happiest minds
AI FIRST. AGILE ALWAYS.

# Unlocking EdTech's Potential with Item Response Theory (IRT) - Smarter Assessments, Better Learning



# ABSTRACT

As learning becomes more personalized and AI-driven, the way we measure progress needs to evolve too. Traditional tests treat all students and all questions the same, often missing the finer details — like how tough a question really is, or whether a student is guessing.

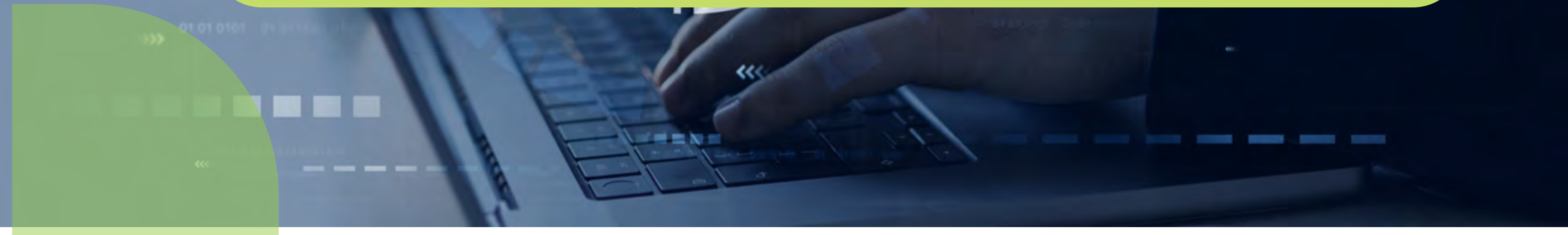
Item Response Theory (IRT) offers a smarter way. Instead of just recording right or wrong answers, it interprets what each response says about a learner's true ability. This whitepaper explores how IRT can power adaptive assessments, smarter diagnostics, and more personalized learning journeys. We also take a product-focused view of how EdTech platforms can put IRT into practice.

# CONTENTS

# INTRODUCTION

In digital education, the demand for fair, accurate, and adaptive assessments has never been higher. Fixed tests and raw scores rarely capture the uniqueness of a learner's journey.

Item Response Theory changes this by asking a deeper question: What does a student's answer really reveal about their ability? Instead of treating all items equally, IRT brings in context — difficulty, discrimination, and even the chance of guessing. The result: adaptive tests that adjust on the fly, calibrated question banks, and insights that go far beyond scores.



## WHAT IS IRT?

Item Response Theory (IRT) is a group of mathematical models used to predict how a student will respond to individual questions, based on their ability and characteristics of the question.
Key concepts in IRT include:

- **Ability**: The latent skill level of the student.
- **Difficulty**: How challenging a question is.
- **Discrimination**: How well does a question differentiate between stronger and weaker students.
- **Guessing**: The likelihood of a student guessing the answer correctly (relevant for MCQs).

This makes IRT very different from traditional scoring. Instead of simply tallying correct answers, it gives a richer, more adaptive picture of both learner progress and item quality.



© Happiest Minds 2025

## HOW IRT WORKS IN THE EDTECH WORLD?

Here's how IRT enhances different areas of an EdTech platform:

**Adaptive Testing (Uses ability θ to select next item):**

Instead of fixed tests, IRT enables assessments that change in real-time based on each answer. If a student answers a question correctly, the next one is a bit harder. If they answer incorrectly, the system tries a slightly easier one. This helps reach an accurate estimate of ability quickly.

**Calibrated Question Bank (Uses b, a, c values to auto-label questions):**

Teachers and content creators often struggle to tag questions manually as 'easy', 'medium', or 'hard'. IRT assigns these values statistically based on actual student performance. So, when a new student takes a test, the system already knows what type of question it is.

