

**RESEARCH ARTICLE****Perceived Desirable Difficulty in AI Tutoring: Engagement, Retention Confidence, and Continued Use****Sumit Khanal<sup>1</sup> and Abhijeet Sen Gupta<sup>1</sup>**<sup>1</sup> Kingsford Institute of Higher Education, Australia

Sumit.Khanal@kihe.nsw.edu.au | abhijeet.sengupta@Kihe.nsw.edu.au

**ARTICLE INFORMATION****Published:** August 2026**Volume/Issue:** Vol. 2, No. 1**KEYWORDS**

AI tutoring; perceived desirable difficulty; learner engagement; intrinsic motivation; knowledge retention confidence; continued use intention; PLS-SEM.

**ABSTRACT**

This study examines the role of perceived desirable difficulty in AI tutoring systems and explains how productive learning challenge influences learner engagement, productive struggle orientation, intrinsic motivation, knowledge retention confidence and continued use intention. Although AI tutoring is increasingly used in higher education, much of the existing discussion mainly focuses on usefulness, satisfaction and general acceptance. Due to this reason, less attention has been given to whether students accept AI-supported difficulty as a meaningful learning condition. To address this gap, the study developed a structural model by integrating the Desirable Difficulties Framework, Self-Determination Theory and technology continuance logic. A quantitative cross-sectional survey design was applied, and data were collected from 358 university students who had experience using AI tutoring or AI-assisted learning tools. The proposed model was tested through Partial Least Squares Structural Equation Modelling using SmartPLS 4. The findings show that perceived desirable difficulty is the main psychological and learning driver in the model. It significantly influenced learner engagement, productive struggle orientation, intrinsic motivation, knowledge retention confidence and continued use intention. The result also shows that intrinsic motivation, learner engagement and productive struggle orientation significantly improved knowledge retention confidence, which later influenced continued use intention. Moreover, perceived AI adaptiveness significantly moderated the relationship between perceived desirable difficulty and continued use intention. From the above findings, it can be understood that difficulty in AI tutoring is not harmful when it is productive, adaptive and psychologically meaningful. Rather, manageable challenge can support deeper engagement, stronger motivation, better retention confidence and sustained AI tutoring use.

## 1. INTRODUCTION

### 1.1 Background of AI Tutoring and Adaptive Learning

In the context of higher education, artificial intelligence-based learning systems have created a significant transformation in the teaching and learning environment. One of the main reasons behind this transformation is that AI is not now only considered as a supportive digital tool for classroom activities. In the earlier stage of digital education, technology was mainly used for distributing lecture materials, managing online classes and supporting basic assessment-related activities. However, with the present development of AI tutoring systems, this understanding has changed in a significant manner. AI tutoring systems can now provide personalised explanation, real-time feedback, step-by-step academic guidance and adaptive learning support based on the individual needs of learners. From the above situation, it can be understood that AI tutoring is not only connected with academic convenience. Rather, it has a strong pedagogical support role within the higher education learning process.

Recent empirical findings also support the argument that AI tutoring can develop meaningful learning benefits when it is designed through proper instructional planning. Kestin et al. (2025) found that students who used an AI tutor achieved better learning outcomes than students in an active learning classroom, and also reported higher engagement and motivation. The result shows that AI tutoring may influence both cognitive learning and motivational learning at the same time. Similarly, Goldstone (2026) explains that AI-powered instruction is becoming more visible in the 2026 classroom because it can provide personalised support, identify learning gaps and assist teachers in managing different learner needs. From the above findings, it can be said that AI tutoring is becoming a part of a broader educational transformation, where learning support is no longer limited only to fixed classroom interaction or teacher-centred explanation.

Adaptive learning is one of the important mechanisms through which this transformation becomes more visible. In the adaptive learning condition, the system changes content, difficulty level, feedback and learning pace according to the learner's level of understanding. This creates a more individualised form of academic support because students do not learn at the same speed and also do not respond to learning difficulty in the same way. Li et al. (2025) also found that students' motivation to learn with AI can be understood through different motivational dimensions. This reflects that AI use in education is not only technological in nature, but also psychological in nature. For this reason, the present study examines AI tutoring as a personalised learning environment where perceived difficulty, engagement, motivation and continued use are connected within one learning relationship nexus.

