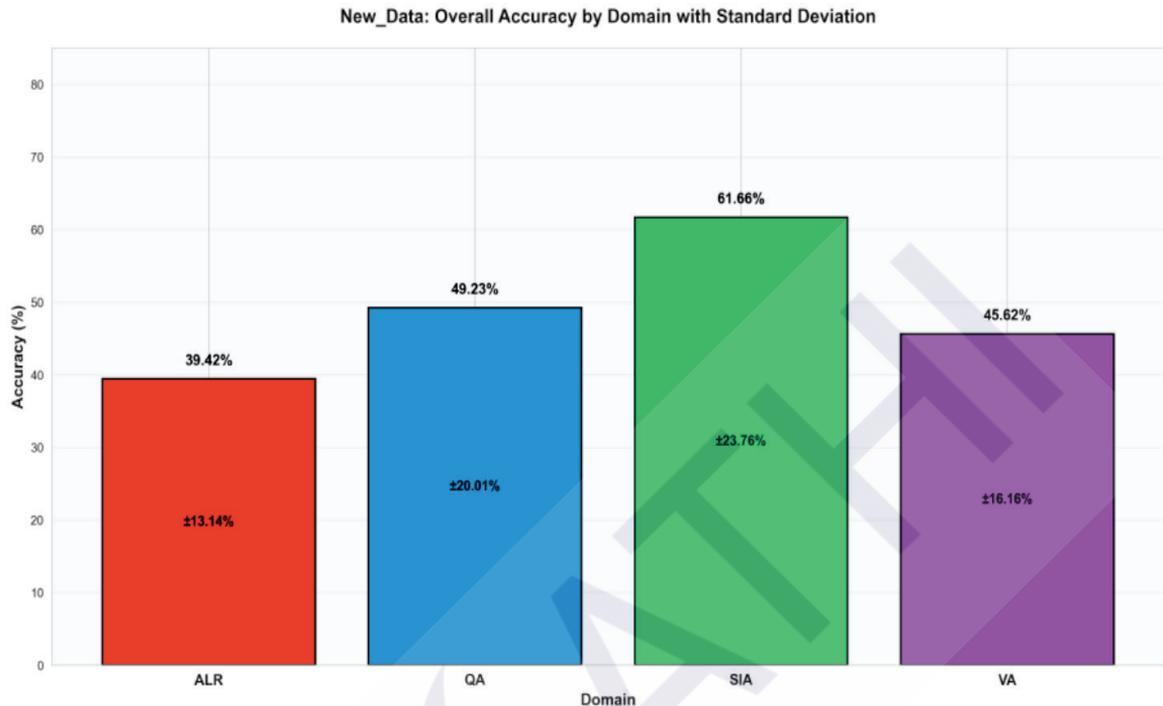


Figure 2: Accuracy by Domain

- **SIA** (Situational Judgment & Information Analysis) stands out prominently with the highest accuracy of **61.66%**, accompanied by the largest standard deviation of **±23.76%**. This indicates strong average performance, but also highlights considerable dispersion, likely reflecting a wide range of question difficulty within the domain.
- **VARC** (Verbal Ability & Reading Comprehension) attains an accuracy of **45.62%** with a standard deviation of **±16.16%**, performing better than ALR while exhibiting more stable performance than QADI and SIA.

(1) *Difficulty-wise Performance Analysis Across Domains:* Figure 3 presents the difficulty-wise accuracy of the *Final_Saathi* dataset across four domains: ALR, QADI, SIA, and VARC. The analysis reveals substantial variation in performance trends across domains and difficulty levels, highlighting both strengths and calibration issues within the dataset.

- ALR Domain:** The ALR domain is the weakest-performing domain overall. While Easy questions achieve a moderate accuracy of 48.70%, performance drops sharply for Medium questions to 33.40%, marking the lowest accuracy across all domains and difficulty levels. Hard questions recover slightly to 39.7%.
- QADI Domain:** The QADI domain exhibits relatively strong performance on Easy (57.0%) and Medium (52.1%) questions. However, accuracy drops significantly for Hard questions to 38.10%, indicating a clear increase in cognitive complexity at higher difficulty levels.



Figure 3: Difficulty-wise Performance Analysis Across Domains

- (c) **SIA Domain:** The SIA domain consistently outperforms all other domains across all difficulty levels, making it the strongest domain in the dataset. Accuracy remains high for Easy (68.80%), Medium (60.30%), and Hard (59.60%) questions.
- (d) **VARC Domain:** In the VARC domain, Medium questions achieve the highest accuracy (50.7%), outperforming Easy questions (44.30%). Hard questions show a moderate decline to 41.20%.

Overall, this analysis highlights SIA as a domain of strong and stable performance, while ALR emerges as the most challenging and least consistent domain, particularly at the Medium difficulty level.

C. Overall Category-wise Performance and Variability

Figure 4 presents the overall accuracy of the *Final_Saathi* dataset across four categories (C1–C4), along with the corresponding standard deviation, capturing both performance and variability within each category.

- (a) **C4 Category (Best Performing):** C4 achieves the highest overall accuracy at 61.66%, clearly outperforming all other categories. However, it also exhibits the largest variability, with a standard deviation of $\pm 23.76\%$. This indicates that while C4 contains many well-answered questions, performance varies substantially across instances, suggesting a mix of very easy and very challenging questions within this category.
- (b) **C2 and C3 Categories (Mid-range Performance):** C2 attains an overall accuracy of 47.29% with a standard deviation of $\pm 16.92\%$, while C3 follows closely with 45.62% accuracy and a standard deviation of $\pm 16.16\%$. Both categories show moderate performance and relatively comparable variability, indicating more balanced difficulty distributions compared to C4.

Among the two, C2 performs slightly better on average, though both remain significantly below C4.

- (c) *C1 Category (Weakest Performing)*: C1 records the lowest overall accuracy at 40.84%, making it the most challenging category for students. Its standard deviation of $\pm 15.07\%$ is the smallest among all categories, suggesting consistently low performance rather than the presence of extreme outliers. This consistency indicates that questions in C1 are uniformly difficult rather than unevenly distributed in difficulty.

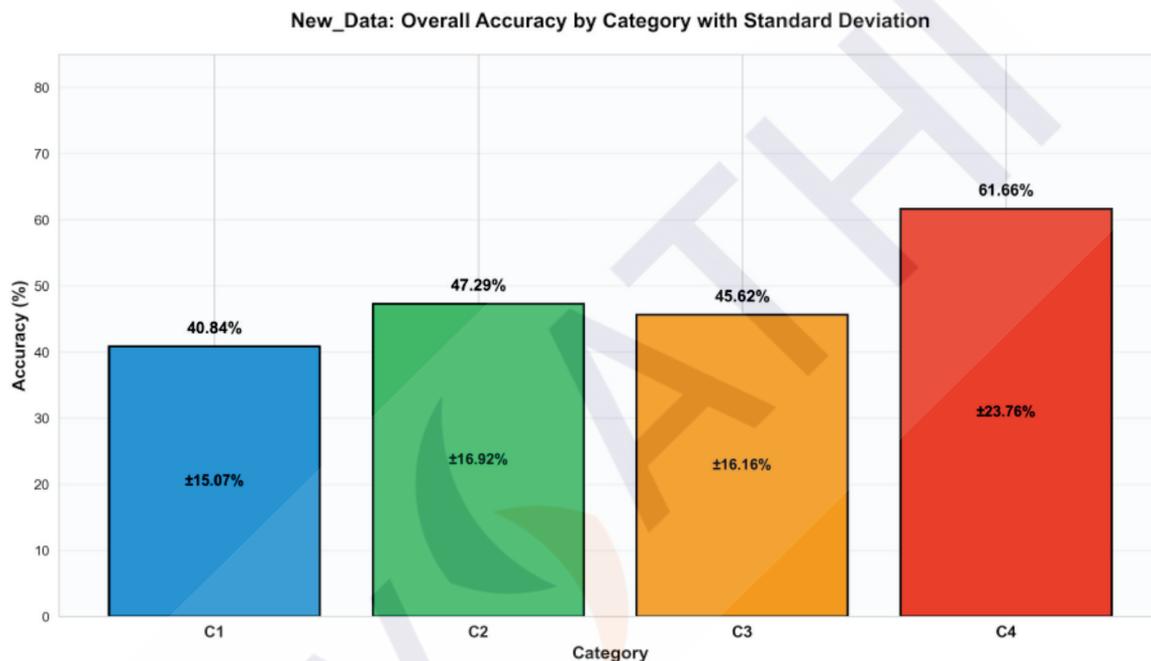


Figure 4: Overall Category-wise Performance and Variability

(1) *Difficulty-wise Performance Across Categories*: Figure 5 illustrates the accuracy of the *Final Saathi* dataset across four categories (C1–C4), stratified by difficulty levels (Easy, Medium, Hard). The results reveal distinct performance patterns and highlight category-specific difficulty behaviors.

- (a) *C4 Category (Consistently Strong Performance)*: C4 demonstrates the strongest and most consistent performance across all difficulty levels, achieving 69.8% accuracy on Easy questions, 60.3% on Medium, and 59.6% on Hard questions. The relatively small drop from Easy to Hard indicates well-calibrated difficulty and robust understanding within this category.
- (b) *C1 Category (Most Challenging)*: C1 emerges as the weakest category overall, with accuracy steadily declining from Easy (49.8%) to Medium (40.2%) and reaching its lowest point on Hard questions (34.3%). This monotonic decrease aligns with expected difficulty progression but also highlights significant challenges faced by students in this category.
- (c) *C2 Category (Moderate and Stable Performance)*: C2 exhibits moderate performance with a gradual decline across difficulty levels: 55.6% (Easy), 46.5% (Medium), and 42.86% (Hard).

(Hard). Compared to C1, the performance drop is less severe, suggesting a more balanced distribution of question difficulty.

