American Journal of Interdisciplinary Research and Innovation (AJIRI)

ISSN: 2833-2237 (ONLINE)

VOLUME 5 ISSUE 4 (2026)

PUBLISHED BY

E-PALLI PUBLISHERS, DELAWARE, USA

American Journal of Interdisciplinary Research and Innovation (AJIRI)

Volume 5 Issue 4, Year 2026
ISSN: 2833-2237 (Online)
DOI: https://doi.org/10.54536/ajiri.v5i4.8896
https://journals.e-palli.com/home/index.php/ajiri

AI-Powered Tutoring Quality and Perceived Mathematics Learning Outcomes among Higher Secondary School Students

Md. Shahiduzzaman[1*]

Article Information

Received: August 12, 2026

Accepted: October 24, 2026

Published: November 09, 2026

Keywords

Artificial Intelligence, Learning Engagement, Mathematics Education, Partial Least Squares Structural Equation Modelling, Perceived Learning Outcomes, Self-Efficacy, Tutoring Quality

ABSTRACT

AI Tutoring has the potential to facilitate learning in mathematics, including through the quality of instruction, perceived usefulness, engagement, and confidence. This paper analyses these associations on the basis of provided partial least squares structural equation modelling result of 470 records. Perceived AI tutoring quality, perceived usefulness, mathematics learning engagement, mathematics self-efficacy and perceived mathematics learning outcomes are reflected in five reflective constructs and twenty-five indicators. The analysis records 5,000 bootstrap samples and tenfold repeated prediction. The suggested seven direct relationships are all positive with confidence intervals not equal to zero. The quality of tutoring is linked with usefulness, engagement and self-efficacy. The engagement is linked to self-efficacy and perceived outcomes and self-efficacy is linked to perceived outcomes independent of engagement. The model is used to explain 39.5 percent of the perceived outcomes. The indicators loading, internal consistency, convergent, and discriminant validity are reported to fulfill the assessment criteria as stated in the workbook. However, prediction is above the linear standard of nineteen of twenty endogenous indicators, but the differences are minor. These results set engagement and self-efficacy as complementary explanatory paths. Nevertheless, an empirical sample cannot be verified due to hypothetical platform labels and missing fieldwork documentation. The results thus detail the provided records of analysis and are not able to determine causal achievement gains. Substantive educational conclusions will still require independent evaluation and validated recruitment documentation.

INTRODUCTION

Presentation of the correct solutions is not all that mathematics education needs. Students are required to process concepts, select the right procedures, realise mistakes, and persevere when they cannot get answers instantly. Such needs generate an ongoing necessity of assistance that is responsive to personal challenges (García-Martínez *et al.*, 2023). The classroom teaching offers a common starting point, yet students have differences in background knowledge, speed, confidence, and availability of further clarification. One of the potential options aimed at making instructional assistance outside the regular lessons is AI powered tutoring. It is of educational importance because of the way the students experience that support and utilize it.

An AI tutoring system here is conceived as an application or a Web site that will offer explanations, hints, practice tasks, or feedback in response to student queries or responses (Samuelsson, 2023). This distinction divides responsive tutoring with normal calculators, taped classes and websites that stay at the same answer level. The difference is at the core of the research in the sense that not all access to digital content presupposes the instructional adaptation. Resource can convey information without diagnosing a misconception or assisting a learner to reconsider an unsuccessful approach (Kochmar *et al.*, 2020). The current structure deals with perceived quality of instruction in responsive mathematics support.

The fact that there is an answer does not necessarily imply learning. The students can recreate a process without knowing what it assumes, or an assignment without developing the ability to handle a similar problem on their own. On the other hand, properly developed explanation can make learners begin to do more challenging tasks and analyze their own thinking (Zhang *et al.*, 2023). All these possibilities render the interaction between tutoring and learning conditional. A study of AI tutoring would thus consider the mechanisms between teaching implementation and learning impressions to educational, instead of taking the use of technology as an adequate definition of enhancement.

The perceived usefulness is the perceived usefulness of the tutoring. Students can find a system helpful when it facilitates the learning objectives, ensures that studying can be more efficient, and helps to complete the tasks (Marrone *et al.*, 2022). The learning engagement in mathematics is related to attention, efforts, persistence, active efforts, and focus. Mathematics self-efficacy relates to belief in conceptual and problem-solving self-sufficiency. These constructs, though similar, outline various issues about learning. Usefulness deals with the value of a resource, engagement deals with the engagement in learning and self-efficacy deals with the perceived personal capability (Almusaed *et al.*, 2023).

[1] Assistant Professor, Dept of CSE, Dhaka City College, Bangladesh
[*] Corresponding author's e-mail: torunshahid33@gmail.com

The result in this research is the felt mathematics learning outcomes. This wording is deliberate. The provided measurement specification has a student reported outcome construct and no curriculum aligned test scores or authorised achievement records are provided with the results. Therefore, positive outcome rates cannot be seen as confirmed gains in mathematical achievement. Perceived progress is educationally viable since it can influence subsequent involvement and intentions to practice (Neo *et al.*, 2022). However, the difference between perceived progress and actual competence should be apparent during interpretation.

The research comes up with a cohesive model whereby perceived tutoring quality is linked to usefulness, engagement and self-efficacy. Usefulness has a connection with engagement, and engagement with self-efficacy and perceived outcomes (Wang *et al.*, 2023). Self-efficacy brings in another association with perceived results. This arrangement enables the quality of instruction to relate to outcomes via a variety of indirect routes. It also acknowledges that engagement can be related to consequences directly as well as in terms of confidence. The framework thus offers a more differentiated explanation than a single relationship among the adoption of technology and academic performance.

The key aim is to test the existence of the proposed relationships based on the provided analytical results. The investigation is organised in three subsidiary objectives. The former tests the reliability and validity of the reflective measurement model. The second investigates direct, specific indirect and total associations between the constructs (Kumar *et al.*, 2023). The third measures explanatory performance and stated forecast against a linear standard. These goals relate to the provided results, as opposed to an independent reproduction analysis, since the materials were not accompanied by respondent level observations and code that could be executed to produce estimates.

The contribution is thus limited. This manuscript incorporates a specification of a questionnaire, a conceptual framework, and fourteen analytical tables into a logical interpretation of a model. It distinguishes engagement and self-efficacy as independent explanatory pathways and divides between the statistical support and educational effectiveness. The hypothetical tutor labels are used in the workbook, but the provided materials lack the definition of the recruitment setting and the provenance of data. The method and conclusions include these limitations (Fitzmaurice *et al.*, 2021). The analysis is not provided as independently verified evidence of a named school population.

LITERATURE REVIEW

Perceived AI tutoring quality and perceived usefulness

Perceived AI tutoring quality relates to the features of instructions that are encountered by a learner. The provided materials focus on providing clear explanations, clear steps to follow in finding a solution, the appropriateness of the provided solutions to the learner and the provision of explanatory feedback and the overall standard of instructions (Kshirsagar *et al.*, 2022). These aspects indicate the focus on the educational content of tutoring. They do not quantify the level of computational complexity, business brand awareness, or the rate at which systems are updated (Cunska, 2021). Even the technically sophisticated platform can be regarded as bad when its explanations cannot be accessed or its response fails to explain mathematical mistakes.

Perceived usefulness relates to whether tutoring can help with the mathematics work of the student. There is an analytic importance in the difference between quality and usefulness (Soesanto *et al.*, 2022). Quality refers to a judgement of instruction and usefulness refers to its practical contribution to learning activities. Clarity in explanations can be appreciated as it minimizes fruitless confusion or can assist a learner to spend his or her time studying more efficiently (Hossein-Mohand & Hossein-Mohand, 2023). Incidents that are not familiar can be led by logical steps. Proper feedback can help rectify and allow an effective continuation of a task. They experience these things, which gives a plausible basis of a positive association.

