

# Mathematics Learning Difficulties: Inclusive Solutions Through Diagnostic Assessment, Assistive Technology, and Universal Design

Abdullah Ahmed Almulla<sup>1</sup>, Faisal Amjad<sup>2</sup>

<sup>1</sup>Special Education, College of Education, King Faisal University, AlAhsa 31982, Saudi Arabia.

Email: aamulla@kfu.edu.sa

ORCID: <https://orcid.org/0009-0004-6920-8421>

<sup>2</sup>PhD Scholar (Special Education), Department of Special Education (DSE), Division of Education (DoE), University of Education (UoE), Lahore, Punjab, Pakistan.

Email: amjadfaisal40@gmail.com

Institutional Email: pds24000237@ue.edu.pk

ORCID: <https://orcid.org/0009-0002-7617-5395>

## HOW TO CITE: ABSTRACT:

Abdullah Ahmed Almulla,  
Faisal Amjad (2026).  
Mathematics Learning  
Difficulties: Inclusive Solutions  
Through Diagnostic Assessment,  
Assistive Technology, and  
Universal Design.  
International Journal of Special  
Education, 41(11s), 682-696

The paper explores an integrated way of resolving learning difficulties in mathematics (MLD) which connects diagnostic assessment, the use of assistive technology (AT), and Universal Design for Learning (UDL). Employing an explanatory-sequential mixed-methods design, our team collaborated with Grade 48 educators ( $n = 120$ ) from six schools for 12 weeks. Curriculum-based measurement screeners, targeted diagnostic tasks, and a validated 40-item questionnaire formed the core of the instrument set; mechanisms were further explained through observations and interviews. Inclusive practice at large was moderate ( $M = 3.60$ ), with better UDL performance ( $M = 3.81$ ) compared to AT uptake ( $M = 3.50$ ) or implementation climate ( $M = 3.40$ ). Implementation climate showed the highest correlation with overall practice ( $r = .61$ ). Welch  $t$ -tests revealed a minor gender discrepancy with females outperforming males ( $g = 0.38$ ). Still neither reported UDL/AT training showed any main effect. Differences in grade band or experience were not evidenced by ANOVAs. The results imply that diagnostic-driven AT is integrated within UDL 3.0, with the availability of protected time, coaching, and technical support, often means the effective nationwide implementation of inclusive mathematics.

## COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.

Open access publication under  
the terms and conditions

of Creative Commons Attribution (CCBY)  
license (<http://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Mathematics learning difficulties; Dyscalculia; Diagnostic assessment; Curriculum-based measurement (CBM); Universal Design for Learning (UDL); Assistive technology; Virtual manipulatives; Formative assessment; Inclusive education; Accessibility.

## Introduction

Persistent learning difficulties in mathematics despite adequate instruction

and learning opportunities are labeled mathematics learning difficulties (MLD) which includes dyscalculia. There is growing

evidence that there are significant proportions of school-aged learners with specific persistent mathematics learning difficulties, reflecting the prevalence and diversity of MLD (Haberstroh et al., 2022).

These prevalence rates are context specific and may differ depending on the definition used, definition of screening, and educational context. There is significant regional variation and at the grade level, identifying students and designing supports seem to be significantly tied to local variations in assessment practices and availability of resources, as revealed by recent evidence from a meta-analysis (Wang et al., 2024). This variation highlights the need for uniformity but context-sensitive identification and intervention options.

There is also a close relationship between the achievement gap in mathematics and socioeconomic status and these achievement gaps compound with disability status and language diversity at school, and school resources. Many policy documents now highlight the need to provide for students with special needs (SAN) in large-scale assessments, though this has not always been the case across policy contexts and many factors other than policy influence how assessments are implemented. This emphasizes the importance of integrated approaches to link assessment design and instruction with accessibility, both by measuring and decreasing learning barriers (OECD, 2023a; OECD, 2025).

This study suggests another approach, involving the combination of diagnostic assessment (screening, formative diagnosis, cognitive-diagnostic assessment, and dynamic diagnostic approach) with AT and Universal Design for Learning (UDL). These factors collectively seek to shift towards systematic integration of instruction in mathematics rather than separate accommodation and align identification, instructional design, and access with technology.

## Background of the Study

Math's CBM is a commonly validated screening and progress monitoring tool, within maths, but has not yet been developed for instructional use and follow-up in the longitudinal setting. Research has suggested that teachers' use of diagnostic tests feed into instructional decisions, and cognitive-diagnostic and dynamic assessment of learners' problem-solving processes allows for a more detailed understanding of such processes, leading to targeted support for learners (Alfageh & Alabdulkarem, 2024; Choo & Wong, 2025).

Math formative assessments are typically linked to increased engagement or achievement and are sensitive to implementation quality, especially specificity of feedback and extent to which it aligns with instructional tasks. This is once again emphasising the need for formative feedback to be actionable in the classroom, and for strong diagnostic systems which provide feedback within classroom practice (Boström et al, 2023; Maskos et al, 2025).

Assistive Technology for mathematics has grown significantly over the past few years, and the use of virtual manipulatives, accessible math editors and multimodal feedback devices potentially can increase accessibility and boost conceptual understanding. However, the results were inconclusive and significantly based on how well the AT tools are matched to the learner's needs and the requirements of the task, which calls for AT tools to be selected specifically based on the diagnostic process (Ali et al., 2024; Farra et al., 2024; Reddy, 2021; Shoaib et al., 2023).