- (d) *C3 Category (Irregular Difficulty Pattern):* C3 shows an unusual trend where Medium questions (50.7%) outperform Easy questions (44.3%). Hard questions record a lower accuracy of 41.2%. This inverted pattern suggests potential ambiguity or misclassification of Easy questions within C3, warranting further inspection of difficulty labeling.
- (e) *Cross-Category Observations:* Across categories, Easy questions are most effectively answered in C4 (68.8%) but prove relatively difficult in C3 (44.3%). While C1 and C2 largely follow the expected Easy → Medium → Hard progression, C3 deviates from this pattern, highlighting inconsistencies in difficulty calibration. Overall, these findings emphasize the need for category-specific difficulty validation to ensure consistent assessment standards.



Figure 5: Difficulty-wise Performance Analysis Across Categories

II. Analysis of Category Wise Test, Re-Test & Split-Test

A. Test Category (2,291)

(1) *Accuracy by Difficulty Level:* Figure 6 reports the accuracy across three difficulty levels for the Test category. The results show a clear and well-calibrated difficulty progression.

- **Easy:** The highest accuracy, **57.64%**. Students demonstrate strong foundational knowledge, with nearly half of the responses correct.



Figure 6: Accuracy by Difficulty

- **Medium:** Accuracy drops to **50.70%**, a decline of 6.44 percentage points from Easy items. This indicates moderate difficulty in intermediate concepts.
- **Hard:** The lowest accuracy at **46.24%**, suggesting significant difficulty with advanced or conceptually heavy items.

(2) *Accuracy by Domain:* Figure 7 presents the domain-wise accuracy distribution for the Test dataset.

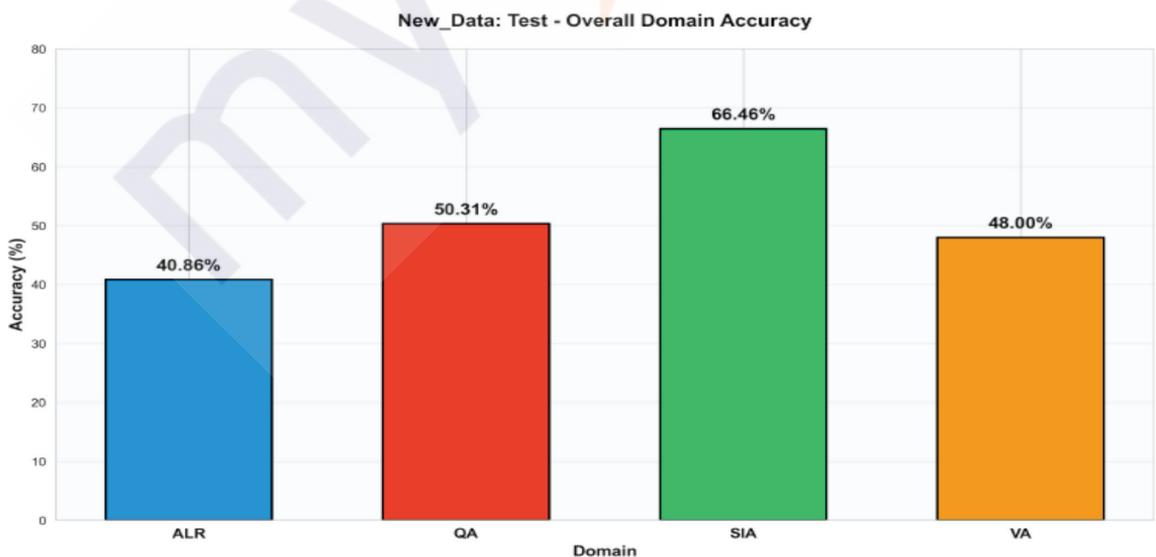


Figure 7: Accuracy by Domain

- **SIA** achieves the highest accuracy, **66.46%**.
- **VARC** records **48.00%**, near the overall Test accuracy.
- **QADI** achieves **50.31%**, indicating moderate proficiency.
- **ALR** exhibits the lowest accuracy at **40.86%**, highlighting substantial conceptual challenges.

(3) *Accuracy by Category*: C4 achieves the highest performance at 66.46%, followed by C2 (49.58%) and C3 (48.00%), while C1 records the lowest accuracy at 41.80%, Figure8.

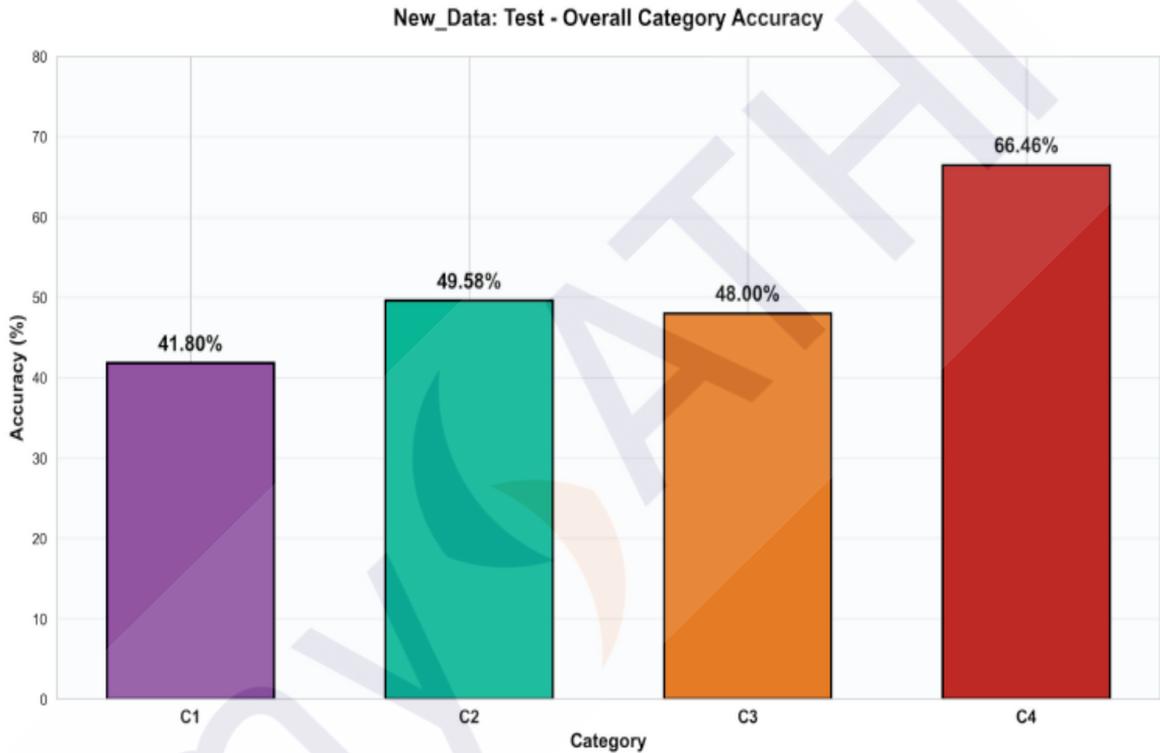


Figure 8: Accuracy by Category

B. Re-test Category (622)

(1) *Accuracy by Difficulty Level*: Figure 9 illustrates difficulty-level performance in the Re-test dataset. The pattern deviates sharply from the Test results.

- **Easy: 57.25%**
- **Medium: 51.55%,**
- **Hard: 47.41%**



Figure 9: Accuracy by Difficulty

(2) Accuracy by Domain: Figure 10 reports domain-wise performance in the Re-test dataset.

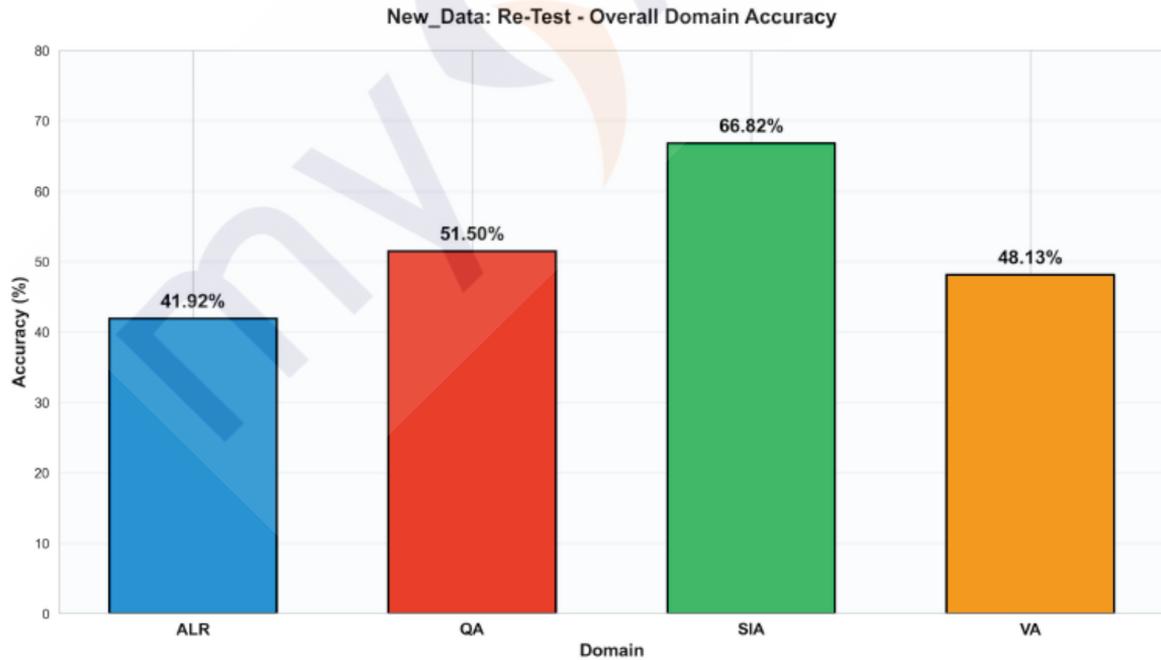


Figure 10: Accuracy by Domain

- **ALR: 41.92%**,
- **QADI: 51.50%**.
- **SIA: 66.82%**.
- **VARC: 48.92%**

(3) *Accuracy by Category*: The Re-test set (622 students) exhibits performance patterns closely aligned with the Test set, suggesting stability and consistency in student understanding Figure 11. Accuracy slightly improves across all categories, with C4 again leading at 66.82%, followed by C2 (49.83%), C3 (48.13%), and C1 (43.84%).