The correlation does not have to be deterministic. Students might enjoy explicit teaching and discover it to be incongruent with their immediate curriculum or assessment demands (Utterberg Modén *et al.*, 2021). On the same note, a convenient tool can be viewed as helpful even when teaching is not delivered in an even manner. The proposed association thus relates to a projected inclination throughout ratings, but not a statement that the constructs are interchangeable. H1: Perceived AI tutoring quality has a positive relationship with perceived usefulness among the higher secondary school students (Radović, 2024).

Perceived tutoring quality and mathematics learning engagement

Engagement comprises the distribution of attention and effort towards mathematics learning. The provided tool has perseverance of hard problems and trying out before verifying solutions (Lopes *et al.*, 2023). Such behaviours are peculiar to active learning and passive exposure. A learner who keeps on demanding the full solution can end up wasting a lot of time on a system without involving himself in the process of thinking mathematically (Chai *et al.*, 2022). Participation should then be viewed based on quality of participation, and not just device use or length of session.

Practical and suitably timed teaching can help to facilitate participation. Students might be more willing to continue when they know what the next step in a problem entail (Yue *et al.*, 2022). Explanatory feedback can also give a definite path of returning to a task. Conversely, a correction without reason can disrupt the effort or promote copying of the answer. The suggested framework considers the

perceived instructional quality as a possible correlate of fruitful involvement and acknowledges that motivation, fatigue, classroom anticipations, and preliminary knowledge might also influence involvement (Moliner *et al.*, 2022).

This association is especially pertinent in cases when tutoring is not held under the supervision of the teacher. The learning condition can give less prompting to stay concentrated or describe a solution (Astuti, 2022). Tailored teaching support might be beneficial to have students stay on track, yet the survey itself is not able to know whether or not it actually alters behaviour in the long run. H2: Perceived AI tutoring quality correlates positively with mathematics learning engagement with higher secondary school students (Weinhandl *et al.*, 2022).

Perceived tutoring quality and mathematics self-efficacy

Mathematics self-efficacy is about the beliefs in the ability to perform mathematical activities. The prominent phrase of the questionnaire requires students to think about what they can afford not to get a prompt response in the AI system (Cunska, 2022). This qualification safeguards the conceptual difference between trust in an instrument and trust in individual ability. A learner might have faith in the ability of the system to come up with an answer which is correct but not sure of being able to solve problems on their own (Fornons *et al.*, 2021). The framework focuses on the latter capability judgement.

Instructional explanations can assist confidence in the event that they render a task intelligible and unveil a strategy that can be used. Feedback can also assist students to make sense of errors as being particular and fixable (ÇetiN *et al.*, 2023). These experiences might be related to increased confidence in conceptual grasp, choice of approaches, and multistep problem-solving. There might also be a poor calibration of confidence though when learners confuse successful assisted performance with independent mastery (Alomari & Jabr, 2020). The positive association must then be combined with the lack of explicit measure of achievement.

The suggested direct relationship will enable quality to be associated with confidence in addition to its relationship with engagement (Akın, 2022). This specification acknowledges that a description can still change the way a learner comprehends a task despite the general involvement that does not change. It does not predetermine that there is no need to engage and confidence is developed instantly (Taha & Abdulrahman, 2023). H3, perceived AI tutoring quality has a positive relationship with mathematics self-efficacy in higher secondary school students.

Perceived usefulness and mathematics learning engagement

Students decide on how to allocate scarce time in studying. A resource which is believed to be relevant to learning objectives can get a longer attention span than a resource which is believed to be peripheral (Tiwari, 2023). Usefulness may thus be an effective justification to keep practicing and reviewing feedback and to revisit challenging material. This appraisal in the model proposed relates perceived instructional quality and engagement. The relationship is plausible between the value attributed to tutoring and the effort that was put in learning via it.

The usefulness must not be diminished to speed. Quick achievement is appealing, though effective learning involves learning, verifying and memorizing (Azevedo & Jaques, 2020). The instrument contains support of the learning goals and value of lessons, which is of use even beyond instant generation of answers. However, the survey will not be able to know whether every respondent defined efficiency as deeper learning or just quicker completion. This confusion applies to the assessment of the educational consequences of a positive relationship.

Usefulness may also impact engagement (García-Martínez *et al.*, 2023). Students that put effort in may find additional uses of a system and then rate it more positively. The hypothesised direction does not imply a real-life temporal sequence, but rather a conceptual sequence of the model. H4: The positive relationship between mathematics learning engagement and perceived usefulness among higher secondary school students exists (Samuelsson, 2023).

Mathematics learning engagement and mathematics self-efficacy

The active learning provides the possibility to play with the problems, observe the progress, and reconsider the unsuccessful strategies. Such experiences can be linked to confidence since they give information about the personal capability (Kochmar *et al.*, 2020). This concept of persistence applies especially to mathematics, and a solution that starts out unsuccessful may need a new representation or a new process (Zhang *et al.*, 2023). An engaged student will have an opportunity to experience the fact that a challenging task can be handled. The connection between engagement and self-efficacy is thus about the meaning of the active learning experiences.

The reciprocal relationship can be the same. Already empowered students can challenge more challenging tasks and will persevere longer. These directions cannot be separated by a cross-sectional model (Marrone *et al.*, 2022). The current specification uses engagement as a predictor of self-efficacy to test the explanatory sequence proposed. Alternative ordering is still possible and would need longitudinal or experimental evidence to compare it. H5; mathematics self-efficacy has positive relationship with mathematics learning engagement among school students.

Engagement and perceived mathematics learning outcomes

Concentration, hard work, perseverance and self-initiated effort could give students a greater feeling of improvement. When learners are engaged, they get more chances to

draw a comparison between their initial uncertainty and subsequent knowledge. They can also report that they have better procedure-following or explanation-giving skills (Almusaed *et al.*, 2023). These experiences yield a possible correlation amid engagement and perceived outcomes. This direct association is provided in the model whilst a separate pathway is permitted by the model using self-efficacy (Neo *et al.*, 2022).

Perceived outcomes are susceptible to appraisal bias. Students who like a learning experience can show improvement even where there is little change in performance (Wang *et al.*, 2023). On the other hand, rigorous practice can expose areas of weakness that were not apparent before and have a short-term deceptive noticeable effect. The hypothesised positive correlation thus does not presume a perfect fit between subjective progress and objective attainment (Kumar *et al.*, 2023). H6 postulates that the engagement in mathematics learning is positively related to perceived mathematics learning outcomes among the higher secondary school students.

Self efficacy and perceived mathematics learning outcomes

Confidence can have an impact on student interpretation of mathematics learning. Students who think that they can comprehend and choose relevant approaches can have a more positive attitude to their learning (Fitzmaurice *et al.*, 2021). They can also be more willing to use their knowledge when performing assessment tasks. Self-efficacy offers a proximal correlate of perceived outcomes in the proposed framework. Its inclusion acknowledges that involvement might not reflect the interpretation of capability by the learner (Kshirsagar *et al.*, 2022).

There should be a difference between confidence and the outcome appraisal. Self-efficacy is a matter of belief of what the learner can do whereas perceived outcomes are the improvements attributed to learning. Discriminant validity is particularly crucial in case of their conceptual proximity (Cunska, 2021). There may be high measurement overlap that would overinflate their structural association. The research thus takes into account the heterotrait monotrait rating as well as the path estimate. H7 provides the statement that mathematics self-efficacy has positive relationship with perceived mathematics learning outcomes in higher secondary school student.

Integrated conceptual framework

The seven suggested structural paths are given in figure 1. The only exogenous construct is perceived tutoring quality. The endogenous constructs are usefulness, engagement, self-efficacy and perceived outcomes (Soesanto *et al.*, 2022). There is no direct path between tutoring quality and perceived outcomes that are mentioned and no direct path between usefulness and perceived outcomes. Their associations are consequently by design passed in the incorporated pathways (Hossein-Mohand & Hossein-Mohand, 2023). This limitation is relevant in analyzing indirect relationships and will guard against an unwarranted assertion that an established direct association has vanished.