**Personalized Learning (Based on θ progression over time):**

Once a student's ability is estimated, the system can recommend lessons, videos, and practice questions matched to their level. A student with $\theta = 0.5$ might get moderate-level practice, while a student with $\theta = -1.5$ will be recommended more foundational material.

**Data for Educators (Generates skill mastery predictions):**

Instead of just test scores, teachers now get dashboards with ability estimates, strengths/ weaknesses by topic, and even which questions are performing well or poorly. In these learning analytics, IRT helps generate skill mastery predictions.




© Happiest Minds 2025

# TECHNICAL VIEW – IRT MODEL

IRT comes in levels, depending on how much detail you want to capture:

## LEVEL 1:
## One-Parameter Logistic Model (1PL) – Rasch Model

- Considers only item difficulty (b)
- Assumes all items are equally effective at distinguishing between students

Formula:

$$P(\theta) = \frac{1}{1 + e^{-(\theta - b)}}$$

### Interpretation:

The probability a student with ability (θ) answers an item correctly depends only on how difficult the question is.

### Use Case:

- Simple assessments where you assume all questions are well-designed.
- Often used in educational tests with high standardization (e.g., some language assessments).

### The Algorithm (a student's experience in an adaptive test):

**Step 1**

The test begins.

The student is given a question of average difficulty (b = 0).

**Step 2**

Student attempts the question.

**Step 3**

The system checks:

Did the student get it right or wrong?

**Step 4**

Based on the answer:

→ If correct, the estimated ability (θ) increases slightly.

→ If wrong, the θ decreases slightly.



© Happiest Minds 2025

**Step 5**

Using the updated θ, the system calculates:

For each remaining question:

P(correct) = 1 / (1 + exp(-(θ - b)))

**Step 6**

The student gets the next question based on this logic.

**Step 7**

This loop continues:

Question - Response - θ updated - Next best question Until the student's ability estimate stabilizes or time is up.

**Step 8**

At the end, the student receives:

→ His/her ability score (θ)

→ A personalized report showing strengths and improvement areas.

# LEVEL 2:

## Two-Parameter Logistic Model (2PL)

- Involves estimating item difficulty (b) and item discrimination (a)
- Adds discrimination, allowing questions to vary in how well they distinguish students.

**Formula:**

$$P(\theta) = \frac{1}{1 + e^{-a(\theta - b)}}$$

## Interpretation:

- Steeper curves (higher 'a') mean the question is very effective at telling apart students.
- Flat curves (low 'a') mean the question is less useful diagnostically.

## Use Case:

- When you want more precise measurement and not all your items are equally strong.
- Suitable for adaptive testing and personalized learning systems.



© Happiest Minds 2025

## The Algorithm (a student's experience in an adaptive test):

**Step 1**

The student begins the test (starting ability θ = 0).

**Step 2**

System picks a question based on:

→ Difficulty (b) ≈ your θ

→ Discrimination (a) = how "informative" the question is

**Step 3**

The student attempts the question

**Step 4**

System updates the ability:

→ It checks how strongly the question discriminates.

→ Discrimination (a) = how "informative" the question is

**Step 5**

Using the updated θ, system calculates for all remaining items: For each remaining question:

P(correct) = 1 / (1 + exp(-a × (θ - b)))

**Step 6**

System picks the next question where:

→ Discrimination is high (a >> 0).

→ Probability ≈ 50% > maximizes learning about the student

**Step 7**

Repeats until enough info is gathered:
Question > Response > θ updated > Next best question
Until the student's ability estimate stabilizes or time is up.

**Step 8**

At the end, the student gets a detailed profile showing the strengths based on how he/she performed on high-discrimination questions.

## LEVEL 3:
### Three-Parameter Logistic Model (3PL)

- Estimating item difficulty (b), item discrimination (a), and an additional parameter, pseudo guessing (c)
- The guessing parameter accounts for the chance of answering correctly by guessing.