### 1.2 Research Problem and Study Gap

Although AI tutoring use in higher education is increasing, the existing research discussion is still largely concentrated on learning improvement, perceived usefulness, motivation, satisfaction and general technology acceptance. These areas are important because AI tutoring cannot become sustainable in the educational environment if students do not perceive value from using this system. However, this kind of research direction creates only a partial understanding of AI-supported learning. It explains whether students receive benefit from AI tutoring and whether they are motivated to use this technology. However, it does not properly explain how students respond when AI tutoring creates productive difficulty inside the learning process. From this above condition, it can be understood that the present literature still gives less attention to the difficulty-based learning relationship nexus in AI tutoring.

This gap is important because meaningful learning is not always developed through easy interaction or immediate academic support. In many educational situations,

learning becomes stronger when students face manageable challenge, apply cognitive effort and remain engaged with difficult learning materials. However, AI tools are often discussed as systems that reduce effort, simplify tasks and provide quick academic solutions. In this regard, a possible tension can be seen in AI-mediated education. If AI tutoring only makes learning faster and easier, students may use it mainly for convenience rather than for developing deeper understanding. Therefore, the important issue is not only whether AI tutoring improves learning performance. Rather, it is also whether students accept, interpret and value the difficulty that may be necessary for stronger knowledge retention.

From the above research condition, it can be said that perceived desirable difficulty is still an underdeveloped area in the AI tutoring literature. Kestin et al. (2025) show the learning and engagement potential of AI tutoring, while Li et al. (2025) provide an important motivational understanding of AI-based learning. However, limited attention has been found on the psychological process through which students interpret challenging AI tutoring as useful, engaging and worth continuing. In this particular study, difficulty is not considered as a barrier to learning. Rather, it is examined as a possible learning-supportive condition when it creates productive struggle, intrinsic motivation, learner engagement and confidence in knowledge retention. Therefore, this study addresses this gap by examining how perceived desirable difficulty shapes learning and continuance outcomes in AI tutoring, and how these elements are connected within one learning relationship nexus.

### **1.3 Aim, Objectives, and Research Questions**

The purpose of this study is to examine the role of perceived desirable difficulty in the context of AI tutoring. More specifically, this study attempts to explain how perceived desirable difficulty influences learner engagement, productive struggle orientation, intrinsic motivation, knowledge retention

confidence and continued use intention. In the context of this particular study, desirable difficulty is understood as a learning condition where the AI tutoring system creates a certain level of academic challenge that requires effort from students, but at the same time supports their understanding and learning progress. Therefore, this study does not assume that all forms of difficulty are useful for learning. Rather, the study focuses on the type of difficulty that students perceive as productive, manageable and connected with learning improvement. From this above position, it can be said that perceived desirable difficulty is treated as a learning-supportive condition, not as a learning barrier.

Based on this purpose, four research questions have been developed for this study. First, the study examines how perceived desirable difficulty in AI tutoring systems directly influences learner engagement, knowledge retention confidence and continued use intention among university students. Second, the study investigates whether productive struggle orientation and intrinsic motivation work as psychological mechanisms between perceived desirable difficulty and knowledge retention confidence. Third, the study examines whether the pathway from learner engagement to knowledge retention confidence and continued use intention can explain the sustained use of AI tutoring systems. Fourth, the study also examines whether perceived AI adaptiveness strengthens the relationship between perceived desirable difficulty and continued use intention. From the above research questions, it can be understood that this study does not only observe the direct effect of AI tutoring difficulty. Rather, it tries to identify the psychological and learning relationship nexus through which difficulty becomes meaningful for students.

These research questions are developed from the argument that AI tutoring cannot be understood only through usefulness, satisfaction or general technology acceptance. These factors are important, but

they do not fully explain how students experience challenge during AI-supported learning. Rather, the learning value of AI tutoring may depend on how students interpret difficulty, how they remain engaged with that difficulty and whether the system appears adaptive enough to make the challenge acceptable. For this reason, the present study develops a structural model where learning difficulty, psychological response, retention confidence and continued use are examined together. This reflects that AI tutoring is not only a technological support system. It is also a learning environment where productive challenge and adaptive support may jointly shape students' future learning behaviour.

#### **1.4 Research Contribution and Structure of the Paper**

This study contributes to the existing literature in three specific ways. First, this study introduces perceived desirable difficulty as a central construct in the AI tutoring context. Most of the previous studies in AI education explain tutoring systems through learning improvement, motivation, satisfaction and acceptance of technology. However, this study argues that the real pedagogical value of AI tutoring may also depend on whether the system can create a productive level of learning challenge for students. From this point of view, AI tutoring should not be considered only as a system that gives quick answers or immediate academic support. Rather, it can be understood as a system that may develop deeper learning through manageable cognitive effort. For this reason, perceived desirable difficulty creates an important new direction for understanding how AI tutoring supports the learning process.