The model allows a number of sequential indirect relationships as well. The quality of tutoring can be associated with usefulness and self-efficacy linked with engagement (Radović, 2024). Its relationship with perceived outcomes may be by engagement, self-efficacy or combinations of the intermediate constructs. These paths give a systematic understanding of covariance (Lopes *et al.*, 2023). They are not statistical decompositions of the given model and fail to reveal that the given stages took place in the given order among specific learners.

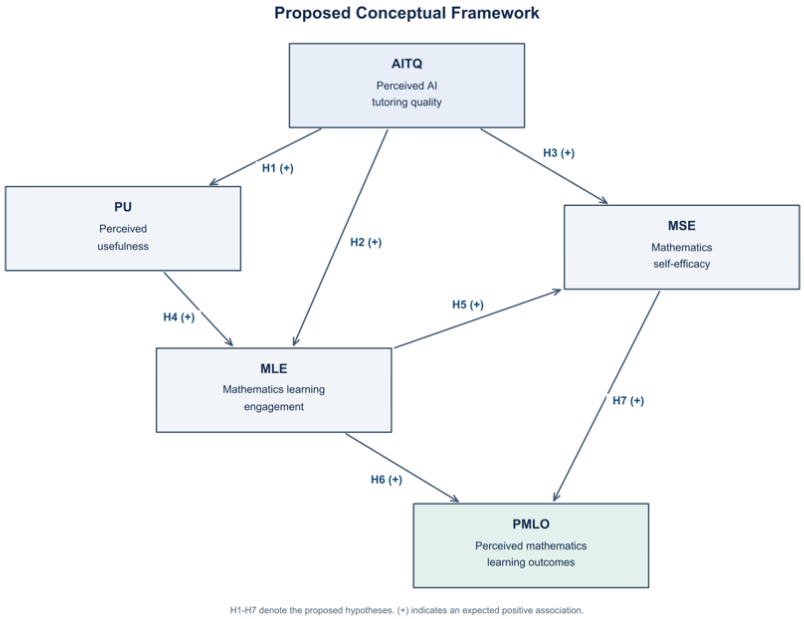


Figure 1: Conceptual framework of the proposed relationships

MATERIALS AND METHODS

Research design and analytical materials

This is a cross-sectional model analysis study through a questionnaire provided, conceptual diagram, structural diagram, and results workbook. The workbook records 470 records of analysis under descriptive statistics and structural evaluations. The materials available can be used to interpret the estimates reported but cannot be used to independently reconstruct respondent selection or reestimate the model. The current document thus differentiates between documented and undocumented analytical procedures and fieldwork practices.

The categories of the platform are clearly marked with Hypothetical Tutor A, Hypothetical Tutor B, and Hypothetical Tutor C. Such labels do not determine whether or not the underlying observations are simulated, anonymised, or are based on real students. There is no country, list of institutions recruited, dates of recruitment, sampling frame, response rate or respondent level dataset provided (Chai *et al*., 2022). To this end, the records are not presented as a tested probability sample or an artifact of an individual national education system. Substantive generalisation involves explanation of provenance and the fieldwork record taken on the original.

Eligibility and participant safeguards

There are four eligibility criteria mentioned in the questionnaire. Respondents must be enrolled in higher secondary school, have utilized AI tutoring in mathematics over the past three months, have gotten responsive explanations or assistance, and can name the system that is used most often. Before proceeding all four screening responses should be positive. This design aims at restricting the responses in relation to the actual experience and creating a uniform reference system. The workbook, however, does not have screening frequencies and exclusions.

The participant information explains voluntary participation, withdrawal prior to submission and report aggregation. It says that there are no applications of names, identification numbers, telephone numbers and account passwords. The questionnaire requires parent or guardian consent and school permission of students who are not yet of the age of independent consent. These are provisions of instruments and not records of verification of implementation (Yue *et al*., 2022). The supplied documents cannot guarantee formal approval, completed consent procedures, data collection permissions, as well as actual storage practices.

Measurement specification

There are five constructs that are defined reflectively and five indicators per construct. All indicators have a five-point agreement scale where strongly disagree or strongly agree is used. According to the questionnaire, there are no reverse coded items. The latent measurement model does not include background and screening questions. Students are asked to reply based on the past three months and the single system that they use as their most used maths tutor. Unanswerable statements must not be answered on the neutral basis but are simply left unanswered (Moliner *et al*., 2022).

The questionnaire provided has the full wording of tutoring quality, usefulness and engagement. It presents three complete self-efficacy statements, and then an incomplete row. The construct specification and workbook however, recognize MSE1 to MSE5 and PMLO1 to PMLO5. The rest of self-efficacy items and all the perceived outcome items do not have complete wording. They can report their numerical outcomes but cannot be absolutely content valid. None of the statements that are missing are reconstructed or presented as administered items.

The tool suggests professional scrutiny, cognitive interviews with students, and a pilot test. It also suggests the need to translate it and review it independently as to its wording in case of administration in Bangla. These suggestions do not prove that (those activities) took place. Measurement evidence in this manuscript will thus be limited to reported indicator statistics and available item content (Astuti, 2022). The constructs mean explain the provided scores; the precise procedure of missing the response treatment and score creation needs the initial record of analysis to be ascertained.

Analytical approach

The workbook documents a biased least squares structural equation modelling and mentions a SEMinR typical PLS summary. Measurement assessment includes outer loadings, Cronbach alpha, rho A, composite reliability, average variance extracted and heterotrait monotrait ratios. The structural assessment includes variance inflation factors, direct paths, specific indirect relationships, total relationships, explained variance and effect sizes. There are 5,000 valid resamples that are reported by the bootstrap tables. The confidence intervals can be given in the form of the 2.5th and the 97.5th percentile in the workbook.

Normal approximations to bootstrap standard errors are reported probability values. Values less than 0.001 are shown as $p < 0.001$ instead of being rounded to 0. Hypothesis testing takes into account the direction of the coefficient and also whether the reported range does not include zero. This convention does not change a statistically supported relationship into a causal effect (Weinhandl *et al*., 2022). There is also repeated prediction with ten folds and ten repetitions in the workbook, where 4,700 of the held-out predictions are used per indicator in the repetitions.

Predictive assessment compares root mean squared error of PLS with a benchmark linear model and takes into account the reported values of predictive relevance. Not additional independent respondents are repeated held out predictions. Additional global fit indices are indicated as not estimated. Assumption does not provide any missing fit statistic, recruitment detail, or any analytical

setting. The presentation of the fourteen table numbers in the order, which they appear in the original work is reproduced in the manuscript and the results are rounded off so that the same figures could be used in comparison with the original values.

RESULTS AND DISCUSSION

Background characteristics

Table 1 characterises the 470 records. Fifteen years is the largest age group with 23.4 percent, then 14 years with 22.6 percent. There are 49.8 and 48.9 percent male and female respectively. The classes eight, ten, and nine share significant proportions whereas the smallest class category is class six. The percentage of 61.1 is attributed to the non-government schools whereas 36.6 is attributed to the government schools. These distributions characterize the analytical records themselves and do not determine whether they are representative or not.

In urban schools, locations are 54.0 percent and rural schools' 44.3 percent. In 64.9 percent of records, there is a report of private mathematics tuition. This background feature is also applicable since tutoring experiences could be accompanied by considerable extra instructional support. Structural model does not regulate on the private tuition, school type, age and location. The estimated relationships must therefore, not be taken to show associations with these background differences taken into consideration (Cunska, 2022). The tiny residual groups are kept so as to maintain the distribution provided.