Figure 11: Accuracy by Category

C. Split-test Category (344)

(1) *Accuracy by Difficulty Level*: Figure 12 displays the difficulty-level accuracy for the Split-test dataset. Since Split-test uses the same alternative question pool as Re-test, the results are identical.

- **Easy: 21.39%**
- **Medium: 25.88%**
- **Hard: 19.52%**

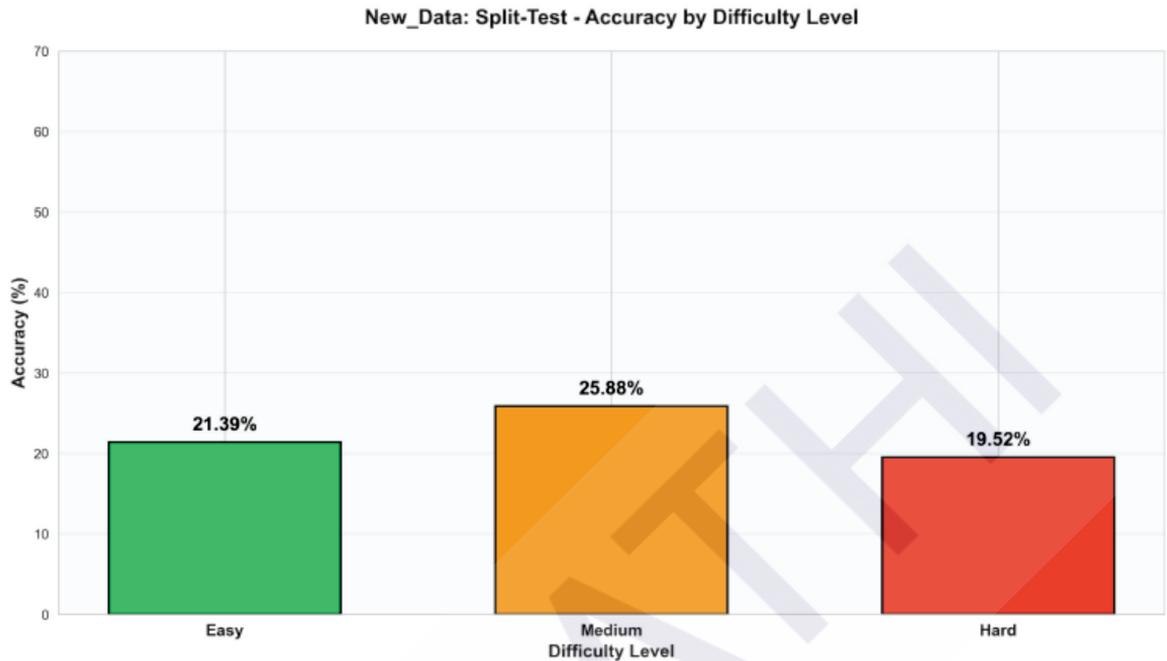


Figure 12: Accuracy by Difficulty

(2) *Accuracy by Domain*: Figure 13 shows domain-wise accuracy for the Split-test dataset. As expected, results mirror the Re-test data.

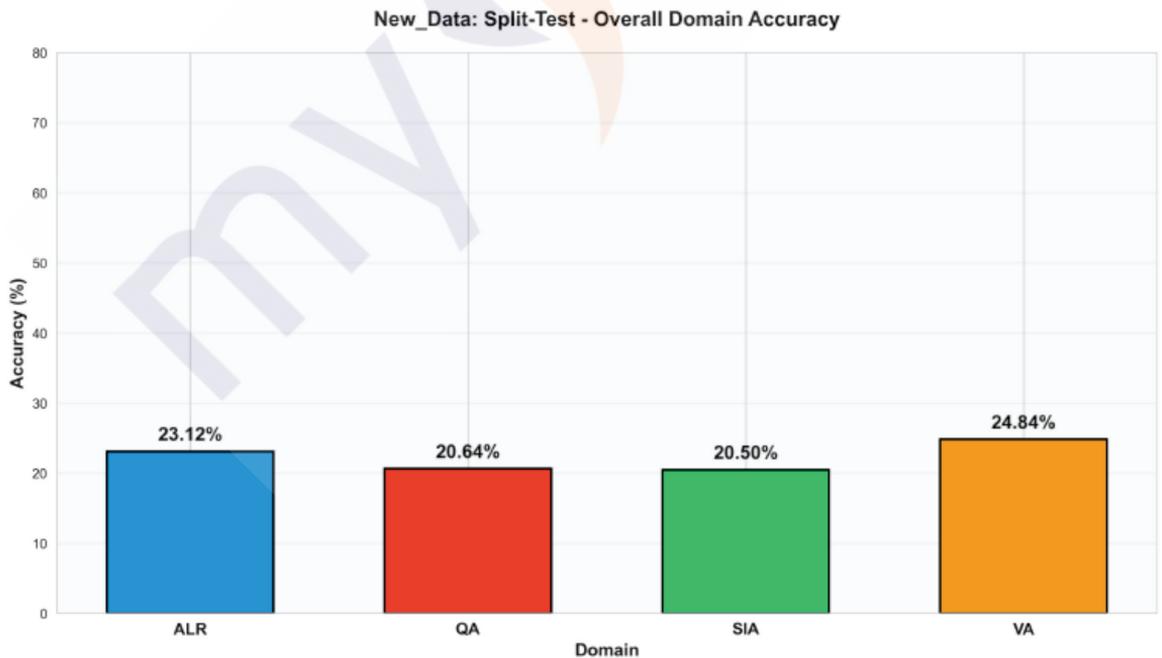


Figure 13: Accuracy by Domain

- ALR: 23.12%
- QADI: 20.64%
- SIA: 20.50%
- VARC: 24.84%

(3) *Accuracy by Category*: In contrast, the Split-test set (344 students) shows a substantial drop in accuracy across all categories Figure 14. C3 attains the highest accuracy at 24.84%, while C1 (23.07%), C4 (20.50%), and C2 (20.54%) perform comparably poorly.

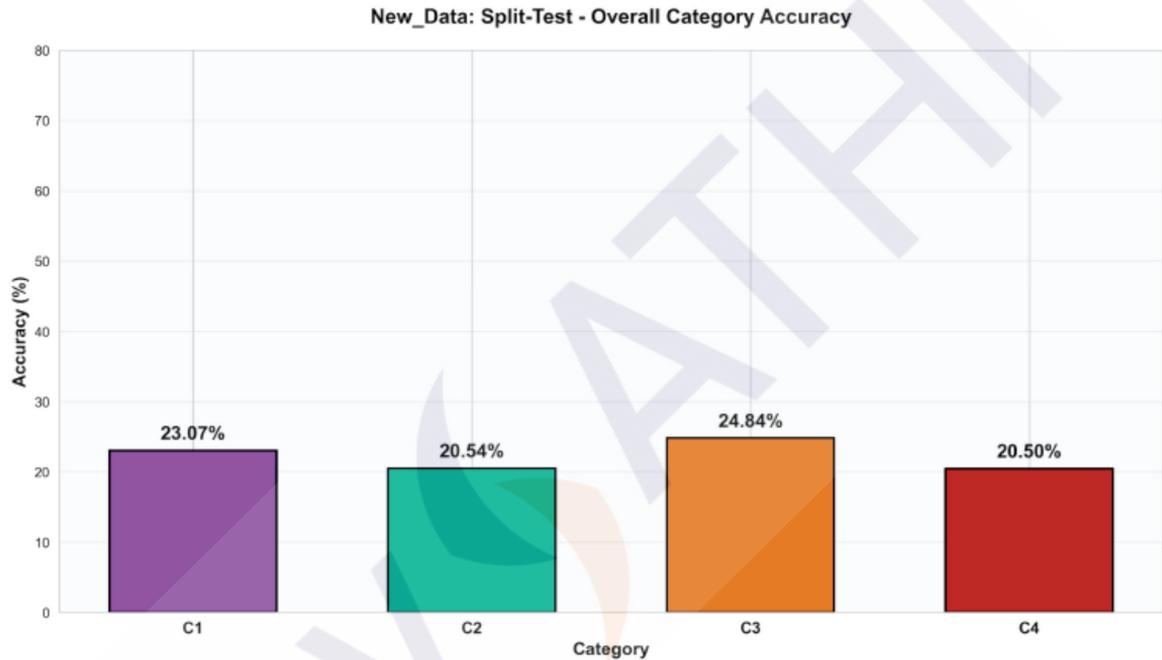


Figure 14: Accuracy by Category

III. Comparison of all the Datasets (Student Responses)

A. Cross-Dataset Comparison: Difficulty-Level and Domain-Wise Trends

Figures 15 and 16 present a holistic comparison of accuracy trends across all six datasets. The results reveal several consistent patterns in both difficulty progression and domain-wise performance.

(1) *Difficulty-Level Comparison Across All Datasets*: Across all six datasets, we observe a robust and well-calibrated difficulty progression from Easy → Medium → Hard. This monotonic decline holds consistently across all dataset versions, indicating strong internal validity and stability of the difficulty design across independently constructed question sets.