Second, this study explains the psychological relationship nexus through which perceived desirable difficulty can influence learning outcomes. In this regard, learner engagement, productive struggle orientation and intrinsic motivation are used as important psychological mechanisms that connect difficulty with knowledge retention

confidence. This contribution is important because it shows how difficulty may become educationally meaningful instead of becoming frustrating for students. If students understand difficulty as productive and manageable, then they may become more engaged and motivated in the learning process. Knowledge retention confidence is also important in this study because it reflects whether students believe that AI-supported learning can be remembered and applied later. From the above arguments, it can be said that this study does not only examine whether difficulty exists in AI tutoring. Rather, it examines how this difficulty works through psychological and learning-related pathways.

Third, this study examines perceived AI adaptiveness as a boundary condition for continued use intention. The reason behind this is that difficulty may not always create positive outcomes for learners. If the AI tutoring system appears rigid, poorly matched or not suitable with the learner's level of understanding, then difficulty may reduce the willingness to continue using the system. However, when the system is perceived as adaptive, the same difficulty may be understood as personalised, manageable and helpful for learning improvement. This reflects that the effect of perceived desirable difficulty may depend on the adaptive capacity of the AI tutoring system. The rest of the paper is organised in the following way. The next section reviews the relevant literature and develops the hypotheses. The following section explains the methodology of the study. After that, the data analysis and findings are presented. Finally, the paper discusses the findings, implications, limitations and conclusion.

## **2. LITERATURE REVIEW**

### **2.1 Theoretical Foundation**

The theoretical foundation of this study is developed by integrating the Desirable Difficulties Framework, Self-Determination

Theory and technology acceptance logic. The Desirable Difficulties Framework explains that some learning conditions may look more difficult during the early stage of learning, but they can create stronger long-term learning when learners are required to process, retrieve and reorganise information in a more active way (Bjork, 1994). In the context of AI tutoring, this framework is important because AI-based learning should not be understood only as a system that reduces student effort. Rather, it can also be understood as a system that creates productive cognitive effort when the level of challenge is properly managed.

Self-Determination Theory provides the motivational base of this study. Ryan and Deci (2000) explain that intrinsic motivation and psychological growth are supported when learners experience autonomy, competence and relatedness. This is relevant in the context of AI tutoring because students may respond more positively to challenging learning tasks when they feel that the challenge is manageable and also supportive for developing competence. Therefore, perceived desirable difficulty may not create a positive learning outcome in every situation. Rather, it needs to be understood through the motivational condition of the learner.

Technology acceptance logic adds the behavioural intention part of the model. Davis (1989) argued that technology use is strongly influenced by perceived usefulness and ease of use. However, in the case of AI tutoring, continued use may not be shaped by ease alone. If the system creates useful challenge, supports learning confidence and appears adaptive, students may still develop the intention to continue using it. Therefore, the integration of these three theoretical positions helps this study to explain the full relationship nexus between learning difficulty, motivation, retention confidence and continued use intention.

## **2.2 AI Tutoring Systems and Adaptive Learning**

AI tutoring systems are becoming a visible part of higher education because they can provide learning support in a more flexible way than traditional digital learning platforms. In the earlier stage, educational technologies were mostly used for accessing materials, submitting tasks online and managing classroom activities. However, AI tutoring systems are now able to respond to student questions, provide explanations, create practice opportunities and give feedback during the learning process. Due to this reason, AI tutoring is no longer only a supportive technology. Rather, it can be understood as a more active learning partner within higher education.

Kestin et al. (2025) found that students who used an AI tutor showed stronger learning outcomes and also reported higher engagement and motivation compared with students in an active learning classroom. This result indicates that AI tutoring can influence both academic achievement and learning experience when it is developed with suitable pedagogical design. On a similar note, Goldstone (2026) explains that AI-powered instruction is becoming more important because it can identify knowledge gaps and support individual learner needs in real time. From these findings, it can be understood that AI tutoring is not only connected with academic convenience. Rather, it is also connected with personalised learning support.

Adaptive learning is the main mechanism through which this personalisation can become possible. In an adaptive system, the learning content, feedback and level of difficulty can be adjusted according to the present condition of the learner. Li et al. (2025) also show that students' motivation to learn with AI includes different forms of motivational regulation. This means that AI learning is not only technological, but also psychological in nature. Therefore, in this study, AI tutoring is examined as an adaptive learning environment where personalised

support, learning challenge and student motivation are structurally connected.