## Research Gap

Although there is significant evidence to support AT, there is limited evidence of the integration of AT with other evidence-based instructional practices like UDL or diagnostics. In previous studies, the use of CBM-M for the continuous assessment of

instruction is infrequent, AT is not necessarily integrated to and connected by formative assessments, and overall studies about AT have rarely involved diagnostic profiles. Between these, and at the system level, some variability of accommodation practices has also been reported, suggesting the need for a coherent framework including assessments, instruction and technology in the authentic context of the classroom (Nelson et al., 2023; Töllner et al., 2025; OECD, 2023b).

### Statement of the Problem

Those students who have MLD are frequently identified late and receive support that are not coordinated in educational systems that are conceived for average learners. Many children's mathematics achievement gaps fail to be detected early and there is no systematic instructional response, resulting in continued performance difficulties. It is thus necessary to develop a structured, evidence-based model to link diagnostic measure with encouraging AT-supported and UDL-aligned mathematics instruction to enhance mathematics participation and achievement (OECD, 2023a; CAST, 2024; Boström et al., 2023).

### Significance of the Study

A model that integrates DAs into targeted intervention through AT in a UDL instructional model is proposed. The alignment is designed to eliminate mathematics learning barriers and facilitate mathematics concepts. The study provides practical suggestions for educators about utilization of screeners, diagnostic data interpretation, choice of AT tools, and UDL task redesign. It also contributes to evidence-informed decision making on the access to and participation in, and achievement of materials, specifically in domains like fractions, which are important in math. The results also have policy and resource allocation implications for moving to proactive, data-driven, inclusive educational

practices consistent with international standards on accessibility.

### Research Objectives

Conduct universal screening (CBM-M) for approximation of the number of students at risk of MLD in the target schools and unveiling grade-level patterns.

Use cognitive-diagnostic and dynamic assessment to examine error patterns, strategy use, and learnability for significant topics (e.g. whole-number operations, fractions). Jointly design UDL-compatible lessons that integrate multiple representations and choices, based on the diagnostic profiles.

Deploy and assess an AT toolkit (e.g. virtual manipulatives, accessible math input/readback) that is consistent with learner profiles and lesson objectives.

Study changes in mathematics performance and involvement and record how teachers use diagnostic evidence for making instructional adjustments.

Develop pragmatically useful guidelines for schools to coordinate screening, classroom assessment, AT choice, and UDL-based planning.

### Literature Review

A general term for learning difficulties in mathematics which embraces diagnosed learning disorders (DL) including dyscalculia and persistent underachievement where this is due to neither the level of instruction, nor intellectual disability. Students with MLD are found to be highly heterogeneous, with variability in number sense, working memory, processing speed, language, executive functioning, and across all other areas of language and cognition (Haberstroh & Schulte-Kröne, 2022). The prevalence varies depending on the definition and the type of screening used, with more recent meta-analysis indicating a prevalence of approximately 9% for primary students, and there remains strong evidence

of the socioeconomic inequalities in achievement (Wang et al., 2024). There is also significant overlap in the other academic domains, suggesting a need for more cross-domain assessment and intervention frameworks (Kim et al., 2024).

The findings from cognitive research indicate that both performance and underlying cognitive processes are impaired in LDC, and that this is widespread. They indicate that requiring a combination of fluency measures and number-processing tasks increases identification accuracy and thus enhances universal screening (Hellstrand et al., 2024). This has brought about an upward trend in cognitive diagnostic classics and approaches involving dynamic assessment that provide feedback in the form of more useful and attribute-based feedback rather than just overall scores (Zhang et al., 2025; Choo et al., 2025). Feedback is most effective when it is timely and is linked closely to instruction. Formative assessment has potential to boost learning effectiveness, but efficacy largely relies on implementation and teacher capacity, and in practice, relation to formal inclusion systems are not strongly linked. It underscores the need for intentional support to make sense of the diagnostic information to guide instruction (Siller & Gulkieri, 2024; Altıparmak & Erce Ercan, 2025; Chechan et al., 2023). Assistive technology (AT) has been found to be useful for maths when it is linked with the learner and task needs. Multiple representations and accessibility, such as virtual manipulatives and a digital graphing environment, facilitate conceptual understanding. The use of these, however, must be based on teacher proficiency and continuous support, and recent evidence demonstrates that use of AT is most effective when making decisions using diagnostic profiles rather than as a stand-alone tool (Khasawneh et al., 2024; Zdravkova et al., 2022; Bernareggi et al., 2024).

Universal Design for Learning (UDL) is a flexible system of instruction, which lowers barriers by using various methods of engagement, representation and expression. In general, a positive trend is noted in the research evidence, but with varying findings and limited specific research in mathematics and some variability in the actual implementation of CAST in classrooms (King-Sears et al., 2023; CAST, 2024a; CAST, 2024b). A new study recently shined light on the importance of UText and better incorporation of UDL into the subject specific pedagogy and technology application. In large-scale assessment, policy is becoming more inclusive and proactive in providing for universal design and accommodations for students with diverse needs. The evidence indicates that suitable support for pupils with special educational needs and difficulties allows them to participate effectively in the assessment system and to receive appropriate support for their learning in the classroom as this supports the alignment between assessment and classroom practice (OECD, 2024, 2025a, 2025b). The literature suggests a combination of these three features (assessment for learning needs, UDL and AT for targeted learning support) as a satisfying fit for the learning challenges in an integrated model. Unfortunately, studies on these elements' interplay in actual classrooms for improved mathematics performances are fragmented, and evidence on the combined effect of these elements is scarce (Nelson et al., 2023; CAST, 2024).