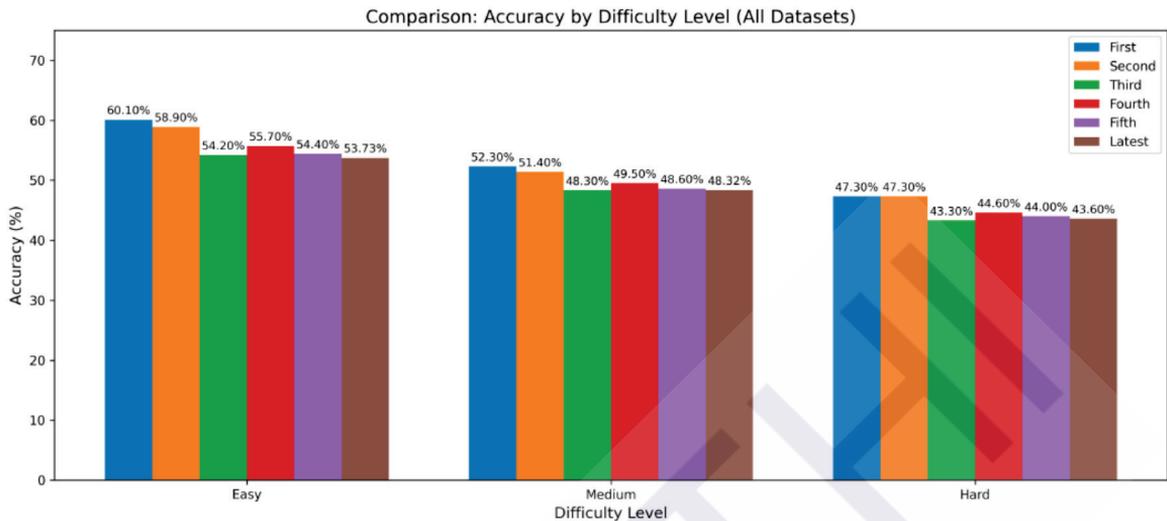


Figure 15: Difficulty-Level Comparison Across All Datasets. (First Dataset: 468), (Second Dataset: 1243), (Third Dataset: 2081), (Fourth Dataset: 2645), (Fifth Dataset: 2953), (Latest Dataset: 3257)

- **First and Second datasets** exhibit the steepest decline, with an approximate **12–13 percentage point** drop from Easy to Hard. This sharp gradient reflects a clear separation of difficulty levels and suggests strong discrimination between varying levels of student proficiency.
- **Third, Fourth, and Fifth datasets** demonstrate a similarly reliable decline, typically in the range of **10–11 percentage points**. Although absolute accuracy values are slightly lower compared to earlier datasets, the relative progression remains stable and pedagogically coherent.
- The **Latest dataset** closely mirrors the trend observed in the Fifth dataset, maintaining nearly identical Medium-to-Hard transitions. This consistency suggests that recent refinements preserve the underlying difficulty calibration rather than introducing distributional shifts.
- **Overall**, all datasets follow a consistent difficulty ordering, indicating that observed performance differences are driven by genuine variations in task complexity rather than artifacts of dataset construction or evaluation noise.

Students consistently achieve the highest accuracy on Easy items (mid-50% to low-60% range) and the lowest accuracy on Hard items (low-40% range), with Medium items occupying the expected intermediate position. The close alignment of performance curves across all dataset versions demonstrates that the difficulty model is both reliable and transferable across assessment iterations.

(2) Domain-Wise Comparison Across All Datasets: A clear and stable domain performance hierarchy emerges across all six datasets, demonstrating strong consistency in how different competency areas are assessed.

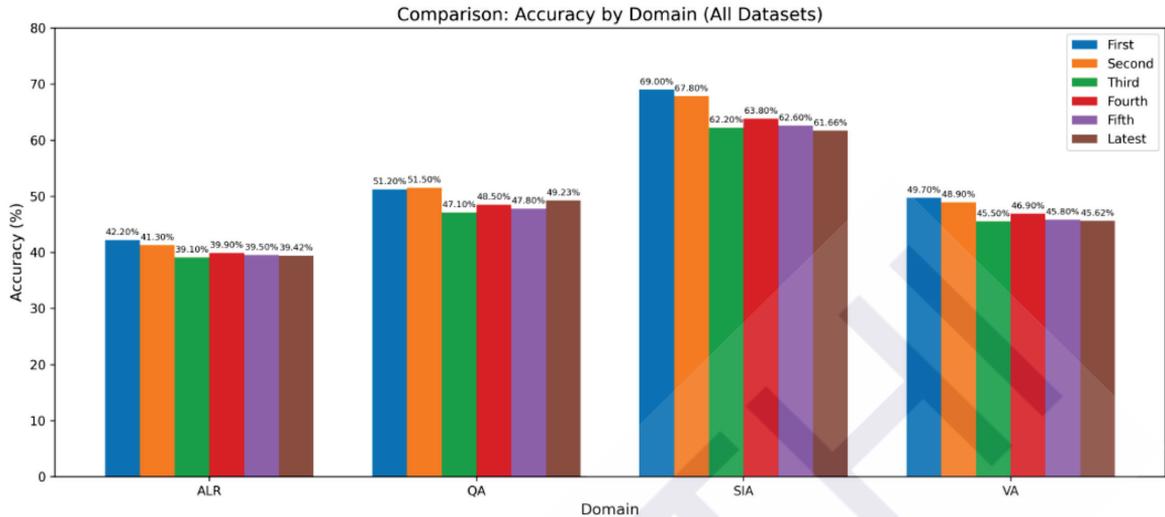


Figure 16: Domain-Wise Comparison Across All Datasets. (First Dataset: 468), (Second Dataset: 1243), (Third Dataset: 2081), (Fourth Dataset: 2645), (Fifth Dataset: 2953), (Latest Dataset: 3257)

- **SIA is consistently the strongest domain** across all dataset versions, achieving the highest accuracy in every case, with values ranging approximately from **61% to 69%**. This indicates robust student performance in scenario-based interpretation and applied reasoning tasks.
- **ALR remains the weakest domain**, with accuracies concentrated in the **39–42%** range across all datasets, including the Latest version. This persistent gap suggests that analytical and regulation-oriented reasoning continues to pose the greatest challenge.
- The mid-tier domains, **QADI** and **VARC**, exhibit stable and reproducible behavior. QADI consistently outperforms VARC by a small margin, with both domains maintaining similar relative positions across all dataset iterations.

The invariance of domain ordering across all six datasets highlights strong reproducibility in domain design and confirms that observed differences are driven by domain-specific cognitive demands rather than dataset construction artifacts.

(3) Differences Between Dataset Groups: Based on overall accuracy patterns, the datasets can be grouped into two broad performance clusters:

- **First and Second datasets** show higher overall performance, characterized by:
 - SIA accuracy near **68–69%**,
 - QADI accuracy near **51–52%**,
 - VARC accuracy near **49–50%**.

These results suggest strong question discrimination and potentially higher baseline proficiency in the earlier dataset versions.

- **Third, Fourth, Fifth, and Latest datasets** exhibit slightly lower but tightly clustered performance, with:

- SIA accuracy in the **61–64%** range,
- QADI accuracy near **47–49%**,
- VARC accuracy near **45–46%**.

The close alignment of accuracy curves across these four datasets indicates highly consistent difficulty calibration and comparable student response distributions.

(4) Overall Interpretation: Taken together, the domain-wise comparison demonstrates strong structural stability across dataset iterations. While absolute accuracy levels vary modestly, the persistence of domain rankings and performance gaps confirms that all dataset versions measure the same underlying competencies in a reliable and transferable manner. The Latest dataset, in particular, closely tracks the Fifth dataset, suggesting continuity rather than drift in domain difficulty design.

Conclusion: Across all assessments, the difficulty progression is coherent, domain ordering is stable, and dataset groups exhibit predictable accuracy patterns. These findings confirm the robustness of the assessment design and the reproducibility of student performance across multiple test forms.

IV. Summary of Key Observations for tSNE and Clustering Analysis (Datasets–1 to 5)

In our prior analysis across four datasets (ranging from 1,243 to 2,953 students), and now also in the fifth dataset (3,212 students), we observe a stable and highly interpretable behavioural structure as the scale increases:

- **Test Type dominates the global geometry of the behavioural embedding.** Whenever multiple assessment modes are present (*Test*, *Re-Test*, *Split-Test*), t-SNE consistently separates students primarily by Test Type.
- **Each Test-Type region reliably decomposes into (at least) two behavioural archetypes.** Within each Test Type, students separate into a higher-performing (high attempt / high accuracy) subgroup and a lower-performing (lower attempt / lower accuracy) subgroup.
- **Conceptual metadata (4C category, Section, Difficulty) exhibits weak alignment with the major clusters.** These labels remain mixed within clusters, indicating that behavioural style (how students attempt and respond) explains the dominant structure more than question-content groupings.
- **A simple 1D clustering on mean correctness yields a stable two-group split across all datasets.** K-Means with $k = 2$ on mean correctness consistently separates students into high- vs. low-performing groups, with near-parallel performance profiles across 4C, Sections, and Difficulty.

In what follows, we include Dataset–4 as a reference baseline (previously reported) and add Dataset–5 (new). Earlier datasets (1–3) show the same qualitative patterns and are omitted here for brevity.

V. T-SNE ANALYSIS

We perform a comparative t-SNE analysis using behavioural features (attempt rate, blank rate, accuracy, and accuracy-conditioned-on-attempt) aggregated across 4C category, Section, and Difficulty groups. To highlight the stability of the observed structure at larger scale, we reproduce Dataset-4 (reference) and present the new embedding for Dataset-5.

A. Dataset-4 (2953 samples) — Reference

Figure 17 shows the t-SNE projection for Dataset-4. As observed previously, **Test Type forms the primary axis of separation**: *Test*, *Re-Test*, and *Split-Test* occupy well-separated regions, with *Split-Test* typically forming the tightest and most distinct grouping due to its different question structure. Within each Test-Type region, the embedding further splits into **two consistent sub-clusters** corresponding to higher-performing vs. lower-performing behavioural archetypes. Importantly, 4C, Section, and Difficulty labels remain **intermixed** within these regions, reinforcing that the dominant structure is behavioural rather than content-driven.