**Formula:**

$$P(\theta) = c + \frac{1 - c}{1 + e^{-a(\theta - b)}}$$



© Happiest Minds 2025

### Interpretation:

- Useful in multiple-choice settings where students might guess. .
- The curve doesn't start from 0 on the y-axis but from the value of c

### Use Case:

- Online quizzes or standardized tests with MCQs.
- Useful in EdTech platforms that use auto-generated or low-stakes practice tests.

## The Algorithm (a student's experience in an adaptive test):

**Step 1**

The student begins the test (starting ability $\theta = 0$).

**Step 2**

System picks a question based on:

→ Difficulty (b) ≈ your θ

→ Discrimination (a)

→ Guessing Factor (c)

**Step 3**

The student attempts the question

**Step 4**

System estimates the probability of answering correctly using:

$P(\text{correct}) = c + (1 - c) / (1 + \exp(-a \times (\theta - b)))$

**Step 5**

Based on the response, system updates θ:

→ If your answer seems like a guess (correct on a hard item), it's discounted:

→ The model assumes the student *might* have guessed

**Step 6**

System picks the next question where:

→ Discrimination is high

→ Guessing chance is low

→ Probability of success ≈ 50%

**Step 7**

Repeats until enough info is gathered:

→ Guessing gets filtered out

→ Real ability emerges

**Step 8**

At the end, the student gets a report about:

→ his/her estimated ability

→ how confident the system is (precision)

→ where he/she guessed vs where they actually knew the answer



© Happiest Minds 2025

## Advanced Levels & Concepts

### 4PL Model:

- Adds a slipping parameter (d) — accounts for errors from high-ability students (e.g., careless mistakes).
- Rarely used in practice but useful in simulations and research

**Formula:**

$$P(\theta) = c + \frac{d - c}{1 + e^{-a(\theta - b)}}$$

## Multidimensional IRT (MIRT):

- Models more than one ability at the same time (e.g., Algebra skills + Reading Comprehension).
- Each item may measure a combination of skills.
- Helpful in complex diagnostics like cognitive ability profiles or integrated STEM assessments.

## WHAT IS ICC IN IRT?

An Item Characteristic Curve (ICC) is a graphical representation that shows how likely a student is to answer a question correctly based on their ability level (θ).

**In simple terms, imagine plotting:**

X-axis: Student ability (θ) — from low (-4) to high (+4)