Table 1: Demographic Characteristics of the Respondents

Variable	Category	n	%
Age group	12 years or below	68	14.5
Age group	13 years	83	17.7
Age group	14 years	106	22.6
Age group	15 years	110	23.4
Age group	16 years	89	18.9
Age group	17 years or above	14	3.0
Age group	Prefer not to answer	0	0
Gender	Male	234	49.8
Gender	Female	230	48.9
Gender	Another identity	2	0.4
Gender	Prefer not to answer	4	0.9
Class	Class 6	61	13.0
Class	Class 7	91	19.4
Class	Class 8	114	24.3
Class	Class 9	99	21.1
Class	Class 10	105	22.3
School type	Government	172	36.6
School type	Non-government	287	61.1
School type	Other	11	2.3
School location	Rural	208	44.3
School location	Urban	254	54.0
School location	Unsure	8	1.7
Private mathematics tuition	Yes	305	64.9
Private mathematics tuition	No	157	33.4
Private mathematics tuition	Prefer not to answer	8	1.7

Note. N = 470 analytical records. Percentages may differ from 100 because of rounding.

Patterns of tutoring use

The analysis in Table 2 indicates that Hypothetical Tutor A counts 50.6 percent, Tutor B counts 30.0 percent and Tutor C counts 19.4 percent of the records. There are no claims of these categories supporting named commercial platforms. Four to six months is the most common and one or two days per week is the most common. The average sessions of 15-20 nine are 39.8 percent. The distributions indicate different exposure in the records as opposed to one uniform intervention.

The major device category is Smartphones, which form 66.8 percent of the total and the main location category is home use, which constitutes 71.1 percent. The account of independent use is 46.2 percent and 20.4 percent respectively with teacher support. In 59.6 percent of the records, there is reported of reliable or very reliable

Table 2: Patterns of AI-Powered Tutoring System Use

Variable	Category	n	%
Most-used AI tutoring system	Hypothetical Tutor A	238	50.6
Most-used AI tutoring system	Hypothetical Tutor B	141	30
Most-used AI tutoring system	Hypothetical Tutor C	91	19.4
Duration of use	Less than 1 month	46	9.8
Duration of use	1-3 months	108	23.0
Duration of use	4-6 months	127	27.0
Duration of use	7-12 months	121	25.7
Duration of use	More than 12 months	68	14.5
Frequency of use	Less than once a week	56	11.9
Frequency of use	1-2 days a week	180	38.3
Frequency of use	3-4 days a week	159	33.8
Frequency of use	5-7 days a week	75	16.0
Typical session length	Less than 15 minutes	73	15.5
Typical session length	15-29 minutes	187	39.8
Typical session length	30-44 minutes	129	27.4
Typical session length	45-59 minutes	55	11.7
Typical session length	60 minutes or more	26	5.5
Main device	Smartphone	314	66.8
Main device	Tablet	53	11.3
Main device	Laptop	79	16.8
Main device	Desktop computer	20	4.3
Main device	Other	4	0.9
Internet reliability	Very unreliable	20	4.3
Internet reliability	Unreliable	64	13.6
Internet reliability	Neither reliable nor unreliable	106	22.6
Internet reliability	Reliable	194	41.3
Internet reliability	Very reliable	86	18.3
Main location of use	Home	334	71.1
Main location of use	School	71	15.1
Main location of use	Coaching centre	50	10.6
Main location of use	Other	15	3.2
Usual support person	Mainly used independently	217	46.2
Usual support person	Teacher	96	20.4
Usual support person	Parent or guardian	44	9.4
Usual support person	Private tutor	71	15.1
Usual support person	Classmate or friend	38	8.1
Usual support person	Other	4	0.9
Mathematics learning activities	Understanding mathematical concepts	309	65.7
Mathematics learning activities	Receiving step-by-step explanations	347	73.8
Mathematics learning activities	Practising mathematics problems	303	64.5
Mathematics learning activities	Checking answers and mistakes	263	56.0
Mathematics learning activities	Completing homework	279	59.4
Mathematics learning activities	Preparing for examinations	299	63.6
Mathematics learning activities	Other learning activities	23	4.9

Note. N = 470. Activity items permit multiple responses. Hypothetical platform labels are retained exactly as supplied.

internet. Such characteristics can be used to shape questions to be used in future implementation research, but the current analysis does not test the difference in devices, connectivity or support. Their prevalence through description should not be construed as evidence that one of the arrangements is more effective.

The most common activity selected is step by step explanations at 73.8 percent. Another popular choice is conceptual knowledge, practice and examination preparation. Activity percentages take the number of records (470 records) as the denominator and can therefore add up to more than a hundred since you can have more than one response. The dominance of explanations is in line with the didactic nature of the framework. However, an activity selection does not show the use of accurate explanations, understanding, and their subsequent application without help.

Construct descriptive statistics

Construct means of between 3.218 and 3.297 are reported in Table 3 on the five-point scale. Self-efficacy is the highest in mean, closely after is the perceived outcomes, and lastly usefulness. The contrasts are not too big and have not been tested as mean contrasts. The standard deviations are the following: 0.976 to 1.016, which means that the dispersion between the records is significant. Constructs all lie within the entire response range. The outcomes thus report fairly positive average ratings and a large amount of variation.

The skewness values are a little negative and excess kurtosis values are also negative in all constructs. These summaries do not indicate extreme asymmetry in the reported construct scores. They do not also standardize single signs or the lack of such strong observations (Fornons *et al.*, 2021). Response patterns, missingness, duplications and unusual cases cannot be analyzed without respondent level data. The descriptive evidence is then supposed to be taken as a summary of the variables of the analytical variables which are exported.

Table 3: Descriptive Statistics of the Study Constructs

Construct	Mean	SD	Min	Max	Skew	Kurtosis
AITQ	3.282	0.998	1	5	-0.222	-0.784
PU	3.218	0.989	1	5	-0.105	-0.873
MLE	3.265	1.016	1	5	-0.317	-0.726
MSE	3.297	0.981	1	5	-0.269	-0.740
PMLO	3.295	0.976	1	5	-0.163	-0.710

Note. N = 470 for each construct. Kurtosis is excess kurtosis.

Indicator reliability

All the twenty-five outer loadings are given in Table 4. Values range from 0.784 for PU3 to 0.867 for MLE5. All the loadings are greater than the 0.708 guideline applied in the workbook, and all the confidence interval reported is positive. These findings help to justify the correspondence of each indicator with its reflective construct in the provided model. There was no loss in

Table 4: Indicator Loadings of the Reflective Measurement Model

Construct	Item	Loading	SE	Lower CI	Upper CI
AITQ	AITQ1	0.820	0.015	0.787	0.847
AITQ	AITQ2	0.857	0.012	0.832	0.880
AITQ	AITQ3	0.841	0.016	0.808	0.869
AITQ	AITQ4	0.833	0.016	0.801	0.861
AITQ	AITQ5	0.835	0.016	0.803	0.862
PU	PU1	0.808	0.017	0.771	0.839
PU	PU2	0.843	0.014	0.814	0.868
PU	PU3	0.784	0.021	0.740	0.820
PU	PU4	0.859	0.013	0.833	0.882
PU	PU5	0.843	0.014	0.814	0.869
MLE	MLE1	0.840	0.014	0.811	0.866
MLE	MLE2	0.842	0.014	0.812	0.867
MLE	MLE3	0.800	0.018	0.761	0.833
MLE	MLE4	0.833	0.015	0.799	0.860
MLE	MLE5	0.867	0.012	0.842	0.889
MSE	MSE1	0.822	0.015	0.791	0.849
MSE	MSE2	0.837	0.014	0.808	0.862

MSE	MSE3	0.799	0.017	0.762	0.831
MSE	MSE4	0.817	0.016	0.781	0.846
MSE	MSE5	0.829	0.014	0.799	0.855
PMLO	PMLO1	0.810	0.017	0.775	0.842
PMLO	PMLO2	0.856	0.014	0.827	0.881
PMLO	PMLO3	0.816	0.017	0.782	0.847
PMLO	PMLO4	0.806	0.018	0.767	0.839
PMLO	PMLO5	0.848	0.014	0.817	0.873

Note. All 25 indicators retained; 5,000 valid bootstrap resamples per indicator. All reported p < 0.001. CI denotes the reported 95% interval.

all the indicators specified. No deletion of indicators is introduced in the process of writing a manuscript, and each construct has the original five-item structure.