### **2.3 Desirable Difficulties and Productive Struggle**

The concept of desirable difficulty provides the central logic of this study. Bjork and Bjork (2020) explain that learning conditions which create effort during learning may improve long-term retention when they support deeper encoding and retrieval. This means that difficulty is not always harmful for learning. Some forms of difficulty can become useful when they make learners think more actively and engage more seriously with the learning content. Cepeda et al. (2006) also showed through their review of distributed practice that spaced learning conditions can improve memory performance compared with massed learning. Therefore, the learning process may become stronger when students are required to invest cognitive effort over time.

In the context of AI tutoring, this idea becomes more important because AI systems can either reduce effort too much or create a more balanced form of challenge. Structural Learning (2025) discusses desirable difficulty as a teaching principle where learning is strengthened through strategies such as spacing, retrieval and interleaving. At the same time, Chung et al. (2026) suggest that personalised AI tutors can become more effective when they guide learning proactively rather than only giving immediate answers. From this point, it can be said that adaptive AI tutoring should not remove all forms of struggle. Rather, it should preserve the level of struggle that supports learning.

Productive struggle is closely related to this argument. Young et al. (2024) show that productive struggle has become an important learning concept because it refers to effortful engagement with challenging tasks. The Bellwether report also raises an important question about AI in education: how much cognitive effort should AI reduce, and how much should it preserve for learning development (Kulesa, 2025). Baule (2025) similarly argues that failure and struggle can

make learning more meaningful when they are used carefully. Therefore, perceived desirable difficulty and productive struggle orientation are connected in this study because both explain how students may treat challenge as a learning opportunity rather than as a barrier.

### **2.4 Intrinsic Motivation, Learner Engagement, and Retention Confidence**

Intrinsic motivation and learner engagement are important psychological mechanisms for explaining why difficulty can become useful in AI tutoring. According to Deci and Ryan (2000), motivation becomes stronger when the learning environment supports autonomy, competence and meaningful involvement. Ryan and Deci (2020) also explain that intrinsic motivation is not only connected with interest. It is also connected with the internal feeling that the activity has value and that the learner is able to grow through the activity. From this theoretical position, it can be understood that difficulty does not automatically reduce motivation. Rather, if the difficulty helps learners feel competent and active, it can support stronger internal motivation.

In the context of AI-based learning, this issue becomes more important because AI tools can either create passive dependency or active learning involvement. Galindo-Domínguez et al. (2025) found that artificial intelligence can support learning motivation when it is applied through a structured educational intervention based on self-determination logic. Similarly, Yang et al. (2025) showed that AI dependence may influence students' innovation capability through cognitive and contextual mechanisms. This indicates that AI use is not always positive or negative by itself. The effect depends on how students psychologically process the AI-supported learning experience.

Learner engagement also works as a central factor in this relationship. Almufarreh (2024) found that students' satisfaction with AI tools is influenced by factors such as perceived utility and learning-related

experience. Annamalai et al. (2025) also showed that students' continuous use of ChatGPT in higher education is shaped by motivational conditions such as autonomy, competence and relatedness. Moreover, Katalinic et al. (2026) reviewed AI in online education and showed that AI can support learner engagement and satisfaction when it is properly integrated into learning environments.

From the above literature, it can be said that perceived desirable difficulty may become useful when it activates engagement and motivation rather than creating frustration. In this particular study, knowledge retention confidence is used as an important outcome because it reflects the learner's belief that AI-supported learning can be remembered and used later. Therefore, when difficulty produces active engagement, productive effort and intrinsic motivation, it may also strengthen confidence in knowledge retention. This creates a meaningful relationship nexus between challenge, psychological involvement and learning confidence.

### **2.5 Perceived AI Adaptiveness and Continued Use Intention**

Perceived AI adaptiveness refers to the learner's belief that the AI tutoring system can adjust its explanation, feedback, pace and difficulty according to the learner's level of understanding. This construct is important because difficulty is not always accepted positively by students. If a task is difficult but the system does not appear responsive, then the difficulty may be interpreted as pressure or poor system design. However, if the same difficulty is presented through an adaptive system, students may understand it as a personalised learning challenge.

Recent studies show that adaptive learning technologies can improve learning experience when they are used to provide more personalised educational support. Contrino et al. (2024) found that an adaptive learning tool improved student performance and satisfaction in both online and face-to-face learning environments. Demartini et al.