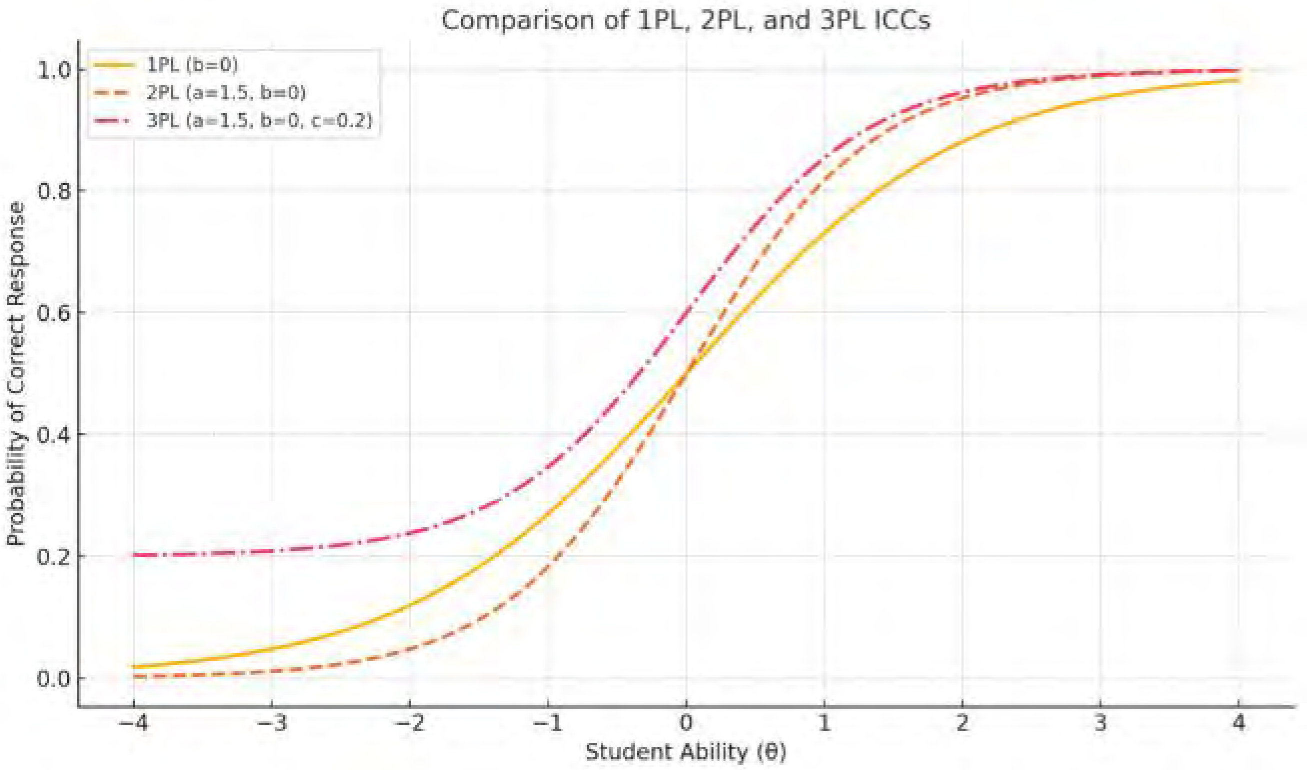
Y-axis: Probability of answering the question correctly — from 0 to 1



© Happiest Minds 2025

**The curve you get shows:**

- Where the question is easy or hard
- How well the question distinguishes between high- and low-performing students
- Whether students can guess the answer




For the various IRT models:

- In **1PL** (Rasch) Model, the ICC shows a smooth S-curve, centred around difficulty (b=0), all items assumed to discriminate equally.
- In **2PL** Model, the ICC shows a similar S-shape, but steeper. Discrimination (a=1.5) affects how sharp the curve is — higher 'a'  steeper curve  better at distinguishing abilities.
- In the **3PL** Model, the ICC starts above 0 on the Y-axis because of guessing (starts at 0.2 on Y-axis), reflects realistic behaviour is MCQs.

For **4PL**, it usually adds an upper limit ($d < 1$), modelling student carelessness or fatigue — not all high-θ students always get the question right.


© Happiest Minds 2025

**Example:**

If a question's ICC is flat, it doesn't help differentiate between students — even low- and high-ability students have the same chance. That means it's a weak question and should be revised or removed.

**Why it is important?**

- Teachers can see if an item works well — e.g., is it too easy or too hard?
- TeaProduct teams can design better adaptive tests — picking items whose ICCs match a student's θ levelhers can see if an item works well — e.g., is it too easy or too hard?
- Analysts can use ICCs to evaluate and calibrate item quality statistically

## USE CASE - Adaptive Quiz Engine for a High School Math Platform:

Imagine a high school math platform called MathEdge, used by thousands of students across India. The platform's goal is not only to test students but to understand their true capabilities and help them improve.

Previously, MathEdge gave every student the same 20-question test. While this seemed fair, it failed to reflect students' real strengths and weaknesses.

To make assessments smarter, MathEdge implemented an IRT-powered Adaptive Quiz Engine. Here's how it transformed the learning experience:

### Priya, a 10th-grade student, logs in to take a 15-question algebra test.

- The first question is of medium difficulty ($b = 0$). Priya answers it correctly
- The system estimates her ability to be above average and selects a harder question.
- As Priya progresses, the system adjusts question difficulty based on each response.