Pattern of loading amongst and between constructs is fairly similar. This eliminates the anxiety of a construct being represented as a symbol of a strong indicator and a number of weak indicators. But statistical loading strength cannot substitute an analysis of item meaning. The lack of words in the instrument does not allow fully evaluate whether the indicators reflect a sufficient amount of the intended content. The results are therefore in favor of the reliability of the numerical indicators and leave part of the content validity test unresolved.

Internal consistency and convergent validity

Table 5 reports Cronbach alpha values from 0.879 to 0.893 and rho A values from 0.880 to 0.896. The composite reliability is between 0.912 and 0.921. These values lie within the range of assessment that is given in the workbook, and shows that there is a high level of internal consistency in records provided. Mean variance extracted is between 0.674 and 0.701 (self-efficacy and tutoring quality) respectively. All constructs meet the mentioned 0.50 level of convergent validity.

These results show that indicators have a high degree of variance with the constructs. They fail to show that all the facets of the larger educational concept have been quantified. Reliable ratings can still indicate a small area of content or a general response bias. Reliability and convergence are thus considered as required evidence of measurement as opposed to full validation of the questionnaire. This difference is especially applicable in the case of perceived outcomes as the full wording of item is not available.

Table 5: Internal Consistency Reliability and Convergent Validity

Construct	Alpha	rho A	rho C	AVE
AITQ	0.893	0.894	0.921	0.701
PU	0.885	0.889	0.916	0.685
MLE	0.893	0.896	0.921	0.700
MSE	0.879	0.880	0.912	0.674
PMLO	0.884	0.886	0.916	0.685

Note. Five indicators per construct.

Discriminant validity

Table 6 also gives heterotrait monotrait ratios between 0.334 and 0.646. Ten construct pairs all fall short of the 0.85 standard of the workbook. They also have lower

Table 6: Discriminant Validity Assessment Using HTMT

Construct 1	Construct 2	HTMT	Lower CI	Upper CI
AITQ	PU	0.543	0.465	0.619
AITQ	MLE	0.504	0.422	0.583
AITQ	MSE	0.471	0.385	0.552
AITQ	PMLO	0.359	0.266	0.447
PU	MLE	0.607	0.538	0.671
PU	MSE	0.418	0.329	0.506
PU	PMLO	0.334	0.240	0.420
MLE	MSE	0.636	0.567	0.698
MLE	PMLO	0.604	0.533	0.673
MSE	PMLO	0.646	0.575	0.712

Note. All HTMT values < 0.85; all upper confidence limits < 1; 5,000 valid bootstrap resamples.

confidence limits which are less than one. The first ratio is about self-efficacy and perceived outcomes, then there are engagement and self-efficacy. The two conceptually related pairings are distinguished by the reported values which justifies their empirical difference in the provided measurement model. The least ratio is on usefulness and perceived results.

The structural paths are critical to be interpreted with the help of discriminant validity. In the event that the measure of engagement, confidence, and perceived outcomes were, in fact, the same, their approximate associations would provide a limited explanatory ability (Alomari & Jabr, 2020). The listed ratios lower the specified concern. They do not rule out similar evaluative tendencies or shared method variance. A single source design is able to hold distinct constructs without generating any association due to response style or overall satisfaction with tutoring.

Structural collinearity

Table 7 gives a variance inflation factor of 1.000 to 1.468. Values are all less than three, which is the benchmark as in the workbook. The best values are on engagement and self-efficacy as co-predictors of perceived outcomes. The findings do not suggest that there was any problematic predictor collinearity in the specified equations. They hence encourage the interpretation of the estimated coefficients as different conditional associations between the constructs included.

The absence of low structural collinearity does not imply that there is no common method bias. It is also silent on the omixed variables, measurement inaccuracies other than the model or being selected to use AI tutoring. These problems relate to various sources of uncertainty. The collinearity test is instead limited to the stability aspect that it will specifically test and does not serve as a broad-based check on model validity.

Table 7: Structural Model Collinearity Assessment

Hypothesis	Predictor	Outcome	VIF
H1	AITQ	PU	1
H2	AITQ	MLE	1.305
H3	AITQ	MSE	1.255
H4	PU	MLE	1.305
H5	MLE	MSE	1.255
H6	MLE	PMLO	1.468
H7	MSE	PMLO	1.468

Note. All reported VIF values are below 3.

Direct relationships and hypotheses

All the seven directional relationships proposed are supported in the provided analysis in Table 8. There is a positive correlation between tutoring quality and usefulness, where the 2- ends of the confidence interval are 0.415 to 0.552 ($\beta = 0.483$). It is less associated with engagement and self-efficacy with $\beta = 0.245$ and $\beta = 0.205$ respectively. They also exclude zero in their intervals. These findings confirm H1, H2 and H3 and inform that the strongest direct relationship of quality involves perceived practicability.

Engagement is positively related to usefulness with 0.425 as the 0.425 value supports H4. Engagement is positively correlated with self-efficacy, and the correlations are 0.472, which favors H5. These two relationships are relatively narrow and both reporting intervals and p values are lower than 0.001. The trend follows an approach of connecting practical value, active involvement, and confidence. It does not prove that usefulness comes first before engagement and engagement comes before confidence on actual student experience.

Perceived outcomes are independently related to engagement and self-efficacy. These are 0.320 and 0.390 which favors H6 and H7 respectively. The self-efficacy coefficient is numerically greater, although the results provided do not present an actual test of the difference between these coefficients. It would be unjustified thus to say that there is statistically determined superiority

Table 8: Direct Relationships and Hypothesis Testing

Hypothesis	Path	β	t	Lower CI	Upper CI
H1	AITQ → PU	0.483	13.622	0.415	0.552
H2	AITQ → MLE	0.245	5.810	0.161	0.328
H3	AITQ → MSE	0.205	4.993	0.125	0.286
H4	PU → MLE	0.425	10.947	0.348	0.501
H5	MLE → MSE	0.472	12.505	0.395	0.543
H6	MLE → PMLO	0.320	7.353	0.232	0.402
H7	MSE → PMLO	0.390	9.177	0.308	0.476

Note. All seven hypotheses supported within the supplied analysis. All $p < 0.001$; 5,000 valid bootstrap resamples. Probability values are normal approximations.

of self-efficacy (Akın, 2022). The workbook provides the structural estimates and the values of the explained variability which are reproduced in figure 2.

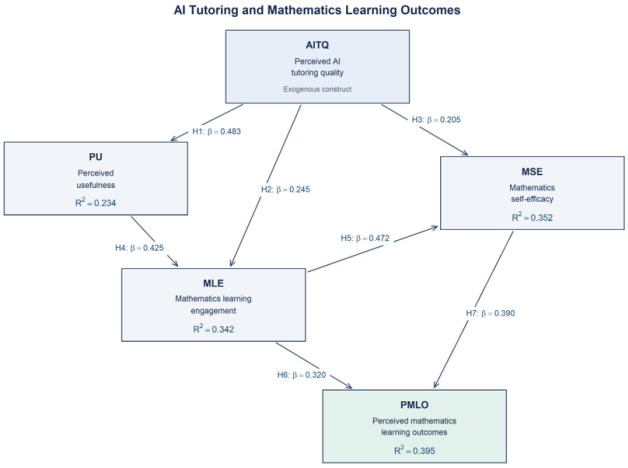


Figure 2: Structural model with standardised coefficients and explained variance

Specific indirect relationships

There are twelve individual indirect relationships with positive confidence intervals and reported in Table 9. The correlation between the quality of tutoring and usefulness on the one hand to engagement on the other hand is 0.206. The relationship between engagement to self-efficacy and perceived outcomes is 0.184. These findings are in line with the hypothetical mediating functions of usefulness and confidence. They also demonstrate the incompleteness of the implied associations of the model using pure coefficients.