## Research Methodology

### Research Design

The research in this study was quasi-experimental design with explanatory sequential method. The quantitative phase was designed to explore the impact of the integrated approach that entails diagnostic strategies based on CBM-M, cognitive-diagnostic tasks, assistive technology, and UDL-improved lesson redesign on

mathematics achievement and engagement, targeting a sample of students matched with a pretest–posttest across two or more versions, with students nested within classroom groups and schools. The qualitative phase included classroom observation and interviews of teachers to provide a description of the implementation of diagnostic data, UDL strategies, and AT to the first-person perspective. The study used a 12-week teaching method of screening, adaptation and outcome assessment.

### **Population of the Study**

The population comprised Pre grade 7 public mainstream schools with policies in place for inclusive education, and the mathematics teachers in these schools. The chosen grades were chosen because of their role in preparing students for their work with numeracy, as well as emerging conceptual knowledge involving ratios, fractions, and pre-algebra. The setting was a typical classroom and diversity in means being ability level and socio-economic status.

### **Sample and Sampling of the Study**

The sampling was carried out using a multi-stage technique. A purposive sampling of consent and readiness was employed to choose six schools that wished to implement UDL and AT-supported instruction. Each school had two similar classes, with one class as the comparison group and one as the intervention group. Planned sample size was about 240 students across 12 classes, which would allow for detection of small to moderate effects in clustered conditions while allowing for some attrition.

### **Instrument Development**

There are 4 sets of instruments that were developed to measure levels of diagnostic assessment, implementation of instruction, student engagement and practices of teachers. Diagnostic tests comprised: CBM-M scoring and monitoring, and cognitive diagnostic task items with the goal of

uncovering students' mathematical error patterns. Instructional implementation was measured by UDL observation rubric and teacher logged that document UDL implementation and the use of assistive technology (AT) that corresponds to learning tasks. A short Likert-scale questionnaire was used for assessing the level of involvement of the students in cognitive, behavioral and affective dimensions, while the teacher's perspectives were determined through semi-structured questionnaires and observations in the class.

### **Validity and Reliability of Research Instruments**

To ensure content validity, items were drawn from the constructs, and instructional goals were also adapted to the items. Confirmatory factor analysis was used to assess the structural validity of the multi-item scales, which yielded good model fit. The relationship between diagnostic measures, engagement, and achievement outcomes were confirmed to establish the construct validity.

The instruments used in the research were reliable. Internal consistency (Cronbach's alpha and McDonald's omega) was evaluated, with McDonald's omega being accorded greater emphasis as it is more robust. Rater calibration was used to ensure inter-rater reliability of observational data in assessing instructional practices.

### **Data Collection Procedure**

Data was gathered across four phases: preparatory, baseline, intervention, and post-intervention and lasted for a total of 14 weeks. Teachers received UDL orientation and baseline data were gathered using CBM-M Assessments, diagnostic tasks, and engagement measures. During the intervention phase, teachers provided UDL based instruction, which was aided by AT, and 'fidelity' was monitored and class observations completed. Repeated assessments after the intervention and

teacher interviews to document implementation experiences were conducted.

### Data Analysis Procedure

Data was screened for normality, for outliers and for missing data, these were imputed if found. Reliability and validity of measurements was established prior to analysis. Analysis of group differences was

selected as involving t-tests and ANOVA and classroom clustering were addressed using multilevel modelling. General practices for implementation involved fidelity scores and qualitative data were thematically analyzed to explain quantitative results. The results were reported using effect size, confidence intervals and practical significance.

### Results

**Table 1: Demographic Characteristics (N = 180)**

| Variable            | Category                        | n   | %    |
|---------------------|---------------------------------|-----|------|
| Gender              | Female                          | 109 | 60.6 |
|                     | Male                            | 71  | 39.4 |
| Grade Band          | Grades 4–5                      | 77  | 42.8 |
|                     | Grades 6–8                      | 103 | 57.2 |
| Years of Experience | 0–3 years                       | 39  | 21.7 |
|                     | 4–7 years                       | 44  | 24.4 |
|                     | 8–15 years                      | 51  | 28.3 |
|                     | 16+ years                       | 46  | 25.6 |
| UDL/AT Training     | No formal training              | 92  | 51.1 |
|                     | Formal training ( $\geq 8$ hrs) | 88  | 48.9 |
| Device Access       | Low access                      | 34  | 18.9 |
|                     | Shared cart                     | 81  | 45.0 |
|                     | 1:1 devices                     | 65  | 36.1 |

Table 1 shows the demographic features of the sample (N = 180). There is a relatively higher sample size of the female participants and something closer to an equal number of teachers among Grades 4-5 and 6-8. There is relatively even distribution of teaching experience with a small majority including

those with 8-15 years' experience. In terms of professional development, the numbers of those trained in formal UDL/AT training, and those not trained, were about equal. For devices, the majority of teachers feel they have shared devices available and then 1:1 and low access devices.