© Happiest Minds 2025

**Behind the scenes, each question has values like:**

- Difficulty (b): Ranges from -2 (easy) to +2 (hard).
- Discrimination (a): Measures how sharply a question differentiates learners.
- Guessing (c): Adds realism for MCQs (e.g., 0.25 for 4 options).

**As Priya answers, the system continually updates her ability score (θ) and recommends the most suitable next question**

By the end, Priya receives a personalized report showing:

- Ability score (θ): 1.2 (above average)
- Strengths: Linear equations
- Areas to improve: Word problems

– Teachers use dashboards to get actionable insights across their class.

– Students benefit from smarter feedback. Everyone wins.

**NOTE**: Please note that to implement **1PL, 2PL, or 3PL IRT models** in an EdTech assessment system, there are some essential **pre-requisites**—ranging from content readiness to data infrastructure and psychometric validation.

# CONCLUSION

Item Response Theory (IRT) offers a transformative approach to learner assessment, moving beyond traditional one-size-fits-all models to enable adaptive, data-driven evaluation. By accounting for item difficulty, discrimination, and guessing behaviour, IRT empowers educators and platforms to measure learner ability more accurately and fairly. In the context of EdTech, integrating IRT enhances everything from personalized quizzes to high-stakes testing. It allows for smarter item selection, better feedback loops, and scalable adaptive learning systems that evolve with each learner's journey. The simulations and visualizations presented in this paper demonstrate how different IRT models (1PL, 2PL, and 3PL) provide increasing levels of precision, offering clear benefits for product teams building intelligent assessment tools

As education continues its digital evolution, adopting IRT models isn't just a technical upgrade—it's a strategic step toward equity, personalization, and deep learning analytics. For EdTech leaders and product managers, this presents a unique opportunity to build evidence-based platforms that not only test knowledge but unlock true learning potential.


© Happiest Minds 2025

# AUTHOR BIO

Upasana Das is a Product Manager in the EdTech domain at Happiest Minds, driving innovation across learning platforms, assessment systems, and AI-enabled learner experiences. She has led end-to-end product strategy and delivery for multi-tenant LXPs, proctoring solutions, adaptive assessments, AI mentors, virtual tutors, and other scalable digital learning products.

With a background in business analysis and product ownership, she partners closely with cross-functional teams and customers to align technology with learning outcomes and institutional goals. Upasana has delivered sessions on K12 systems, proctoring trends, AI in EdTech, and Item Response Theory, contributing to thought leadership within the industry.

Recognized for her structured problem-solving, user-centric mindset, and focus on impactful solutions, she continues to build future-ready products that enhance teaching, learning, and digital transformation in education.

## About Happiest Minds

Happiest Minds Technologies Limited (BSE, NSE: HAPPSTMNDS) is an AI-led custome-centric digital engineering and Mindful IT company. With an integrated approach that spans from chip to cloud, Happiest Minds delivers secure and scalable solutions across product engineering, cybersecurity, analytics and automation platforms. Happiest Minds brings purpose and precision to every engagement, helping enterprises solve complex business challenges and fast-track their digital evolution across industry sectors such as Banking, Financial Services & Insurance (BFSI), EdTech, Healthcare & Life Sciences, Hi-Tech and Media & Entertainment, Industrial, Manufacturing, Energy & Utilities, and Retail, CPG & Logistics.

Happiest Minds has been honored by both the Golden Peacock Awards and the Institute of Company Secretaries of India (ICSI) for its exemplary Corporate Governance practices. Guided by its mission of 'Happiest People . Happiest Customers' and consistently recognized as a great place to work, Happiest Minds is headquartered in Bengaluru, India, with a global presence across the Americas, UK, Europe, Australia, the Middle East, Africa, and Asia.

happiest minds
AI FIRST. AGILE ALWAYS.

www.happiestminds.com

For more information, write to us at **business@happiestminds.com**