(2024) also discussed how artificial intelligence and learning analytics can improve adaptive learning by supporting educational decision-making and identifying learner needs. Similarly, Strielkowski et al. (2025) argued that AI-driven adaptive learning can make education more personalised, accessible and efficient.

In the user experience context, Jamali et al. (2025) found that adaptive learning features influence engagement, motivation and task completion. Hidayat-ur-Rehman (2024) also showed that AI competence, chatbot use and perceived autonomy are important drivers of student engagement in informal digital learning. These studies suggest that students are more likely to engage with AI when they feel that the system supports their autonomy and learning condition.

Continued use intention is also shaped by acceptance, usefulness and learning value. Wang (2025), Dahri et al. (2024), and Fang and Feng (2026) all indicate that students' sustained use of generative AI or ChatGPT-related learning tools depends on motivational, usefulness and self-determination-related factors. Therefore, in this study, perceived AI adaptiveness is treated as a boundary condition that makes desirable difficulty more acceptable. When adaptiveness is high, difficulty may become personalised support. When adaptiveness is low, the same difficulty may become a reason for disengagement.

### **2.6 Conceptual Framework and Hypotheses**

Based on the above theoretical and empirical discussion, this study develops a conceptual framework that links perceived desirable difficulty with learner engagement, productive struggle orientation, intrinsic motivation, knowledge retention confidence and continued use intention. The framework is developed through the integration of desirable difficulties logic, self-determination theory and technology continuance logic. In this particular model, perceived desirable difficulty is positioned as the main antecedent because the study argues that AI tutoring

should not only be examined based on ease, usefulness or satisfaction. Rather, it should also be examined based on whether learners

perceive AI-supported challenge as meaningful for learning.

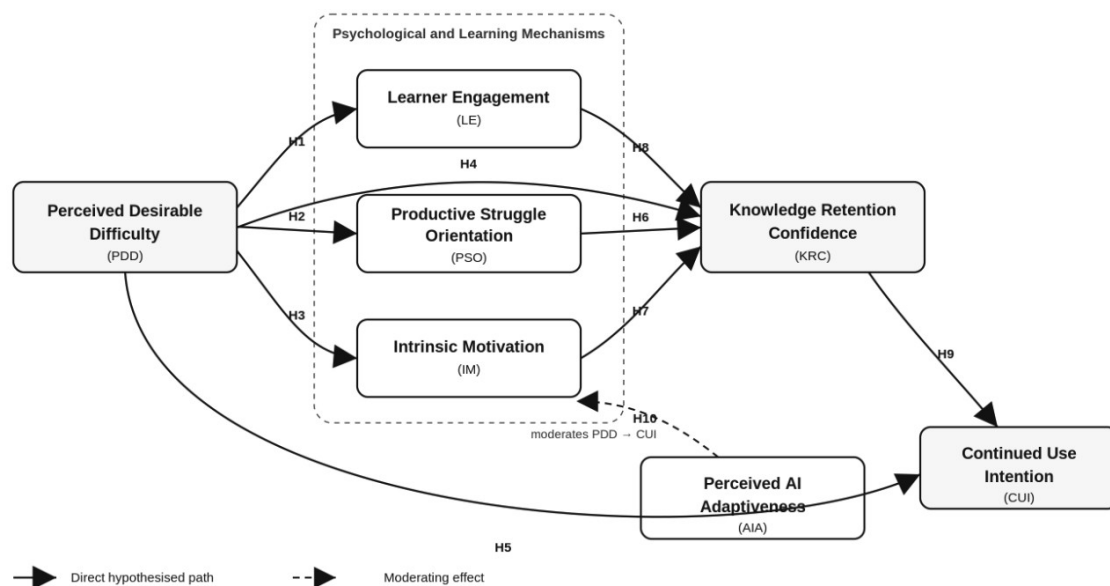


Figure 2.1 Conceptual Framework

The model first proposes that perceived desirable difficulty can directly influence learner engagement, productive struggle orientation and intrinsic motivation. This relationship is developed because manageable difficulty may require students to become more attentive, more persistent and more internally involved in learning. The model also proposes that perceived desirable difficulty directly influences knowledge retention confidence and continued use intention. This means that students may feel more confident about remembering what they learn when AI tutoring creates a productive level of challenge.

The second part of the model explains the mediating mechanisms. Productive struggle orientation, intrinsic motivation and learner engagement are expected to improve knowledge retention confidence because these variables represent effort, interest and active involvement. Knowledge retention confidence is then expected to influence continued use intention because students are more likely to continue using AI tutoring

when they believe that the system helps them remember and apply knowledge. The final part of the model introduces perceived AI adaptiveness as a moderating factor. This means that the relationship between perceived desirable difficulty and continued use intention is expected to become stronger when the AI system is perceived as adaptive.