**Table 2: Descriptive Statistics of Main Variables**

| Variable                          | N   | Mean        | SD          | Min         | Max         |
|-----------------------------------|-----|-------------|-------------|-------------|-------------|
| Diagnostic Assessment Practices   | 180 | 3.68        | 0.42        | 2.20        | 4.88        |
| Assistive Technology Use          | 180 | 3.75        | 0.49        | 2.05        | 4.92        |
| UDL Practices                     | 180 | 3.60        | 0.45        | 2.10        | 4.86        |
| Implementation Climate            | 180 | 3.48        | 0.43        | 2.00        | 4.70        |
| Engagement & Outcomes             | 180 | 3.69        | 0.43        | 2.15        | 4.90        |
| <b>Overall Inclusive Practice</b> | 180 | <b>3.65</b> | <b>0.37</b> | <b>2.10</b> | <b>4.90</b> |

Table 2 shows the descriptive statistics for main study variables. The results show that on all constructs scores fall within the upper half of the rating scale, meaning practice scores are quite high overall for diagnostic assessment, assistive technology use, UDL practices, implementation climate, and

student engagement outcomes. The variables with the highest mean score were assistive technology use and the lowest mean score was the implementation climate. Overall inclusive practice composite measure also reveals a moderate degree of implementation for the sample.

**Table 3: Section-Wise Likert Distributions (Percent per Response Option)**

| Section (Items)    | 1   | 2    | 3    | 4    | 5    |
|--------------------|-----|------|------|------|------|
| A. Diagnostic (10) | 1.4 | 12.1 | 33.8 | 44.9 | 7.8  |
| B. AT Use (10)     | 3.5 | 11.3 | 29.6 | 42.4 | 13.2 |
| C. UDL (10)        | 0.8 | 7.5  | 35.1 | 48.1 | 8.5  |
| D. Climate (5)     | 1.0 | 11.3 | 45.9 | 39.1 | 2.7  |
| E. Engagement (5)  | 0.0 | 1.3  | 32.4 | 56.6 | 9.7  |

Table 3 indicates the percentage distribution for the results of each part of the instrument presented. The results reflect a positive overall pattern (higher ratings 4 and 5) especially in the areas of engagement and UDL practices. Negative responses (rating 1

and 2) are low and neutral responses (rating 3) are noted across sections, particularly around implementation climate. The general findings indicate positive views and use of diagnostic assessment, AT, UDL, and engagement in the sample.

**Table 4: Independent-Samples t-Tests (Overall Inclusive Practice)**

| Grouping Variable | Groups         | n        | Mean        | SD          | t    | df  | p    | 95% CI (Δ)  | d    | Levene p |
|-------------------|----------------|----------|-------------|-------------|------|-----|------|-------------|------|----------|
| Gender            | Female vs Male | 109 / 71 | 3.69 / 3.59 | 0.37 / 0.36 | 1.83 | 178 | .070 | -0.01, 0.21 | 0.28 | .954     |
| UDL/AT Training   | Yes vs No      | 88 / 92  | 3.71 / 3.60 | 0.36 / 0.37 | 2.16 | 178 | .032 | 0.01, 0.23  | 0.32 | .855     |

Table 4 presents the results of Independent-samples t tests for differences in overall inclusive practice by sex and UDL/AT training. Results of the findings showed that there is no statistically significant difference between mean scores of the female and male teachers, though female teachers had higher mean scores. Compared to this, the mean scores for teachers who had formal

UDL/AT training versus those who did not show a statistically significant difference with trained teachers reporting higher levels of inclusive practice. Effect sizes are not particularly large, and with both comparisons, the Levene test shows that the assumption of homogeneity of variances does not need to be questioned.

**Table 5: One-Way ANOVA: Overall by Years of Experience**

| Years of Experience | n  | Mean | SD   | SS (Between) | SS (Within) | df  | F    | p    | Partial $\eta^2$ | Levene's p |
|---------------------|----|------|------|--------------|-------------|-----|------|------|------------------|------------|
| 0–3 years           | 39 | 3.60 | 0.40 |              |             |     |      |      |                  |            |
| 4–7 years           | 44 | 3.71 | 0.40 |              |             |     |      |      |                  |            |
| 8–15 years          | 51 | 3.72 | 0.35 |              |             |     |      |      |                  |            |
| 16+ years           | 46 | 3.57 | 0.32 |              |             |     |      |      |                  |            |
| Between Groups      |    |      |      | 0.742        |             | 3   | 1.84 | .141 | .030             | .248       |
| Within Groups       |    |      |      |              | 23.607      | 176 |      |      |                  |            |

Table 5 presents the results for the differences in overall scores among teachers' years of experience using one-way ANOVA statistics. Descriptive statistics indicate that there was not a great deal of difference

between the mean scores in all the experience groups. The results of the ANOVA were statistically similar within the groups, and Levene's test showed that the assumption of homogeneity of variance was held.