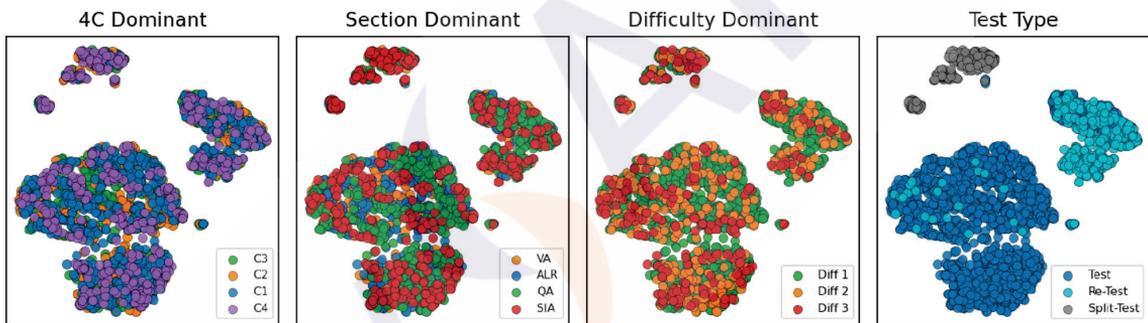


Figure 17: t-SNE Analysis across 4C, Section, Difficulty levels and Test Types for Dataset-4 (2953 samples).

B. Dataset-5 (3212 samples) — New

Figure 18 presents the new t-SNE embedding for Dataset-5 (3212 samples). We observe the **same structural pattern** as in Datasets 2-4:

- **Test Type remains the strongest separator**, producing distinct clusters for *Test*, *Re-Test*, and *Split-Test*.
- **Two sub-clusters persist within each Test Type**, reflecting higher-performing vs. lower-performing behavioural styles (high attempt/high accuracy vs. lower attempt/lower accuracy).
- **4C, Section, and Difficulty labels remain mixed** and do not form isolated regions, indicating weak alignment between conceptual metadata and the global embedding geometry.

Taken together, Datasets-4 and 5 confirm that the behavioural cluster geometry is **stable and scale-**

invariant: as the dataset grows beyond 3200 students, the same Test-Type-driven separation and within-type performance stratification remain clearly visible.

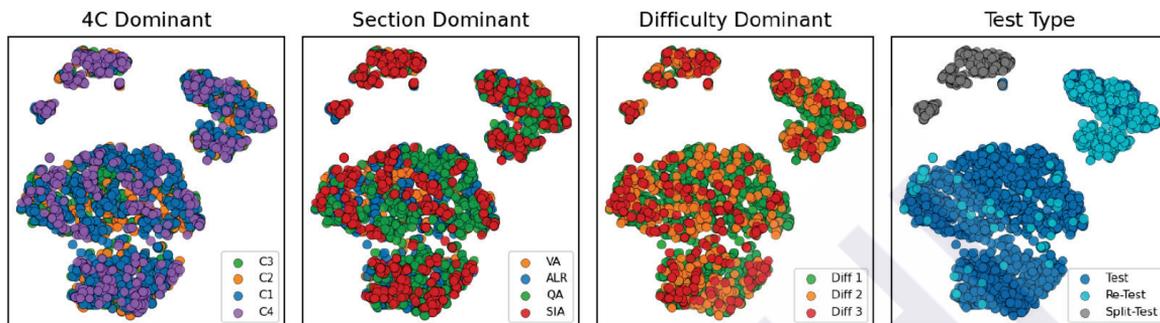


Figure 18: *t*-SNE Analysis across 4C, Section, Difficulty levels and Test Types for Dataset-5 (3212 samples).

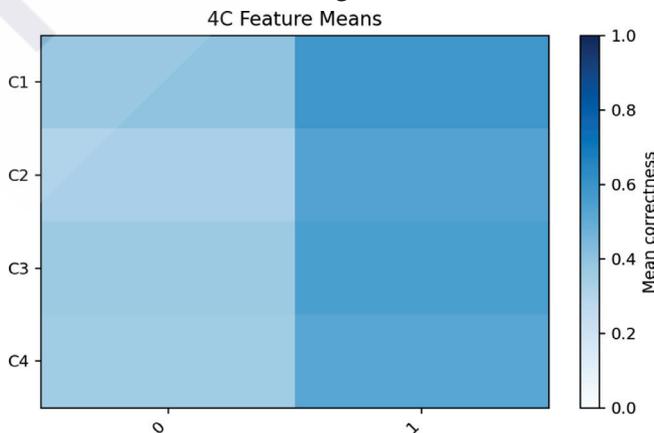
VI. Clustering Analysis

We next use a simple and interpretable clustering approach: each student is represented by a single scalar feature — the *mean correctness (accuracy)* across all answered questions. We apply K-Means with $k = 2$ (selected via elbow method in earlier analysis) to this one-dimensional feature space, and compute the mean correctness of each cluster separately across 4C categories, Sections, and Difficulty levels.

For consistency with the prior report and to emphasize stability at larger scale, we show Dataset-4 as reference and add Dataset-5.

A. Dataset-4 (2953 samples) — Reference

Figure 19 reproduces the cluster-wise mean correctness profiles for Dataset-4. As observed previously, clustering yields a **clean split** into higher- vs. lower-performing groups. The profiles across 4C, Section, and Difficulty are **nearly parallel**, indicating that separation is driven by global performance magnitude rather than selective strengths/weaknesses tied to content groups.



(a) 4C View

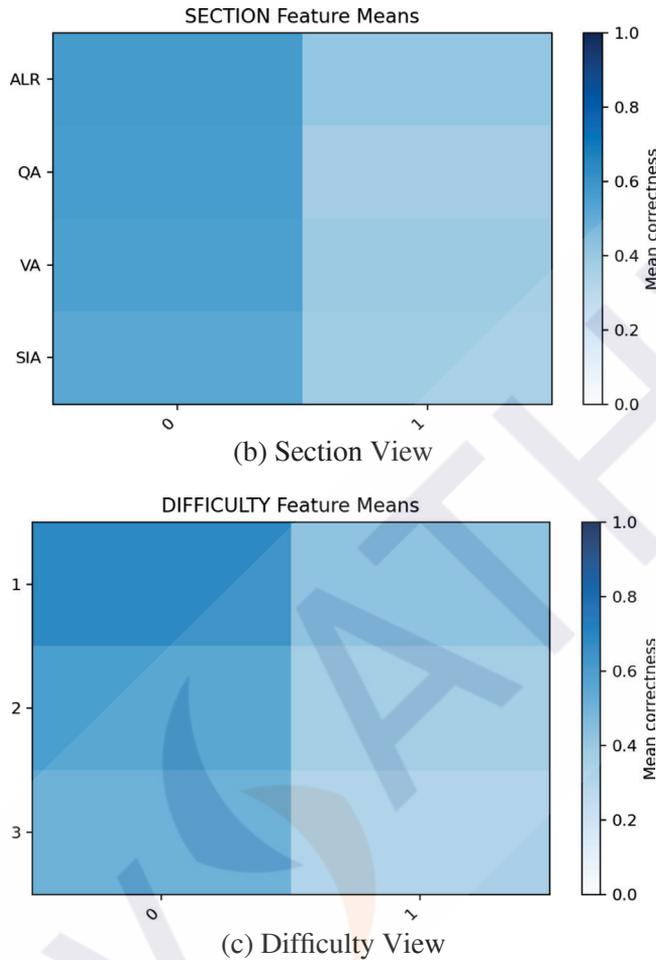


Figure 19: Cluster-wise mean correctness across 4C, Section, and Difficulty for Dataset-4 (2953 samples).

B. Dataset-5 (3212 samples) — New

Figure 20 shows the same clustering analysis for Dataset-5. We again obtain an **extremely consistent** two-cluster structure:

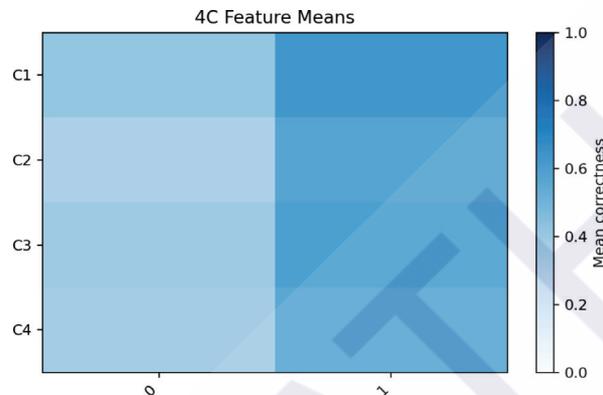
- A **higher-performing cluster** with consistently higher correctness across all groupings.
- A **lower-performing cluster** with uniformly lower correctness values.

Crucially, the **shape** of the 4C/Section/Difficulty profiles remains very similar between clusters (primarily a vertical shift), indicating that the separation continues to reflect a global performance difference rather than content-specific specialization. This confirms that the correctness-based clustering solution remains **stable** as the dataset grows to 3212 students.

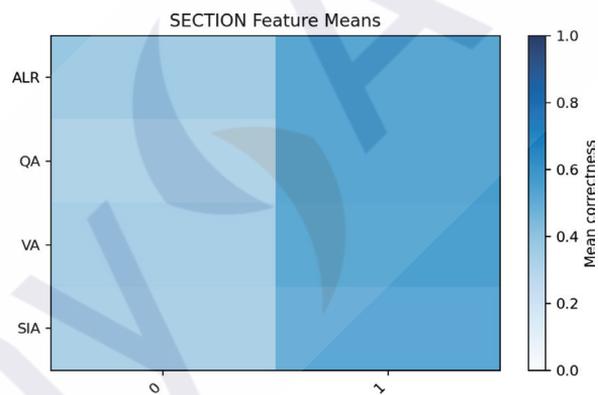
C. Interpretation (Datasets-4 and 5)

Across all prior datasets including Dataset-4 (2953) and Dataset-5 (3212), the clustering results reinforce the same conclusion:

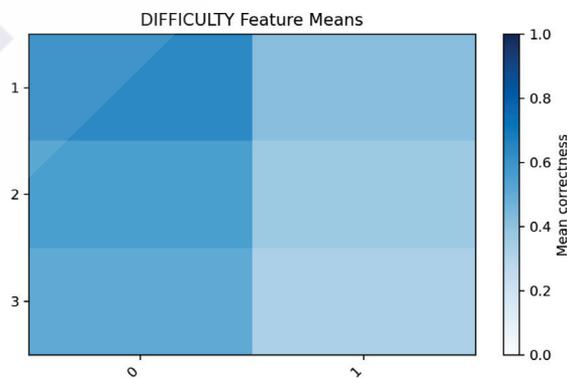
- The student population consistently decomposes into **two stable performance groups** (high vs. low correctness).