Based on this framework, the following hypotheses are developed:

H1: Perceived desirable difficulty has a positive influence on learner engagement.

H2: Perceived desirable difficulty has a positive influence on productive struggle orientation.

H3: Perceived desirable difficulty has a positive influence on intrinsic motivation.

H4: Perceived desirable difficulty has a positive influence on knowledge retention confidence.

H5: Perceived desirable difficulty has a positive influence on continued use intention.

H6: Productive struggle orientation has a positive influence on knowledge retention confidence.

H7: Intrinsic motivation has a positive influence on knowledge retention confidence.

H8: Learner engagement has a positive influence on knowledge retention confidence.

H9: Knowledge retention confidence has a positive influence on continued use intention.

H10: Perceived AI adaptiveness moderates the relationship between perceived desirable difficulty and continued use intention, where the relationship becomes stronger when perceived AI adaptiveness is high.

### 3. Research Methodology

#### 3.1 Research Design and Paradigm

This study was conducted by following a positivist research paradigm, where a quantitative and cross-sectional survey design was applied to examine the structural relationship among the selected variables. The positivist approach is appropriate because the study attempts to test a theoretically developed model through observable measurement indicators, numerical responses and statistical analysis. In this particular study, perceived desirable difficulty, learner engagement, productive struggle orientation, intrinsic motivation, perceived AI adaptiveness, knowledge retention confidence and continued use intention were treated as measurable latent constructs. Therefore, the quantitative design helped to examine the relationship nexus among these constructs in a systematic and empirical manner (Creswell & Creswell, 2018).

The cross-sectional design was selected because the data were collected from respondents at one specific point in time. This approach is useful when the purpose of the study is to understand the current perception, attitude and behavioural intention of a particular respondent group. Since the proposed model includes several direct, indirect and moderating relationships, Partial Least Squares Structural Equation Modelling was selected as the main analytical technique. PLS-SEM is suitable for prediction-oriented

research and for estimating complex models containing multiple latent variables and structural paths (Hair et al., 2019). SmartPLS 4 was used because it is widely applied for variance-based structural equation modelling and provides procedures for measurement model, structural model, mediation and moderation analysis (Ringle et al., 2024).

#### 3.2 Population and Sampling

The target population of this study consisted of undergraduate and postgraduate students in Australia who had used AI tutoring or AI-assisted learning tools for academic purposes. This group was selected because university students are now increasingly exposed to artificial intelligence-supported learning in the form of conversational AI systems, adaptive learning platforms, intelligent tutoring systems and other AI-based academic support tools. Therefore, students with direct experience of AI-supported learning were considered suitable for evaluating perceived desirable difficulty, adaptiveness, engagement, motivation and continued use intention. This population selection is consistent with purposive research logic, where participants are selected because they possess the required experience and information for the study (Etikan et al., 2016).

A non-probability purposive sampling approach was applied because the study required respondents who had already used AI tutoring or AI-assisted learning tools. General students without such experience would not be able to provide meaningful responses to the measurement items. In this regard, purposive sampling was more appropriate than random sampling because the research was not simply interested in the general student population, but in students who could evaluate AI-supported learning difficulty from actual use experience. Undergraduate and postgraduate students were both included because AI tutoring is not limited to one academic level. From this sampling position, it can be said that the sample was selected based on contextual relevance and direct knowledge of the

phenomenon under investigation (Etikan et al., 2016).

### **3.3 Data Collection Procedure**

Primary data were collected through a structured online questionnaire. The online survey method was appropriate because the target respondents were university students who were already familiar with digital platforms and AI-assisted learning tools. Before answering the questionnaire, respondents were informed about the academic purpose of the study, the voluntary nature of participation and the anonymous handling of responses. Informed consent was collected before the respondents proceeded to the main survey. This procedure helped to maintain ethical transparency and also supported more honest responses because no personally identifiable information was collected (Dillman et al., 2014; Podsakoff et al., 2003).

The questionnaire was designed in a structured format so that all respondents answered the same set of items in the same measurement order. This is important in survey research because standardised measurement improves comparability across respondents. To reduce possible response bias, the survey was kept anonymous and respondents were not asked to provide names, student identification numbers or institutional identifiers. From this data collection procedure, it can be understood that the survey method was consistent with the quantitative and cross-sectional nature of the study (Creswell & Creswell, 2018; Dillman et al., 2014).