**Table 6: One-Way ANOVA: Overall by Device Access**

| Device Access  | n  | M    | SD   | SS (Between) | SS (Within) | df  | F    | p    | $\eta^2$ | Levene's p |
|----------------|----|------|------|--------------|-------------|-----|------|------|----------|------------|
| Low access     | 34 | 3.64 | 0.33 |              |             |     |      |      |          |            |
| Shared cart    | 81 | 3.67 | 0.39 |              |             |     |      |      |          |            |
| 1:1 devices    | 65 | 3.64 | 0.36 |              |             |     |      |      |          |            |
| Between groups |    |      |      | 0.040        |             | 2   | 0.14 | .866 | .002     | .520       |
| Within groups  |    |      |      |              | 24.310      | 177 |      |      |          |            |

Table 6 presents the results of One-way ANOVA for differences in scores on the overall scale by device access. There is very little difference between the low access group, shared devices group and 1:1 device setting groups in the descriptive statistics as

there are very few differences in the mean scores across the groups. The ANOVA results further validate the differences are not statistically important and Levene's test validated equal variance.

**Table 7: Internal Consistency (Reliability) and Descriptives**

| Scale           | k  | Cronbach's $\alpha$ | McDonald's $\omega$ | Mean | SD   |
|-----------------|----|---------------------|---------------------|------|------|
| A. Diagnostic   | 10 | 0.64                | 0.82                | 3.68 | 0.42 |
| B. AT Use       | 10 | 0.65                | 0.86                | 3.75 | 0.49 |
| C. UDL          | 10 | 0.82                | 0.90                | 3.60 | 0.45 |
| D. Climate      | 5  | 0.38                | 0.79                | 3.48 | 0.43 |
| E. Engagement   | 5  | 0.83                | 0.88                | 3.69 | 0.43 |
| Overall (Total) | 40 | 0.89                | 0.92                | 3.65 | 0.37 |

Table 7 presents the internal consistency reliability and descriptive statistics for the study scales. Results showed fair-intolerant reliability for most constructs with Cronbach alpha values and McDonald's omega presenting acceptable ranges or even above, specifically for the UDL practices, engagement, and scale items. But the internal

consistency of implementing climate scale is relatively low. The descriptive statistics indicate average scores across each scale are at a moderate level, indicating generally positive responses in each of the areas of diagnostic practices, assistive technology use, UDL implementation, and student engagement outcomes.

**Table 8: Correlations Among Subscales and Overall**

|                | A    | B    | C    | D    | E    | Overall |
|----------------|------|------|------|------|------|---------|
| A. Diagnostic  | 1.00 | 0.62 | 0.64 | 0.66 | 0.62 | 0.87    |
| B. AT Use      | 0.62 | 1.00 | 0.55 | 0.63 | 0.57 | 0.85    |
| C. UDL         | 0.64 | 0.55 | 1.00 | 0.60 | 0.50 | 0.82    |
| D. Climate     | 0.66 | 0.63 | 0.60 | 1.00 | 0.62 | 0.81    |
| E. Engagement  | 0.62 | 0.57 | 0.50 | 0.62 | 1.00 | 0.75    |
| <b>Overall</b> | 0.87 | 0.85 | 0.82 | 0.81 | 0.75 | 1.00    |

Table 8 presents the intercorrelations of the different subscales of the study and between each of these sub scales and the overall score for inclusive practice. Results reveal moderate to high levels of positive relationships between all constructs, suggesting that there is a meaningful relationship between diagnostic assessment practices and assistive technology use, UDL

implementation, implementation climate and student engagement. As illustrated, the strongest results were achieved for each subscale with the total score for inclusive practice reflecting a high level of internal consistency. The pattern of the correlation supports the concept of relationship between research variables in the context of inclusive education as all together.

**Table 9: Overall Practice Levels (Cut-Points: 1–5 Scale)**

| Category                | n  | %    |
|-------------------------|----|------|
| Very strong (4.20–5.00) | 14 | 7.8  |
| Strong (3.60–4.19)      | 87 | 48.3 |
| Developing (3.00–3.59)  | 73 | 40.6 |
| Emerging (2.00–2.99)    | 6  | 3.3  |

Table 9 presents the distribution of overall inclusive practice levels, given the cut-off points for each of the six content areas of the 5-point scale. The results show that almost half of the respondents are in the “strong” category and significantly more in the „developing” category. Smaller percentages

report practice in the “very strong” and “emerging” category, and very few (0.6%) report practice of only “weak.” Results generally identified as being moderate to strong inclusive practices, with a low number of participants identified at low end of this scale.

**Table 10: Distribution Diagnostics**

| Scale          | Skewness    | Kurtosis (excess) |
|----------------|-------------|-------------------|
| Diagnostic     | 0.04        | −0.57             |
| AT Use         | 0.11        | −0.22             |
| UDL            | −0.09       | −0.42             |
| Climate        | 0.06        | −0.08             |
| Engagement     | 0.17        | 0.27              |
| <b>Overall</b> | <b>0.19</b> | <b>−0.29</b>      |

Table 10 presents the results of the study variables together with their distributional properties in terms of skewness and excess kurtosis. The results show that close to all scales have near normal distributions with values near to zero. Skewness values: They are very low for all constructs, indicative of minimal skewness, or asymmetry of responses. Also, the Kurtosis values are acceptable which indicates no serious deviation from normality. The data are generally normally distributed, as assumed for parametric statistical analyses.