(a) 4C View



(b) Section View



(c) Difficulty View

Figure 20: Cluster-wise mean correctness across 4C, Section, and Difficulty for Dataset-5 (3212 samples).

- The cluster profiles across 4C, Sections, and Difficulty are **highly similar in shape**, implying that **content groupings do not drive the split**.
- The observed structure is **robust to scale**, persisting as the dataset expands beyond 3200 students.

Overall, both the embedding-based (t-SNE) and scalar-performance-based (K-Means) analyses provide converging evidence of persistent behavioural archetypes: Test Type explains the dominant separation when present, and within each region a stable high- vs. low-performing stratification is consistently observed.



PART V

CHAPTER 17

The Talent Observatory & Longitudinal Intelligence

The mySATHI Longitudinal Talent Platform

Will a unified data-fuelled platform help the Indian education system to align students, their aspirations, their skills and capabilities and help them choose their career paths” And would it support higher education institutions to identify the right fit students, via a robust assessment engine, and align such students with specific programs? And would this powerful compact help ensure long term success of such identified students?

The thought of building this engine is not idealistic anymore now. Today we have the robust computing, and technology that can combine with a powerful multi-dimensional assessment system to achieve this objective. And why do we need this?

Today, the Indian education system seems to operating in a difficult data vacuum: We have a lot of school enrollment numbers and university admission scores, but they are all separate siloes pillars of information. We don't have the tools to keep track of how people's potential, changes from when they are in Class VIII to when they become Boardroom leaders.

This is what mySATHI seeks to build: a powerful data-led ecosystem that can offer time accurate, deeply analytics filled dashboard to individuals as well as for institutions. At the same time, the assessment should also become a predictor of long term success. And this is precisely why the assessment engine tests the 4Cs – critical thinking, Creativity, Collaboration and Communication skills of the students in a wholesome manner – skills that builds resilience for success on a long term basis. Starting with:

- Assessment for entering University
- Assessment for entering the world of work, in dynamic different roles
- Assessment for long term success in jobs and/ or blooming into entrepreneurs (where the skills needs are radically different or of higher order)

Hence, the radical idea: A “Longitudinal Talent Observatory” (LTO) that aims to close this gap by changing how we think about education from static cross-sectional “snapshots” to

a continuous “movie” of how students learn. The framework proposed by mysATHI, is a key part of this ambitious idea. It can be termed as a unified “Dynamic Passport of Potential,” placed on mySATHI platform for providing the standardized, long-term intelligence needed to connect early aptitude, readiness for higher education, and long-term employability. This would complete the digital public infrastructure trilogy of Identity, Payments, and Learning.

The Crisis of Unrelated Events

Current educational data is fragmented. A student’s 10th grade performance is a separate issue that has nothing to do with how well they do in college engineering classes or how well they do in the job market. The “Chapter on Longitudinal Talent Observatory” calls this a “black box” event. Policymakers are currently unable to address essential inquiries: Is there a correlation between high scores in rote-learning examinations and workplace innovation? Do middle school interventions really help keep students in college? Reforms are still dependent on gut feelings instead of facts since there isn’t a single story that ties them all together.

The Engine of Continuity: mySATHI as the LTO’s Main Part

The LTO framework needs more than just a database; it needs a way to measure things that is always the same. This is where mySATHI turns the theoretical observatory into a real thing. MySATHI is not like regular tests that only measure accomplishment at one moment in time. Instead, it is a “Longitudinal Intelligence Platform” that follows a learner across three important life periods, making a single record of the learner.

- 1. SATHI NAO (Grades 8–10):** This is where the trip starts, and it tests early skills and readiness for college. It sets a standard for what is possible before the stress of board exams starts.
- 2. SATHI 4C (Grades 11–12):** The metric changes to “Future-ready university admissions” as pupils move on to college. This stage doesn’t simply test how well you remember the subject; it also tests the 4Cs: Critical Thinking, Creativity, Collaboration, and Communication. These are the stable qualities that the LTO keeps track of over time to see how much progress has been made.
- 3. SATHI-X (Employability):** The framework finally reaches the job market with “Talent & employability analytics,” which checks to see if the educational inputs from earlier steps really did lead to professional skills.

The LTO goes beyond just storing data by using this three-tiered framework to create “Longitudinal Intelligence.” Not only can the observatory detect where a student is, but it can also measure how quickly they are getting better at certain skills, such as “Decision-Making Under Uncertainty” or “Cross-Disciplinary Innovation.”

From Silos to Synergy: Making the Framework Stronger

The addition of mySATHI makes the LTO framework stronger in three distinct ways:

- 1. Making the “Invisible” Curriculum the Same for Everyone:** One of the hard parts about an observatory is that it’s hard to compare apples to apples between different state boards and colleges. mySATHI fixes this by setting a worldwide standard for the “4Cs.” The “Curriculum & Assessment Committee,” which includes specialists from IIT Delhi, Ashoka University, and ETS Global Research, made it such that when the LTO records a “Critical Thinking” score in Mumbai, it means the same thing as one in Nairobi.
- 2. The Trinity of Intelligence:** The LTO wants to be more than just a place to store things. By combining three different layers, mySATHI adds active intelligence to the system:
 - *Assessment Intelligence:* a detailed look of your skills, personality, and agency.
 - *Learning Intelligence:* Giving personalized advice and support based on areas where performance is lacking.
 - *Opportunity Intelligence:* Finding the right school and job for students in real time as their profiles change.
- 3. Finishing the DPI Trilogy:** India has successfully created Digital Public Infrastructure (DPI) for Payments (UPI) and Identity (Aadhaar). The LTO, which is powered by mySATHI, wants to finish this trio by creating the DPI for Learning. It makes it easier for people to get to talent and opportunities by setting up a “Education Observatory” that helps people move across countries in Asia and Africa.

The Dynamic Passport

Finally, combining the LTO framework with the mySATHI platform changes the way we think about things from “filtering out students” to “enabling learners.” It gives each student a “Dynamic Passport of Potential,” a living document that changes as they move ahead. For the state, it gives them the “precision policymaking” lens, which means that decisions are based on long-term results instead of short-term inputs. We make the invisible visible so that talent isn’t lost in the spaces between institutions; instead, it is nourished through a journey that is constantly floodlit by data.

These three case studies show how the assessment suite (SATHI NAO, SATHI 4C, and SATHI-X) helps students make important transitions in their lives: from school to college, from graduate school to work, and from entry-level jobs to higher-level jobs.

Let’s Look At How This Can Work

Case Study 1: Getting into College from High School

A top private institution is having trouble with a number of students dropping out of its First-Year Engineering program.

The issue is that the university only let students in based on their 12th-grade board marks and standard entry examinations (PCM). Even if the students were good in arithmetic, many of them didn't have the toughness to bear engineering workloads or the critical thinking abilities needed to solve hard problems. This caused a 15% drop-out rate in the first year.

- **The SATHI Solution:** The university used the SATHI 4C Assessment to help them choose new students.
- **Used:** SATHI 4C (Critical Thinking, Creativity, Collaboration, Communication) plus Personality Profiling.
- **Methodology:** The test went beyond just asking questions about physics and math to see how well people could think on their feet and how responsible they were. It found pupils who had somewhat lower board scores but were very “Open to Experience” and “Logical Reasoning.”

Result:

- **Holistic Selection:** The university found a “hidden talent pool” of students who had ordinary board scores but were very good at coming up with new ideas in engineering.
- **Strategic Cohort Analysis:** Faculty got a dashboard that showed the personality of the incoming group. This helped them plan bridge courses for students who didn't have certain attributes, like resilience.

Case Study 2: Getting into Masters/Research

Situation: A research-heavy college or university is picking students for a master's program in data science and policy.

- **The Problem:** Standardized tests like the GRE and CAT were good at finding people who were good at English and Math, but they didn't do a good job of predicting “Research Aptitude.” Students were joining the program, but they were having trouble with difficulties that didn't have a clear structure and group research assignments.
- **The SATHI Solution:** The school used the SATHI Assessment Framework to find out what attributes were important for research success.
- SATHI (Generic Section 1) was used to test Abstract & Logical Reasoning and Verbal Ability, with a focus on comprehension over vocabulary.
- **Methodology:** The test focused on Abstract Reasoning to assess “trainability”—the ease with which a student can understand new, difficult ideas—rather than just testing what they already knew. It also looked at collaboration skills, which are very important for modern interdisciplinary research.

Result:

- **Predictive Validity:** The university discovered a more robust link between the SATHI “Abstract Reasoning” score and the student’s final thesis quality compared to their undergraduate GPA.
- **Diverse Intake:** The test was hard but didn’t favor any one subject, so students from non-traditional backgrounds (such Humanities students who are good at logic) may get into the Data Science stream.

Case Study 3: Newcomers to the Manufacturing and Services Industry

A big manufacturing company is seeking Graduate Engineer Trainees (GETs), and a service sector company is hiring Customer Support Leads.

The Issue:

- **Manufacturing:** New personnel typically didn’t have “process discipline” or “attention to detail,” which led to safety problems.
- **Services:** Hires spoke good English but had low “emotional resilience,” which caused a lot of burnout and attrition in customer-facing professions.

The SATHI Solution: The organizations used SATHI-X (Employability) to find the right behavioral skills for each function.

We used SATHI-X with a bespoke weightage for 4Cs with its sub parts of dimensions.

How to do it:

- **For Making:** The test put a lot of weight on Conscientiousness (being diligent and obeying rules) and Critical Thinking (finding the main cause of a problem). It checked for “Spotting Assumptions & Biases” to make sure everyone was safe.
- **For Services:** The test put Agreeableness (empathy) and Neuroticism (emotional stability) at the top of the list. It looked at “Communication” in terms of “Clarity & Brevity” and “Active Listening.”

Result:

- **Placement Alignment:** The manufacturing company made less mistakes on the shop floor by hiring people who were naturally very conscientious.
- **Retention:** The services company cut down on employee turnover by 20% by not hiring people whose personality profiles showed they wouldn’t do well in high-stress situations.