A number of pathways link the quality of tutoring and perceived outcomes. The path via self-efficacy is 0.080, the path via engagement is 0.078. The path of utility and involvement is 0.066. Other sequential pathways via engagement and self-efficacy add 0.045 and 0.038, but this varies according to the presence of usefulness. They need to be interpreted in terms of their small individual contributions as well as their collective contribution. None give evidence of a specific mechanism having been experimentally isolated.

Indirect relationship is the term that is used here since the records of cross sections are not related by time. Also, the model does not have a direct tutoring quality to outcomes path. Its lack of magnitude in the estimated structure can be seen neither as a tested nonsignificant coefficient nor evidence of full mediation (Taha & Abdulrahman, 2023). The same warnings are usefulness and perceived outcomes. These differences maintain the distinction between a model constraint and an empirical analogy between alternative accounts.

Table 9: Specific Indirect Relationships

Pathway	Indirect β	SE	Lower CI	Upper CI
AITQ → PU → MLE	0.206	0.024	0.162	0.256
AITQ → PU → MLE → MSE	0.097	0.014	0.072	0.128
AITQ → MLE → MSE	0.116	0.021	0.075	0.158
AITQ → PU → MLE → MSE → PMLO	0.038	0.007	0.026	0.053
AITQ → PU → MLE → PMLO	0.066	0.012	0.044	0.091
AITQ → MLE → MSE → PMLO	0.045	0.010	0.027	0.067
AITQ → MLE → PMLO	0.078	0.017	0.048	0.113
AITQ → MSE → PMLO	0.080	0.018	0.047	0.118
PU → MLE → MSE	0.201	0.026	0.152	0.253
PU → MLE → MSE → PMLO	0.078	0.014	0.054	0.108
PU → MLE → PMLO	0.136	0.022	0.095	0.182
MLE → MSE → PMLO	0.184	0.026	0.136	0.237

Note. All intervals exclude zero; all p < 0.001; 5,000 valid bootstrap resamples. Arrows indicate modelled statistical pathways.

Total relationships

Table 10 is a combination of direct and specified relationships. Engagement is also associated with perceived outcomes, having a total association of 0.504 (it has a direct component of 0.320 and an indirect component of 0.184, through self-efficacy). The overall

Table 10: Total Relationships among the Study Constructs

From	To	Direct β	Indirect β	Total β	Lower CI	Upper CI
AITQ	PU	0.483	0	0.483	0.415	0.552
AITQ	MLE	0.245	0.206	0.451	0.377	0.524
AITQ	MSE	0.205	0.213	0.418	0.341	0.493
AITQ	PMLO	0	0.307	0.307	0.252	0.365
PU	MLE	0.425	0	0.425	0.348	0.501
PU	MSE	0	0.201	0.201	0.152	0.253
PU	PMLO	0	0.214	0.214	0.168	0.266
MLE	MSE	0.472	0	0.472	0.395	0.543
MLE	PMLO	0.320	0.184	0.504	0.438	0.569
MSE	PMLO	0.390	0	0.390	0.308	0.476

Note. All total relationship intervals exclude zero. Zero direct entries denote paths not specified, rather than estimated null relationships.

outcome of tutoring quality is 0.307, all in the pathways contained in the model. The overall outcome relationship of usefulness is 0.214. Another direct association of 0.390 between self-efficacy and outcomes is maintained since no additional mediator is given.

The overall correlation of tutoring quality and engagement is 0.451 as compared to its direct correlation of 0.245. Its overall correlation with self-efficacy stands at 0.418 as compared to 0.205 with a direct correlation. All these decompositions explain why it is crucial to take into account the linked model. They cannot be summed up among various predictors to generate a percentage of learning gain. Standardised path coefficients explain on a modelled scale relationships and are not percentage changes in achievement.

Explained variance

Table 11 lists usefulness of 0.234, engagement of 0.342, self-efficacy of 0.352 and perceived outcomes of 0.395. Adjusted values are slightly lower at 0.232, 0.339, 0.350, and 0.393. The low variations are due to the low number of predictors compared to the reported records. The result equation thus explains around 39.5 percent of the difference between the perceived outcome scores in the perceived outcome model provided.

The percentage of outcome variation is large and out of the model that is defined. Other factors, such as prior attainment, teacher support, home learning conditions, motivation, and other factors can also be relevant, but their effects are not estimated here. Explained variance is also an in-sample characteristic and cannot be accordinged to predictive accuracy in new schools (Tiwari, 2023). The prediction evaluation below offers a distinct view of how models perform, but it is conditional upon the provenance and application of the analysis provided.

Table 11: Explained Variance of the Endogenous Constructs

Outcome	Predictors	R^2	Adjusted R^2
PU	AITQ	0.234	0.232
MLE	AITQ, PU	0.342	0.339
MSE	AITQ, MLE	0.352	0.350
PMLO	MLE, MSE	0.395	0.393

Note. N = 470. One predictor for PU and two predictors for each remaining endogenous construct.

Structural effect sizes

The report of f squared values starts at table 12. The value of tutoring quality to usefulness has the highest value of 0.305 with engagement to self-efficacy coming

Table 12: Effect Sizes of the Structural Relationships

Hypothesis	Path	R^2 included	R^2 excluded	f^2	Size
H1	AITQ → PU	0.234	0	0.305	Medium
H2	AITQ → MLE	0.342	0.296	0.070	Small
H3	AITQ → MSE	0.352	0.319	0.052	Small
H4	PU → MLE	0.342	0.203	0.210	Medium
H5	MLE → MSE	0.352	0.175	0.273	Medium
H6	MLE → PMLO	0.395	0.326	0.115	Small
H7	MSE → PMLO	0.395	0.292	0.171	Medium

Note. Effect size classifications follow the supplied workbook.

next at 0.273. The usefulness to engagement 0.210 and the self-efficacy to perceived outcomes = 0.171. The workbook classifies these four relationships as medium. There are small classifications of tutoring quality to engagement, tutoring quality to self-efficacy, and engagement to perceived outcomes with values of 0.070, 0.052, and 0.115 respectively.

The uniformity of statistical support speaks in favor of hypotheses based on the effect size results. A relationship can be accurately estimated and adds a relatively small amount of extra explained variance. This can be seen in respect to the immediate quality relationships about involvement and trust. In contrast, engagement maintains a significant role by both its resultant relationship and linking with self-efficacy. The contributions of models reported do not relate to proposed intervention effects, but are the description of educational intervention effects.

Predictive performance

Table 13 presents all positive values of predictive relevance in all twenty indicators endogenous. Values range from 0.051 to 0.177. The PLS root mean squared error is less than the linear benchmark of nineteen indicators. Exception is PMLO5 where the PLS error is 1.128 as compared to 1.125 in the linear model. In the other indicators, the advantages are not numerically high, as they are within the range of between 0.005 to 0.011. The outcomes justify small comparative prediction in the stated process.

There is positive predictive relevance in the four nonexceptional perceived outcome indicators but comparatively low numerical gains in comparison to the linear benchmark. This trend is consistent with convenient explanatory construction with weak individual forecasting (Azevedo & Jaques, 2020). The positive balance of indicators is not to hide absolute error magnitudes and the minor difference between approaches. No external school sample is provided to be independent and repeated cross validation fails to provide transportability to cross educational settings or platforms.

Table 13: Out-of-Sample Predictive Performance

Indicator	Q^2 predict	PLS RMSE	LM RMSE	Difference
PU1	0.160	1.098	1.104	-0.006
PU2	0.177	1.090	1.101	-0.011
PU3	0.129	1.097	1.102	-0.005
PU4	0.168	1.090	1.097	-0.007
PU5	0.150	1.113	1.121	-0.008
MLE1	0.138	1.135	1.145	-0.010
MLE2	0.114	1.154	1.160	-0.007
MLE3	0.137	1.135	1.144	-0.009
MLE4	0.137	1.099	1.106	-0.008
MLE5	0.165	1.123	1.131	-0.008
MSE1	0.075	1.170	1.176	-0.006
MSE2	0.146	1.089	1.097	-0.007
MSE3	0.120	1.098	1.108	-0.010
MSE4	0.120	1.113	1.123	-0.010
MSE5	0.111	1.157	1.168	-0.011
PMLO1	0.051	1.162	1.170	-0.008
PMLO2	0.055	1.109	1.120	-0.011
PMLO3	0.081	1.124	1.132	-0.008
PMLO4	0.057	1.176	1.185	-0.009
PMLO5	0.094	1.128	1.125	0.002

Note. Ten folds and ten repetitions; 4,700 held out predictions per indicator. Negative differences favour PLS. All Q^2 predict values are positive.