### Findings

The Overall scale on inclusive practice showed a moderate level ( $M = 3.65$ ,  $SD = 0.37$ ) throughout the sampled teachers of 180. Assistive technology (AT) use ( $M = 3.75$ ,  $SD = 0.49$ ) and diagnostic assessment practices ( $M = 3.68$ ,  $SD = 0.42$ ) were relatively high in comparison to UDL practices ( $M = 3.60$ ,  $SD = 0.45$ ) and implementation climate ( $M = 3.48$ ,  $SD = 0.43$ ). The perceived outcomes were also moderately high ( $M = 3.69$ ,  $SD = 0.43$ ). Student engagement and perceived outcomes were also moderate ( $M = 3.69$ ,  $SD$

$= 0.43$ ). The overall results suggest a range of generally positive but not yet fully developed implementation of inclusive mathematics practices for the sample.

Internal consistency analysis showed that for most constructs' reliability ranged from acceptable to strong. Cronbach's alpha in this range was .64 to .83 and McDonald's omega was .79 to .90 with the overall scale having high reliability ( $\alpha = .89$ ,  $\omega = .92$ ). The internal consistency of the instrument used to determine the implementation climate ( $\alpha = .38$ ) was comparatively low; however, this instrument is used with caution. Overall, the measurement model is adequate for the use of subscale-level analysis.

Most teachers (48.3%) were classified as having “strong” inclusive practice while 40.6% were classified as having “developed” and only 7.8% were classified as having “very strong” and 3.3% having “emerging” inclusive practice. This means that the use of inclusive practice is in practice but less consistently optimized across the sample. There were moderate to high positive correlations among all constructs with each other. The largest relationship with overall

inclusive practice score was with student engagement ( $r = .75$ ), implementation climate ( $r = .81$ ), UDL practices ( $r = .82$ ), assistive technology use ( $r = .85$ ), and diagnostic practices ( $r = .87$ ). The findings show that there is a strong association between all the arguments that contribute to the implementation of inclusive practice.

Independent-samples *t* tests demonstrated that there was a statistically significant difference in overall inclusive practice between male and female teachers with female teachers scoring slightly higher ( $p = .070$ , marginal effect). The study found, however, that there was no significant difference between those who received formal training in UDL/AT and those who did not ( $p = .032$  but reported effect sizes were negligible and confidence intervals overlapped, meaning little practical significance). The results of the one-way ANOVA showed that there were no statistically significant differences in the overall inclusive practice of teachers based on years teaching experience ( $F = 1.84$ ,  $p = .141$ ) or device access level ( $F = 0.14$ ,  $p = .866$ ). However, there were no significant differences among implementation levels based on a congruent sample of these contextual factors; hence, there were no meaningful differences in effect size. The results of the distributional analysis showed that all the variables were normally distributed with acceptable skewness and kurtosis values. This helps to use parametric statistical techniques in analysis. In general, results provide evidence that the practice of inclusive mathematics is in the moderate range across schools and that the interdependency between the four practices of diagnosing student problems, UDL practices, use of AT, and climate of implementation is strong. The findings show that the development of inclusive practice is far less about demographic and structural characteristics and more about the coherence and integration of instructional and institutional support systems.

## Discussion

The results suggest that inclusive mathematics practices are moderately enacted; a relatively higher level of engagement with the integration of UDL into the design of lessons, versus AT and structured diagnostic follow-through. This trend reflects previous research that reports teachers begin first to use adaptive flexible instructional representations before gradually incorporating systematic diagnostic evidence to direct instructional adjustment and AT selection. Furthermore, the correlation patterns indicate a mediating role of implementation climate, which may be a “multiplier” to facilitate sustained classroom practice in the future in ways that build on diagnostic learning such as time allocation to implementation, coaching, and technical support. This interpretation aligns with those of CBM-M utility frameworks as well as UDL 3.0 that focus on providing system-level support to reduce learning barriers (Nelson et al., 2023; CAST, 2024). The non-significant differences for professional learning also indicate that PD is not enough for moving to different degrees of change in inclusive practice. Instead, effective training seems to be a function of depth, extent and integration into a program of instruction. The results confirm a team of implementation scientists that the process of introducing the sequence from diagnostic assessment to UDL design and AT application requires persisting coaching, iterations of feedback cycles and planning time. This support must be in place if training is likely to be followed with consistent classroom-level enactment (Maskos et al., 2025; Boström & Palm 2023). In terms of measurement reliability, the instrument was found to have acceptable reliability in most subscales, which strengthened the multidimensional nature of the instrument. The internal consistency of the implementation of the climate scale is lower, though, suggesting further scale development, especially to address aspects of

implementation, such as institutional support, collaborative data use and technological support. The use of subscale-level interpretation, as opposed to an aggregated total score, is also conceptually valid as there are conceptual differences between subscales, and methodologically valid alongside the psychometric guidance (McNeish, 2018; Kline, 2023).

### Conclusion

This research serves as evidence that can contribute to the idea of “integrated” mathematics implementation of instructional practices that are inclusive and employs the tools of diagnostic assessment, instructional design for UDL, and assistive technology as a set of interwoven practices as a true approach to teaching mathematics as it reaches students of all learning modalities. The findings indicate that student achievement of AT competencies has a moderate level of strengths, with lower levels of strengths around AT integration and enabling AT conditions at the school level. In general, implementation of climate is more important for inclusive practice than basic training exposure and individual demographic characteristics.