### **3.4 Measures and Instrumentation**

The research instrument was developed by using a structured questionnaire that measured seven latent constructs. These constructs were Perceived Desirable Difficulty, Learner Engagement, Knowledge Retention Confidence, Productive Struggle Orientation, Intrinsic Motivation, Perceived AI Adaptiveness and Continued Use Intention. In survey-based behavioural research, latent constructs are not measured directly. Rather, they are measured through

observable indicators that represent the underlying theoretical concept. Therefore, each construct in this study was measured through three indicators, which allowed the study to capture the construct meaning while keeping the questionnaire concise for respondents (DeVellis, 2017; Hair et al., 2019).

All constructs were operationalised as reflective measurement constructs. This means that the indicators were treated as manifestations of the underlying latent variable. For example, the items of learner engagement reflected the broader engagement condition of the learner, while the items of continued use intention reflected the respondent's future intention to keep using AI tutoring tools. Reflective specification is suitable when the indicators are expected to be correlated and when the latent construct is assumed to cause the observed item responses. This measurement logic is consistent with PLS-SEM guidelines for assessing indicator reliability, internal consistency reliability, convergent validity and discriminant validity (Hair et al., 2019).

A five-point Likert scale was used for all measurement items, where 1 represented strongly disagree and 5 represented strongly agree. Likert-type scales are commonly used in attitude and perception-based research because they allow respondents to express the degree of agreement with a statement in an ordered response format. In this study, the five-point scale was considered suitable because the constructs were perception-based and required respondents to evaluate their experience with AI tutoring systems. The use of a five-point format also made the questionnaire easier to complete, which is important in online survey-based data collection (Likert, 1932; Koo & Yang, 2025).

Perceived Desirable Difficulty measured whether students viewed AI-supported learning tasks as challenging in a productive and useful way. Learner Engagement measured the degree of active involvement during AI-assisted learning. Productive Struggle Orientation measured the willingness of students to work through

difficult academic tasks with AI support. Intrinsic Motivation measured internal interest and enjoyment in AI-supported learning. Knowledge Retention Confidence measured students' belief that they could remember and later use what they learned through AI tutoring. Perceived AI Adaptiveness measured whether students felt that the AI system adjusted to their learning level. Continued Use Intention measured the likelihood of future AI tutoring use. From the above measurement structure, it can be said that the instrument captured both the learning psychology and continuance behaviour dimensions of the study (DeVellis, 2017; Hair et al., 2019).

### **3.5 Data Analysis Procedure**

The data were analysed by using SmartPLS 4, and the analysis followed the standard two-stage PLS-SEM procedure. In the first stage, the measurement model was assessed to confirm the reliability and validity of the constructs. This included the examination of outer loadings, Cronbach's alpha, composite reliability, average variance extracted and discriminant validity. These tests were necessary because the structural relationships cannot be interpreted confidently unless the measurement model first shows acceptable reliability and validity. Therefore, the measurement model assessment provided the foundation for testing the hypothesised relationships (Hair et al., 2019; Ringle et al., 2024).

In the second stage, the structural model was assessed by examining path coefficients, t-statistics, p-values, coefficient of determination and effect sizes. The significance of the hypothesised paths was tested through bootstrapping with 5,000 resamples. Bootstrapping is appropriate in PLS-SEM because it provides a non-parametric way of estimating standard errors and significance levels without requiring strict normality assumptions. In this study, the bootstrapping procedure was used to test the direct effects among perceived desirable difficulty, learner engagement, productive struggle orientation, intrinsic motivation,

knowledge retention confidence, perceived AI adaptiveness and continued use intention (Hair et al., 2019).

Mediation analysis was conducted to examine whether learner engagement, productive struggle orientation and intrinsic motivation explain the relationship between perceived desirable difficulty and knowledge retention confidence. The indirect pathways from perceived desirable difficulty to continued use intention through knowledge retention confidence were also examined. The interpretation of mediation followed the indirect effect approach, where mediation is supported when the bootstrapped confidence interval does not include zero. This approach is consistent with the argument that mediation should be assessed through the significance of indirect effects rather than only through the traditional stepwise approach (Zhao et al., 2010; Hair et al., 2019).

Moderation analysis was also conducted to examine whether perceived AI adaptiveness changes the strength of the relationship between perceived desirable difficulty and continued use intention. In this model, perceived AI adaptiveness was treated as a boundary condition because difficulty may not be accepted positively by students unless the AI tutoring system is perceived as responsive and adaptive. Therefore, the interaction term between perceived desirable difficulty and perceived AI adaptiveness was tested in SmartPLS 4. This allowed the study to examine whether adaptive AI support strengthens the effect of productive difficulty on continued use intention (Hair et al., 2019; Ringle et al., 2024).