CHAPTER 18

Internationalisation of Higher Education in India — Aligning mySATHI Objectives with Study India Goals

India confronts a moment of profound possibility in higher education, one that echoes the golden age of yore – of the days of Nalanda and Takshashila, where scholars from distant lands converged to wrestle with philosophy, mathematics, and medicine under vast open skies in the Indian subcontinent. This is what the the NITI Aayog–IIT Madras report of December 2025 seeks to revive, viz this legacy with a compelling roadmap, envisioning India as a global knowledge hub that draws 85,000 to 1.3 lakh international students by 2030 and scales to 3 lakh to 11 lakh by 2047.

This ambition unfolds through joint degrees, research collaborations, international branch campuses, and “internationalisation at home,” underpinned by 22 policy recommendations, 76 action pathways, and 125 performance indicators. Yet as we reflect on this blueprint, a quiet tension emerges: policies, however meticulously crafted, risk remaining abstract unless they touch the lived realities of students navigating opaque pathways from Tier II colleges to global opportunities.

However, a tension exists between these high-level policies and the practical realities of students. For example, a student at an Anna University branch in Coimbatore may be interested in a twinning program but faces deeper questions regarding which specific talents are valuable beyond transcripts and how to navigate the uncertainties of visas and cultural adjustment. Currently, India’s outbound mobility costs the government \$50 billion annually and is dominated by eight metro states, while only 50,000 foreign students enroll in India each year.

The mySATHI Counterpoint

MySATHI offers a strategy to bridge the gap between policy and personal choice. The SATHI assessment shifts focus from rote memorization to 21st-century skills: Creativity, Collaboration, Critical Thinking, and Communication (the 4Cs). Using an AI-proctored platform that handled 12 million evaluations in 2024, it enables students from diverse regions, like Patna or Chennai, to participate in global opportunities. The proposed four-band competency framework puts NITI Aayog’s vision into action through student-centered technology and peer mentoring.

A Five-Point Strategy for Global Skills Certification

Here's a hypothesis for a strategic framework positions mySATHI as a global skills credential to help outbound Indian students and foreign universities make evidence-based choices.

1. Create a Global Readiness Skills Transcript

Foreign universities often prioritize skills-based assessment but struggle to evaluate Indian applicants due to the extreme diversity in curricula and grading. A structured “Global Readiness Skills Transcript” provides a standardized credential alongside board marks and language scores.

Implementation: The transcript maps the 4Cs to performance bands and sub-dimensions.

- **Critical Thinking & Analysis:** Evaluating complex problems, distinguishing facts from inference, and logical reasoning.
- **Creativity & Innovation:** Applying original thinking, ideation, and adaptive problem-solving.
- **Collaboration & Teamwork:** Active listening, conflict resolution, and contributing to collective goals.
- **Communication & Expression:** Clarity of expression, audience awareness, and rhetorical effectiveness.

This transcript differentiates students in competitive pools and helps universities identify high-potential applicants from Tier-2 and Tier-3 cities who may lack access to premium international coaching.

2. Embed mySATHI in Pre-Departure Readiness and Program Choice

Outbound mobility is often driven by general aspirations rather than a structured match between student capabilities and program requirements, leading to high costs and potential underemployment.

Implementation: mySATHI testing is integrated at three junctures:

1. **Class 11:** Establishes a baseline to allow for targeted skill-building.
2. **Pre-final Undergraduate Year:** Tracks evolution and triggers interventions like internships or group projects.
3. **Immediately Before Departure:** Ensures profiles are current for applications.

Calibration: Data can be structured by destination and discipline. For example, US graduate schools may require high Critical Thinking, while Canadian co-op programs emphasize Collaboration and practical Communication. This transforms the journey from aspirational to capability-driven.

3. Offer Foreign Universities a Calibrated Admissions and Scholarship Layer

Most foreign universities lack direct visibility into the skill profiles of Indian applicants . mySATHI

provides a scalable, culturally-informed middle layer for better decision-making.

- **Program Thresholds:** Universities can identify specific 4C profiles; a data science program might prioritize Critical Thinking, while a business school prioritizes Collaboration and Communication.
- **Diversity and Equity:** Universities can use scores to find talented students in under-served areas, expanding diversity without lowering standards.
- **Customized Support:** Profiles allow for tailored orientation programs, such as pre-assigning writing center appointments for students with weaker Communication scores.

4. Use Dashboards for Continuous Calibration and Evidence-Based Choice

The current system lacks longitudinal data linking skill profiles to career outcomes . mySATHI provides three tiers of dashboards to fill this gap:

- **Individual Tier:** Enables students to track their trajectories and identify specific gaps.
- **Institutional Tier:** Helps schools and colleges benchmark cohorts and design targeted curricular interventions.
- **University Partner Tier:** Provides foreign institutions with insights into applicant trends and correlations between 4C profiles and academic success.
- **National Evidence:** Aggregated data becomes a national asset, identifying underdeveloped skills and informing policy on internationalization and labor market preparation.

5. Align with Policy Frameworks and Employability Narratives

MySATHI aligns with the National Education Policy 2020 and UGC guidelines on internationalization, which emphasize competency-based education.

- **Domestic Alignment:** The 4Cs map to the National Higher Education Qualifications Framework (NHEQF). High scores can be used to award “credit” or exemptions for specific university courses.
- **Global Alignment:** Task types mirror the learning practices used in the US, UK, Canada, and Australia. Engagement with multinational employers validates the framework’s relevance to the global labor market.
- **Trust and Partnerships:** This alignment signals to foreign partners that Indian institutions are committed to globally-benchmarked, outcomes-oriented education.

MySATHI has the potential to transform Indian student assessment by moving beyond marks-based evaluation to systematic 4C measurement. This strategy aligns with India’s broader aspirations to move from “brain drain” toward intentional global engagement. By demonstrating that Indian institutions invest in internationally comparable competency measurement, India strengthens its position as both a source of talent and a destination for global research, supporting the vision of Viksit Bharat 2047.

Discipline-Specific Competency Needs

Foreign university admissions and scholarship committees often look for specific “recipes” of the 4Cs depending on the field of study.

- **STEM & Data Science:** Requires **Critical Thinking** of high order and **Communication** as the second highest priority to articulate technical findings to non-specialists.
- **Business & MBAs:** Prioritizes **Collaboration** the most, and **Communication** due to the heavy focus on team projects and client-facing roles.
- **Creative Arts & Design:** Places a premium on **Creativity** and **Communication**, with **Collaboration** being essential for team-based studio work.
- **Entrepreneurship:** Requires a balanced high profile across all 4Cs, specifically **Creativity** for ideation and **Communication** for pitching.

Here’s a hypothetical visualized manner that would help learners to align their goals with:

Destination	Primary Skill Emphasis	Secondary Emphasis	Ideal Academic Fit
United States	Critical Thinking & Communication	Creativity	Research-intensive PhDs, Graduate Schools, and Liberal Arts.
United Kingdom	Critical Thinking & Collaboration	Communication	Seminar-style Master’s programs and Architecture/Design.
Canada	Collaboration & Adaptive Problem Solving	Communication (Practical)	Co-op programs and Work-integrated learning.
Australia	Collaboration & Communication	Critical Thinking (Applied)	Industry-linked degrees and Professional certifications.
Europe (e.g., Germany/Nordics)	Creativity & Collaboration	Critical Thinking	Design programs, Digital Media, and Innovation hubs.

CHAPTER 19

Personality and Agency Assessment Suite — Completing the Holistic Readiness Lens

The 4Cs of mySATHI determine if a student has the critical thinking, creative, collaborative, and communication skills necessary for success in today's universities and workplaces. A new suite of assessments that builds on this by delving further into the question: how will every student interact with the system's learning features, feedback, classmates, uncertainty, and opportunities once they're in? Once the cognitive 4C and personality-agency layers are combined, mySATHI goes beyond just being an excellent admission exam and endeavors to become a platform for developmental intelligence that helps universities, individuals and recruiting organizations.

From “Is It Possible?” to “What Are the Odds?”

This Report's main claim is that in order to be predictive, multi-dimensional assessments need to consider not only how well an individual does on exams in the short term, but also what characteristics contribute to their long-term success in school and the workplace or even become an entrepreneur. As we saw in earlier chapters, aggregate ratings reduce several competencies to a single value, which hinders prediction and equity. A personality and behaviour assessment suite enhances the “probability of success” signal without transforming personality into yet another eliminative filter. Can it provide an organized insight into an individual's values, motivation, self-disclosure, locus of control, openness to feedback, and self-awareness.

This suite is based on five supplementary measures that collectively chart values-in-action, learning preferences, agency, reflection, and interpersonal effectiveness. Instead of being used as a rigid admission or selection criterion, each of these is intentionally concise, interpretable, and meant to flow into leadership tracks, advising, and bridge or remedial courses.