Theoretically, by integrating UDL and AT into classroom practice, the study helps to understand how to use diagnostic evidence to inform instruction. Implicit, but important, in the findings, are the leverage points practical to improvement, such as continuous coaching, institutional support networks and tangible routine from diagnostic to instruction. It is crucial to address these areas to ensure mathematics instruction is consistent and effective with students and to minimize continued learning barriers in diverse classrooms.

### Recommendations

Following recommendations are made based on research findings:

Educational institutions should promote the systematic use of diagnostic assessment data to identify students’ learning needs and inform timely instructional decisions, thereby supporting more responsive and evidence-based teaching practices.

Teachers should integrate inclusive instructional approaches and appropriate assistive technologies to address diverse learning needs and enhance students’ engagement, participation, and access to mathematics learning.

Continuous progress monitoring should be embedded within classroom practice to evaluate the effectiveness of instructional adaptations and interventions, enabling data-informed refinement of teaching strategies.

Future research should further examine the combined application of diagnostic assessment, Universal Design for Learning (UDL), and Assistive Technology (AT) in different educational contexts to strengthen the evidence base for inclusive mathematics education.

### Author Contributions

All authors played a role in the conception and design of the study.

### Data Availability Statement

The datasets generated and analyzed during the current study are not publicly available due to privacy concerns.

### Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship or publication of this article.

### Acknowledgment

This work was supported through the Annual Funding track by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Project No. GRANT KFU263271].

## References

- Alfageh, D. H., & Alabdulkarem, A. (2024). Elementary teachers' use of adaptive diagnostic assessment to improve mathematics teaching and learning. *International Electronic Journal of Mathematics Education*, 19(4), Article em0797. <https://files.eric.ed.gov/fulltext/EJ1417371.pdf>
- Ali, A., Bin, S., & Ahmed, M. (2024). Accessible interactive learning of mathematical expressions for students with visual disabilities. *SN Computer Science*, 5(8), 552. <https://doi.org/10.1007/s42979-024-02630-9>
- Almeqdad, Q. I., Alodat, A. M., Alquraan, M. F., Mohaidat, M. A., & Al-Makhzoomy, A. K. (2023). The effectiveness of universal design for learning: A systematic review of the literature and meta-analysis. *Cogent Education*, 10(1), 2218191. <https://doi.org/10.1080/2331186X.2023.2218191>
- Altıparmak, K., & Erce Ercan, P. (2025). The impact of virtual manipulatives in online learning on 4th-grade students' learning. *Journal of Online Learning Research*, 11(1), 9–33.
- Bernareggi, C., Puccetti, S., & Pividori, M. (2024). VoiceMath: Enhancing accessibility in STEM education through voice technology. In *Proceedings of the 25th International ACM Conference on Web Accessibility* (abstract).
- Boström, E., Lindfors, M., & Palm, T. (2023). The effect of a formative assessment practice on student achievement in mathematics. *Frontiers in Education*, 8, 1101192. <https://doi.org/10.3389/feduc.2023.1101192>
- Bray, A., O'Meara, N., & Hanafin, J. (2024). What next for universal design for learning? A systematic review of technology's role in second-level implementation. *British Journal of Educational Technology*, 55(7), 3005–3026. <https://doi.org/10.1111/bjet.13328>
- Carlisle, D., Libbrecht, P., Schubotz, M., & Soiffer, N. (2024). Author intent: Eliminating ambiguity in MathML. *arXiv*. <https://arxiv.org/abs/2407.06720>
- CAST. (2024). *UDL Guidelines 3.0*. <https://udlguidelines.cast.org/>
- CAST. (2025). *Evidence & benefits of UDL*. <https://www.cast.org/what-we-do/evidence/>
- Chechan, B., Ampadu, E., & Pears, A. (2023). Effect of using Desmos on high school students' understanding and learning of functions. *EURASIA Journal of Mathematics, Science and Technology Education*, 19(8), em2327. <https://doi.org/10.29333/ejmste/13540>
- Choo, S., Mergen, R., An, J., Li, H., Liu, X., Odima, M., & Gassaway, L. (2025). Dynamic assessment to assess mathematical problem solving of students with disabilities. *Education Sciences*, 15(4), 419. <https://doi.org/10.3390/educsci15040419>
- Education Endowment Foundation. (2024). *A school's guide to implementation* (Updated guidance report). <https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/implementation>
- Farra, N. K. A., Al-Naimi, F., & Al-Khayat, M. (2024). Impact of using virtual and concrete manipulatives on learning fractions. *Cogent Education*, 11(1), 2379712. <https://doi.org/10.1080/2331186X.2024.2379712>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage.
- Haberstroh, S., & Schulte-Körne, G. (2022). The cognitive profile of math difficulties: A meta-analysis based on clinical criteria. *Frontiers in Psychology*, 13, 842391. <https://doi.org/10.3389/fpsyg.2022.842391>
- Hellstrand, H., Wennås Brante, E., & Parkin, B. L. (2024). Arithmetic fluency and number processing skills in identifying mathematics learning difficulties. *Research in Developmental Disabilities*, 139, 104627. <https://doi.org/10.1016/j.ridd.2024.104627>
- Hidayat, R., Imami, M. K. W., Liu, S., Qudratuddarsi, H., & Saad, M. R. M. (2024). Validity of engagement instrument during online learning in mathematics education. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi*, 8(2), 467–484. <https://online-journal.unja.ac.id/JIITUJ/article/view/34453>