Supplementary model fit

Table 14 documents that saturated and estimated model SRMR, d ULS and d G were not estimated. These indices do not thus draw any global fit conclusion. Their lack does not eliminate the reported reliability, structural or predictive findings, but restricts the scope of model evaluation possible. The omission is explicitly reported in the manuscript instead of replacing the presumed acceptable values. An analytical archive would be needed to find out whether further fit diagnostics can be computed in a suitable manner.

Table 14: Supplementary Model Fit Assessment

Statistic	Status
SRMR: saturated model	Not estimated
SRMR: estimated model	Not estimated
d_ULS with bootstrap reference	Not estimated
d_G with bootstrap reference	Not estimated

Note. The supplied workbook does not provide these indices or their bootstrap reference assessments.

Discussion

Instructional quality and practical value

The greatest direct relationship that ensues out of tutoring quality deals with usefulness. This tendency demonstrates that explicit, sensible, and suitable teaching is most directly expressed in appraisals of usefulness in the provided files. This interpretation is in line with the content of the item as it relates the quality to explain and feedback and utility to study goals and effective learning (García-Martínez *et al.*, 2023). The finding endorses the interpretation of instructional quality as the experience of education, and not as an object of a general attitude toward technological advancedness.

The lower direct correlations with engagement and self-efficacy suggest that the favourable instructional evaluations cannot explain these constructs in the same level as the other stated predictors are used. Such a difference is important to system evaluation. By not always putting in effort or building up confidence, a learner can realise that a resource has good explanations (Samuelsson, 2023). The model then focuses on the context in which the perceived quality is coupled with participation and capability beliefs.

Engagement as a connecting process

Engagement is in the middle since it has both quality and usefulness as its associations and has relations with self-efficacy and perceived outcomes. Its overall outcome association value is bigger than its direct coefficient due to the additional covariance added by the given pathway of interaction (Kochmar *et al.*, 2020). This finding promotes the process-oriented interpretation of tutoring. Whether students approve of the resource or not is not the question of relevance, but an attention, effort, persistence, and independent attempts are the questions that are relevant to learning activities.

The engagement indicators also offer a practical definition between positive participation and dependency on answers. Asking a solution to be given to a problem before one has seen it is not the same as asking it to be done immediately. The two behaviours can be experienced in the same length of time. Usage frequency and duration, therefore, cannot be used as a replacement of engagement measure (Zhang *et al.*, 2023). The descriptive exposure variables provide a context, and the latent engagement construct provides a more specific appraisal of the way in which the students engage in the learning process of mathematics.

Confidence and perceived progress

After the engagement is added, self-efficacy is positively directly related to perceived outcomes. This result suggests that confidence adds information other than reported participation in the given model. Students reporting a high level of confidence are also likely to report better results. The correlation is educationalistically viable but should be interpreted with care of the subjectivity of both constructs. A widespread positive evaluation of self-learning may help to associate it (Marrone *et al.*, 2022).

The instructive power of the questionnaire to take independent performance into consideration is thus useful. It tries to keep personal capability and trust in the system apart. Yet, due to the lack of full self-efficacy and outcome wording, it is not possible to fully investigate whether this distinction is upheld in all indicators. A self-evaluation would also help in making an objective evaluation on whether confidence is being calibrated to competence (Kumar *et al.*, 2023). In the absence of evidence, the current findings relate to perceived ability and perceived improvement as opposed to mastery that has been proven.

Explanatory strength and predictive restraint

Outcome equation records a significant yet unaccounted portion of variance in the reported scores. This contributes to the applicability of engagement and self-efficacy in the model and creates a wide scope of other impacts. The anticipatory results introduce a qualification of their own. The benefits of using PLS are minimal even though the regression method is superior to the linear method in most indicators (Kshirsagar *et al.*, 2022). The body of evidence leans towards a cautious approach that acknowledges consistent associations without purporting to outstanding predictive behavior.

The difference is particularly significant in making educational choices about individual students. The description of average relations could not be relevant to determine who should be supported or who can become better using a tutor. Prediction errors and lack of external validation reported do not warrant personal classification or automatic allocation of learning opportunities (Hossein-Mohand & Hossein-Mohand, 2023). The results are more appropriate in clarifying research questions and assessment criteria than the high consequences decisions of the learners.

Educational and Research Implications

Instructional evaluation is expected to find out the clarity of explanation, logical arrangement of solution steps and quality of error feedback. These scales relate directly to quality construct. An evaluation activity that might be beneficial involves making a tutor explain some

common misconception, and the learner completes the related problem independently. This would produce evidence regarding the instructional support and transfer. The current affiliations encourage the given focus of the evaluation but do not justify any specific product.

The learning activities should also maintain a chance to reason activities and then fully answer the students. Graduated hints, requests for explanation, and subsequent unaided practice are possible design options (Yue *et al.*, 2022). Their efficacy would have to be directly measured. The framework indicates that questions of engagement and confidence should be measured independently since positive scores in one dimension do not predict positive scores in the other. Such segregation would assist teachers to locate students who are engaged but are confused about their potential.

The distributions of devices and location indicate issues of accessibility. The use of Smartphones and home-based activity are dominant in the records provided, yet these cannot define the requirements of a larger group of people. Future field studies might look into how small screens, lack of trustworthy connection, or adult inability to help, affect the way explanations are understood. The comparisons should take into consideration the disparities in the previous achievements and resource access. Otherwise, the apparent differences of the platforms might be in part due to the unequal conditions of learning.

Perceived usefulness, engagement, confidence and achievement should be differentiated in research reporting at its beginning. The way constructs are defined would be clearer, thereby minimizing the chances of describing satisfaction as learning improvement (Astuti, 2022). A chain of permission, consent, and data protection procedures also need to be documented in the case of studying underage individuals. In the case at hand, the fieldwork archive and the recovered questionnaire is a precondition of assessing the evidence as an empirical student study. Uncertainty over the origin of observations cannot be addressed by statistical completeness.

CONCLUSION

The provided PLS SEM outcomes confirm a consistent pattern of the positive relationship between the quality of perceived AI tutoring, its usefulness, the engagement in mathematics learning, self-efficacy, and the perceived outcomes. The confidence interval of all seven suggested direct paths does not contain zero. Numerical reliability, convergence and construct distinction are supported by measurement statistics. Engagement and self-efficacy combine to explain 39.5 percent of the variance in perceived outcomes and the reported procedure of prediction demonstrates slight benefits on a linear gauge of most indicators.

The substantive one is limited by the evidence. The analysis deals with the perceived results and is unable to make causal profits in mathematics performance. The absence of independent verification of an empirical student sample because of hypothetical labeling of tutors, partially worded questions in the questionnaire, and lack of documentation of fieldwork are stopping factors. According to these boundaries, the framework notes the instructional quality, active engagement, and independent confidence as valuable goals to further research. The findings have to be supported with verifiable data provenance and objective evaluation before it can be claimed that it was effective in education.