(1) Personal Effectiveness (PE) Scale - Interpersonal Effectiveness & Feedback Culture

- **What it captures:** Three behavioural facets - **Self-disclosure**, **Openness to feedback**, and **Perceptiveness** - yielding practical profiles (e.g., Effective, Insensitive, Dogmatic, Secretive, Task-obsessed, Lonely-Empathic, etc.) **Format:** 15 items (0-4 scale); profile mapping with interpretation and development suggestions **Value for user:** Flags collaboration readiness; identifies coaching levers (e.g., raise disclosure, build selective

openness to feedback, strengthen situational perceptiveness) **Use in admissions, selection & mentoring:** Surface likely team behaviors for project-based learning; plan faculty feedback loops and peer-learning norms

(2) LOCO (Locus of Control) Inventory - Agency & Attribution

- **What it captures:** Internality (I), Externality-Others (EO), Externality-Luck (EL) plus I/EO, I/EL, I/(EO+EL) ratios
- **Format:** 30 statements; ratio interpretation guides mentoring for agency without hubris (e.g., strong internality with reflection; shifting from luck-attribution to effort-attribution)
- **Value for User:** Anticipates persistence under challenge; informs design of first-year resilience workshops and advising scripts
- **Use in admissions & mentoring:** Interview prompts on effort vs luck; cohort-level dashboards to tailor interventions

(3) “Who Am I?” Self-Profile - Motivation, Self-Awareness, Proactivity, Action Orientation

- **What it captures:** Four reflective domains scored from a short essay: **Motivation** (external purpose, excellence, influence, family, society), **Self-awareness** (strengths, weaknesses, achievements, self-development), **Proactivity** (internal locus, initiative, resource use), and **Action orientation** (self-discipline, planning, future orientation)
- **Formats:** ~ 400 words across four segments (\approx 40 minutes); structured scoring rubric
- **Value for user:** Deepens understanding beyond inventories; validates interview impressions; informs personalized goals in the first semester

(4) Learning Style Inventory (LSI) - Preferences for Learning & Instructional Design

- **What it captures:** Preferences across **Concrete Experience (CE)**, **Reflective Observation (RO)**, **Abstract Conceptualization (AC)**, and **Active Experimentation (AE)** using nine rank-order sets (4-word items)
- **Format:** 9×4 word rankings; quick profile read-out (CE/RO/AC/AE)
- **Value for User:** Helps faculty or mentor design multiple-mode learning (discussion, concept maps, labs, projects) and equips students / learners to study smarter
- **Use responsibly:** A preference is not a capability limit; avoid tracking or fixed labels

(5) ROOHI (Core Values) - Risk-Taking, Openness, Ownership, Honesty, Innovation

- **What it captures:** A values-in-action snapshot for entrepreneurial and research mindsets needed in the 21st century context

- **Format:** Short self-report; campus-ethics scenarios (recommended)
- **University value:** Reinforces campus culture of integrity and initiative; supports leadership and entrepreneurship programs
- **Note:** ROOHI is part of MySATHI's core-values suite; instruments and norms can be shared for institutional orientation for usage as additional insight on incoming students and better mentoring

Connecting Individual's Skills and Abilities to Organisation Vision

The semantically, computer-based adaptive 4C assessment generates large-scale, fine-grained cognitive profiles using rigorous psychometric design and node-based Computer Adaptive Testing, as discussed in earlier chapters. The pilot study did find some persistent behavioural patterns, such as high- and low-attempters and risk-averse and risk-tolerant, but they cannot be explained by only looking at right-and-wrong scores.

Adding the Personality and Behavioural Suite to this architecture allows universities to differentiate between different mentoring challenges that could appear similar in a score report, such as a low-attempter motivated by anxiety and external-luck attribution and a low-attempter who is simply over-deliberative but extremely conscientious. It gives recruiters much benefit on the potential recruit and working professionals or individuals ... more insight into their ability to turn entrepreneurs over a period of time. Start up founders can empanel into the mySATHI ecosystem to get a well-structured database of skill-sets to recruit people that align with their organizations.

The Personality Suite provides measures of non-cognitive qualities in an organized and repeatable manner that can be tracked across admission, academic advancement, and placement, as well as career growth cohorts, thus connecting that vision with institutional practice. Moving evaluation from a static state of gatekeeping to a dynamic state of evidence for educational design, universities and employers can test hypotheses over time, such as whether a shift from external-luck to internal-agency, or an increase in feedback transparency and ownership, predicts lower attrition, higher thesis quality, or better role-person fit.

CHAPTER 20

Looking into the Future — A Manifesto for Predictive and Holistic Assessment

The Convergence of Science and Empathy

(*Purpose:* To bridge the gap between the technical contributions and the human outcome.)

The journey through this volume has taken us from the algorithmic precision of Computer Adaptive Testing to the sociological imperatives of equity in admissions. While the contributors—scientists, academicians, and technologists—approach the problem from different vantages, a singular consensus has emerged: The “snapshot” model of assessment is broken.

We stand at a technological inflection point. As demonstrated in earlier chapters, we now possess the computational power to measure not just what a student *knows* (content), but how they *think* (cognition) and who they *are* (personality). It is well within our imagination to lay a roadmap for the next decade of assessment science.

The Core Shift: From Elimination to Enablement

(*Purpose:* To articulate the central thesis of “Prediction over Selection.”)

The dominant model of the 20th century was the Funnel: pour millions of students in, filter them through a rigid exam, and select the few who fit the mold. This volume proposes a new model: the Prism.

The Flaw of Elimination: Traditional tests generate high “false negatives”—rejecting candidates who have the aptitude but lack the specific exam-cracking training.

The Promise of Prediction: A predictive assessment does not ask, “Did you memorize the formula?” but rather “Do you possess the *Fluid Intelligence* to derive a formula when needed?” Shifting the focus to foundational traits—Critical Thinking and Abstract Reasoning, we can predict professional adaptability better than domain knowledge.

The Technological Backbone: The Next-Gen Adaptive Engine

(Purpose: To summarize the tech inputs from your report.)

The transition from linear testing to AI-driven Computer Adaptive Testing (CAT) is not merely an efficiency upgrade; it is an equity upgrade.

Precision Measurement: As noted by [Insert Name of Tech Provider from your volume], adaptive engines allow us to find the “true floor and true ceiling” of a student’s ability without the fatigue of a 3-hour linear exam.

Scalability: The ability to deploy secure, randomized, and calibrated assessments on-demand allows us to reach the “last mile” learner, breaking the geographical monopoly of coaching hubs.

Scope for Further Research: The Unanswered Questions

(Purpose: To provide the specific research agenda you requested. This is the “Call to Action” for the scientific community.)

While we have established the *need* for change, the *science* of that change requires rigorous expansion. We propose three critical pillars for future research:

I. Longitudinal Validity Studies (The “Time” Variable)

The Research Gap: Most validation studies are cross-sectional—they correlate entrance scores with 1st-year grades. We need to perfect the Longitudinal Talent Observatories that can track cohorts over 5, 10, and 15 years.

Research must test if “grit” (Conscientiousness) and “adaptability” (Openness) measured at age 17 are better predictors of mid-career success (age 30) than PCM scores.

II. The Psychometrics of Non-Cognitive Traits in High-Stakes Settings

The Research Gap: We know personality matters, but how do we measure it without students “faking good”?

The New Frontier: Developing “un-gameable” assessment formats for traits like Integrity and Collaboration. Research into Game-Based Assessments and AI-interview semantics could offer breakthroughs in measuring soft skills objectively.

III. Neuro-Cognitive Correlates of Aptitude

The Research Gap: Current testing is behavioral.

The New Frontier: Interdisciplinary research connecting cognitive science with psychometrics. Can we identify the neural markers of “fluid reasoning” or “working memory” to build assessments that are culturally neutral and purely aptitude-based?

IV. Implications for Industry: The “Employability Quotient”

(Purpose: To address the corporate recruitment aspect.)

For the corporate sector, the “degree” is becoming a depreciating currency. The future of recruitment lies in Skill-Based Taxonomy. Instead of hiring for a generic role, organizations

can use predictive data to match a candidate's inherent personality profile (e.g., High Extraversion + High Logic) to the specific demands of a role (e.g., Sales vs. R&D). Also, it can reduce employee churn, because suggests that alignment between personal aptitude and job role is the single biggest predictor of retention. Assessments must evolve to measure this "fit."

Of course, the idea needs to be stretched taut to gauge and predict skills matrix for becoming an entrepreneur.

The Moral Imperative

We end this volume with a reminder that assessment is not just a statistical exercise; it is a social one. Every data point is a human life. By moving away from the brutal efficiency of elimination toward the nuanced science of prediction, we do more than just improve admissions; we honor the diverse potential of the human spirit. The science is ready. The technology is here. The only variable remaining is our will to adopt it.

Acknowledgements and Contributions

- **Research and Analysis Team:** Debjyoti Das Adhikary; Shubhi Shukla; Illa Sai Ravi Teja; Abhishek Kumar
Dr. Partha P Chakrabarti (Research Lead & Supervisor)
- **Assessments Design :** Dr. Balasubramanian S; Shiva Kumar R
- **Technology and Infrastructure:** Dr Pareshnath Paul; Akbar Ali Major; Shubham Goyal; Unnati Rastogi
- **mySATHI Pilot Assessment team:** Sujatha Kshirsagar; Amitendra Kumar
- **Research Support:** Chintan Turki; Shailesh Vasanji Surti; Arathi Chintan Ghatge

Glossary of Terms Used

4Cs – A bouquet of futuristic skills comprising Critical Thinking, Creativity, Collaboration and Communication

ALR – Analytical and Logical Reasoning

CAT – Computer Adaptive Test

CBT – Computer Based Test

QADI – Quantitative Aptitude & Data Interpretation

SIA – Situational Awareness

VARC – Verbal Ability & Reading Comprehension

CL – Career Launcher (for coordinating the Pilot assessment efforts)

DEXIT – Infra and tech services provider for mySATHI assessment

Research Report 2026

Inaugural Edition

The 21st century is marked by a profound transformation in education and employment. Universities are under pressure to identify students who can succeed academically while also thriving in a world of constant change. Employers, on the other hand, seek graduates who demonstrate transversal skills such as critical thinking, creativity, collaboration, and communication (the 4Cs). Traditional examinations, often focused on rote memory and narrow subject mastery, are increasingly misaligned with these needs. The mySATHI Foundation has introduced mySATHI, a computer-adaptive, on-demand assessment that goes beyond conventional metrics. Yet, for such an innovation to gain acceptance globally, it requires an institutional anchor of credibility—a body dedicated to research, validation, and continuous improvement.

This research report seeks to document the entire journey of mySATHI from its conception to its present state of readiness. The volume includes an introduction to the concepts, its various dimensions, how mySATHI seeks to assess the 4C skills, besides a clinical statistical analysis of the results of the mySATHI pilot conducted in the last quarter of 2025. This Research report also records the various types of inputs from individuals and organizations that gave direction and solidity to mySATHI's work.

The entire team takes pride in presenting this Report which we are sure will be the fount for further discussion and research in the domain.

Distributed by:

CL Educate (Publication Division)

 gkp@gkpublications.com

 www.gkpublications.com

 /gkpub  @gkponline



For more books