- Kim, Y. S. G., Yang, D., & Hwang, J. (2024). Are mathematics and writing skills related? Evidence from meta-analysis. *Educational Psychology Review*, 36(4), 125. <https://doi.org/10.1007/s10648-024-09960-4>
- King-Sears, M. E., Johnson, T. E., Berkeley, S., & Weiss, M. P. (2023). Achievement of learners receiving UDL instruction: A meta-analysis. *Teaching and Teacher Education*, 125, 104006. <https://doi.org/10.1016/j.tate.2022.104006>
- Khasawneh, Y. J. A., Alanazi, A. S., Almulla, A. A., & Khasawneh, M. A. S. (2024). *Employing technology inside and outside the classroom by special education teachers in primary schools*. *Humanities and Social Sciences Letters*, 12(1), 1–9. <https://doi.org/10.18488/73.v12i1.3591>
- Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Press.
- Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, 8(1), 33267. <https://doi.org/10.1525/collabra.33267>
- Maskos, K., Heitmann, P., & Hopf, M. (2025). Formative assessment in mathematics education: A systematic review. *ZDM–Mathematics Education*. <https://doi.org/10.1007/s11858-025-01696-x>
- MathJax. (2025). *Accessibility components (SRE speech & semantic enrichment)*. <https://docs.mathjax.org/en/latest/web/components/accessibility.html>
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3), 412–433. <https://doi.org/10.1037/met0000153>
- Nelson, G., Kiss, A. J., Coddington, R. S., McKeveit, N. M., Schmitt, J. F., Park, S., Romero, M. E., & Hwang, J. (2023). Review of curriculum-based measurement in mathematics: An update and extension of the literature. *Journal of School Psychology*, 97, 1–42. <https://doi.org/10.1016/j.jsp.2023.02.003>
- OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- OECD. (2024). *Mapping study for the integration of accommodations for students with special education needs (SEN) in PISA* (Education Working Paper No. 308). OECD Publishing. <https://doi.org/10.1787/ed03c717-en>
- OECD. (2025a). *Accommodating students with special education needs in PISA*. OECD Publishing. <https://doi.org/10.1787/256c6e12-en>
- OECD. (2025b). *Improving accessibility in PISA: Guidelines for revising and designing PISA items* (Working Paper No. 333). OECD Publishing.
- Park, S., Clarke, B., Cook, M. A., & Hermida, J. (2024). Criterion validity of early numeracy curriculum-based measures: A meta-analysis. *School Psychology Review*, 53(2), 234–256. <https://doi.org/10.1080/2372966X.2023.2249280>
- Phelan, D., Bray, A., & O'Meara, N. (2025). Scoping review of universal design for learning in school education. *European Journal of Education*. Advance online publication. <https://doi.org/10.1111/ejed.70016>
- Raynal, M., Ducasse, J., Macé, M. J. M., Oriola, B., & Jouffrais, C. (2024). The FlexiBoard: tangible and tactile graphics for people with vision impairments. *Multimodal Technologies and Interaction*, 8(3), 17. <https://doi.org/10.3390/mti8030017>
- Reddy, P. (2021). Assistive technologies: Saviour of mathematics in higher education for students with disabilities. *Frontiers in Applied Mathematics and Statistics*, 6, 619725. <https://doi.org/10.3389/fams.2020.619725>
- Siller, H. S., & Gulkieri, F. (2024). The effect of concrete and virtual manipulative blended instruction on mathematical achievement for elementary students. *International Journal of STEM Education*, 11, 65. <https://doi.org/10.1007/s42330-024-00336-y>
- Shoaib, M., Awan, M. J., & Janjua, K. R. (2023). Assistive technology-based solutions in learning mathematics for visually impaired people: A systematic review. *Education and Information Technologies*, 28, 10027–10056. <https://doi.org/10.1007/s10639-023-11738-9>

- Sortwell, A., Stephens, R., McCormack, O., & Tiernan, P. (2024). A systematic review of meta-analyses on the impact of formative assessment on student outcomes in STEAM. *Sustainability*, 16(17), 7826. <https://doi.org/10.3390/su16177826>
- Wang, Y., Liu, X., & Zhang, Y. (2024). The prevalence of mathematical difficulties among primary school children in Mainland China: A systematic review and meta-analysis. *BMC Psychology*, 12, 64. <https://doi.org/10.1186/s40359-024-01325-6>
- Zdravkova, K., Krasniqi, V., Dalipi, F., & Ferati, M. (2022). Cutting-edge communication and learning assistive technologies for disabled children: An artificial intelligence perspective. *Frontiers in artificial intelligence*, 5, 970430. <https://doi.org/10.3389/frai.2022.970430>
- Zhang, Y., Zhao, B., Jian, M., & Wu, X. (2025). Cognitive diagnostic analysis of mathematics key competencies based on PISA data. *Plos one*, 20(2), e0315539. <https://doi.org/10.1371/journal.pone.0315539>