REFERENCES

Akın, A. (2022). The effectiveness of web-based Mathematics instruction (WBMI) on K-16 students' mathematics learning: a meta-analytic research. *Education and Information Technologies, 27*(6), 8015–8040. https://doi.org/10.1007/s10639-022-10931-x

Almusaed, A., Almssad, A., Yitmen, I., & Homod, R. Z. (2023). Enhancing student engagement: harnessing "aied"'s power in hybrid education—a review analysis. *Education Sciences, 13*(7), 632. https://doi.org/10.3390/educsci13070632

Alomari, M. A., & Jabr, M. O. (2020). The effect of the use of an educational software based on the strategy of artificial intelligence on students' achievement and their attitudes towards it. *Management Science Letters,* 2951–2960. https://doi.org/10.5267/j.msl.2020.5.030

Astuti, N. (2022). Pengaruh Metode Pembelajaran Tutor Sebaya Terhadap Hasil Belajar Matematika Siswa Kelas III Sdn Simpang Warga 1 Kecamatan Aluh-Aluh. *Jurnal Terapung : Ilmu - Ilmu Sosial, 4*(1), 33. https://doi.org/10.31602/jt.v4i1.7357

Azevedo, O. B., & Jaques, P. A. (2020). Umstudent engagement: harnessing "aied"'s power in hybrid education—a review analysis. *Anais Do XXXI Simpósio Brasileiro de Informática Na Educação (SBIE 2020*), 762–771. https://doi.org/10.5753/cbie.sbie.2020.762

ÇetiN, İ., ErümiT, A. K., NabiYev, V., Karal, H., Kösa, T., & Kokoç, M. (2023). Theeffect of gamified adaptive intelligent tutoring system artibos on problem-solving skills. *Participatory Educational Research, 10*(1), 344–374. https://doi.org/10.17275/per.23.19.10.1

Chai, C. S., Chiu, T. K. F., Wang, X., Jiang, F., & Lin, X.-F. (2022). Modeling chinese secondary school students' behavioral intentions to learn Artificial intelligence with the theory of planned behavior and self-determination theory. *Sustainability, 15*(1), 605. https://doi.org/10.3390/su15010605

Cunska, A. (2021). Important Factors For More Effective Learning of Mathematics. Society. Integration. Education. *Proceedings of the International Scientific Conference, 2*, 139–150. https://doi.org/10.17770/sie2021vol2.6466

Cunska, A. (2022). Prototype of project AI4Math: interdisciplinary and innovative technology for accelerated learning of mathematics. *Wseas Transactions on Business and Economics, 19,* 1839–1848. https://doi.org/10.37394/23207.2022.19.165

Fitzmaurice, O., O'Meara, N., & Johnson, P. (2021). Highlighting therelevance of mathematics to secondary school students – why and how. *European Journal of STEM Education, 6*(1), 07. https://doi.org/10.20897/ejsteme/10895

Fornons, V., Palau, R., & Santiago, R. (2021). Secondary school students' perception according to their learning style of a mathematics Flipped Classroom. *Journal of Technology and Science Education, 11*(2), 227. https://doi.org/10.3926/jotse.1092

García-Martínez, I., Fernández-Batanero, J. M., Fernández-Cerero, J., & León, S. P. (2023). Analysing the impact of artificial intelligence and computational sciences on student performance: systematic review and meta-analysis. *Journal of New Approaches in Educational Research, 12*(1), 171–197. https://doi.org/10.7821/naer.2023.1.1240

Hossein-Mohand, H., & Hossein-Mohand, H. (2023). Influence of motivation on the perception of mathematics by secondary school students. *Frontiers in Psychology, 13*, 1111600. https://doi.org/10.3389/fpsyg.2022.1111600

Kochmar, E., Vu, D. D., Belfer, R., Gupta, V., Serban, I. V., & Pineau, J. (2020). Automated impact of artificial intelligence and computational sciences on student performance: systematic review and meta-analysis. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Artificial Intelligence in Education* (Vol. 12164, pp. 140–146). Springer International Publishing. https://doi.org/10.1007/978-3-030-52240-7_26

Kshirsagar, P. R., Jagannadham, D. B. V., Alqahtani, H., Noorulhasan Naveed, Q., Islam, S., Thangamani, M., & Dejene, M. (2022). Human intelligence analysis through perception of ai in teaching and learning. *Computational Intelligence and Neuroscience, 2022,* 1–9. https://doi.org/10.1155/2022/9160727

Kumar, D., Haque, A., Mishra, K., Islam, F., Kumar Mishra, B., & Ahmad, S. (2023). Exploring the transformative role of artificial intelligence and metaverse in education: a comprehensive review. *Metaverse Basic and Applied Research, 2*, 55. https://doi.org/10.56294/mr202355

Lopes, P. C. D. C., Abreu, V. R. P. D., Streit, J. A. C., Gameiro, S., & Nunes, P. H. S. (2023). Machine learning na aprendizagem das ciências exatas: análise qualitativa para ensino fundamental e médio. *Revista de Tecnologia & Gestão Sustentável, 2*(5). https://doi.org/10.17271/rtgs.v2i5.3710

Marrone, R., Taddeo, V., & Hill, G. (2022). Creativity and artificial intelligence—a student perspective. *Journal of Intelligence, 10*(3), 65. https://doi.org/10.3390/jintelligence10030065

Moliner, L., Alegre, F., & Lorenzo-Valentín, G. (2022). Peer tutoring and math digital tools: a promising combination in middle school. *Mathematics, 10*(13), 2360. https://doi.org/10.3390/math10132360

Neo, M., Lee, C. P., Tan, H. Y.-J., Neo, T. K., Tan, Y. X., Mahendru, N., & Ismat, Z. (2022). Enhancing students' online learning experiences with artificial intelligence (AI): the merlin project. *International Journal of Technology, 13*(5), 1023. https://doi.org/10.14716/ijtech.v13i5.5843

Radović, S. (2024). Is it only about technology? the interplay between educational technology for mathematics homework, teaching practice, and students' activities. *Journal of Computers in Education, 11*(3), 743–762. https://doi.org/10.1007/s40692-023-00277-9

Samuelsson, J. (2023). Arithmetic fact fluency supported by Artificial Intelligence. *Frontiers in Education Technology, 6*(1), p13. https://doi.org/10.22158/fet.v6n1p13

Soesanto, R. H., Dirgantoro, K. P. S., & Priyanti, N. (2022). Indonesian students' perceptions towards AI-based learning in mathematics. *Journal on Mathematics Education, 13*(3), 531–548. https://doi.org/10.22342/jme.v13i3.pp531-548

Taha, T. B., & Abdulrahman, M. S. (2023). Theimpact of technology on students' Psychological and Educational Performance. *JISA(Jurnal Informatika Dan Sains), 6*(1), 91–95. https://doi.org/10.31326/jisa.v6i1.1661

Tiwari, R. (2023). The integration of AI and machine learning in education and its potential to personalize and improve student learning experiences. *Interantional Journal Of Scientific Research In Engineering And Management, 07*(02). https://doi.org/10.55041/IJSREM17645

Utterberg Modén, M., Tallvid, M., Lundin, J., & Lindström, B. (2021). *Intelligent Tutoring Systems: Why Teachers Abandoned a Technology Aimed at Automating Teaching Processes*. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2021.186

Wang, H., Tlili, A., Huang, R., Cai, Z., Li, M., Cheng, Z., Yang, D., Li, M., Zhu, X., & Fei, C. (2023). Examining the applications of intelligent tutoring systems in real educational contexts: a systematic literature review from the social experiment perspective. *Education and Information Technologies, 28*(7), 9113–9148. https://doi.org/10.1007/s10639-022-11555-x

Weinhandl, R., Mayerhofer, M., Houghton, T., Lavicza, Z., Eichmair, M., & Hohenwarter, M. (2022). Personas characterising secondary school mathematics students: development and applications to educational technology. *Education Sciences, 12*(7), 447. https://doi.org/10.3390/educsci12070447

Yue, M., Jong, M. S.-Y., & Dai, Y. (2022). Pedagogical design of K-12 artificial intelligence education: a systematic review. *Sustainability, 14*(23), 15620. https://doi.org/10.3390/su142315620

Zhang, W., Xiao, S., & Fu, W. (2023). Can "Smart Homework" achieve the goal of chinese "Double Reduction" policy to reduce burden and improve quality?—the positive and negative effects of "Smart Homework" on students. *Sustainability, 15*(12), 9759. https://doi.org/10.3390/su15129759