### **3.6 Ethical Considerations and Methodological Limitations**

The study followed basic ethical research principles by ensuring voluntary participation, informed consent, anonymity and academic use of the collected data. These procedures were important because the respondents were asked to report their own learning perceptions and AI use experiences. However, the study also has methodological

limitations. Since the data were collected through a cross-sectional survey, the findings cannot explain changes in students' perceptions over time. Moreover, self-reported survey data may contain some response bias. Therefore, future studies may apply longitudinal, experimental or mixed-method designs to provide stronger evidence about AI tutoring and desirable difficulty (Creswell & Creswell, 2018; Podsakoff et al., 2003).

## 4. Results and Analysis

### 4.1 Respondent Profile and Preliminary Data Screening

Before testing the measurement and structural model, the data set was screened

**Table 4.1:**  
*Respondent Profile and Preliminary Data Screening*

| Data screening item                        | Result                     |
|--------------------------------------------|----------------------------|
| Initial usable cases in SmartPLS           | 358                        |
| Case index range                           | 0–357                      |
| Demographic table available in output      | Not available              |
| Missing value treatment in SmartPLS output | Complete score matrix used |
| Bootstrapping resamples                    | 5,000                      |
| Final cases retained for PLS-SEM           | 358                        |

From the above table, it can be understood that all 358 usable responses were retained for analysis. Since no separate demographic profile table was available in the output file, the respondent profile is limited to the final valid sample size. However, the availability of a complete indicator matrix and successful bootstrapping procedure supports the adequacy of the data for testing the proposed model.

for its suitability for PLS-SEM analysis. The final usable data set contained 358 responses, where the original case index ranged from 0 to 357 in the SmartPLS indicator data file. The demographic table was not generated in the SmartPLS output. Therefore, this section reports the usable analytical sample and preliminary data condition rather than demographic characteristics. From the SmartPLS execution output, it can be seen that the software read the score matrix as a complete data set, calculated the original sample and then performed bootstrapping with 5,000 resamples. This indicates that the final data set was suitable for further measurement and structural model analysis.

### 4.2 Measurement Model Assessment

The measurement model was assessed by examining indicator reliability, internal consistency reliability, convergent validity and discriminant validity. Following the criteria of Hair et al. (2019), outer loading values should preferably exceed 0.70, while Cronbach's alpha, composite reliability and AVE should meet the accepted threshold levels for reliability and validity. In this study, all observed items showed very high outer loading values. This indicates that the indicators were strongly related with their respective latent constructs. Therefore, the

item reliability of the measurement model can be considered highly satisfactory.

**Table 02:**  
*Convergent Validity and Reliability Results*

| Construct                             | Items | Loading range | Cronbach's alpha | rho_A | CR    | AVE   | Decision  |
|---------------------------------------|-------|---------------|------------------|-------|-------|-------|-----------|
| Perceived AI Adaptiveness (AIA)       | 3     | 0.947–0.953   | 0.946            | 0.948 | 0.965 | 0.902 | Supported |
| Continued Use Intention (CUI)         | 3     | 0.944–0.948   | 0.942            | 0.942 | 0.963 | 0.896 | Supported |
| Intrinsic Motivation (IM)             | 3     | 0.952–0.953   | 0.949            | 0.949 | 0.967 | 0.907 | Supported |
| Knowledge Retention Confidence (KRC)  | 3     | 0.941–0.946   | 0.938            | 0.938 | 0.961 | 0.890 | Supported |
| Learner Engagement (LE)               | 3     | 0.947–0.955   | 0.946            | 0.946 | 0.965 | 0.903 | Supported |
| Perceived Desirable Difficulty (PDD)  | 3     | 0.949–0.954   | 0.948            | 0.948 | 0.966 | 0.905 | Supported |
| Productive Struggle Orientation (PSO) | 3     | 0.948–0.954   | 0.947            | 0.948 | 0.966 | 0.905 | Supported |

Table 02 depicts the convergent validity and reliability results of the measurement model. From the above findings, it can be seen that all item loading ranges were above 0.90. The lowest loading range was found for Knowledge Retention Confidence, where the indicators loaded between 0.941 and 0.946. The highest item loading was found under Learner Engagement, where LE3 loaded at 0.955. This result indicates that the items used for measuring perceived desirable difficulty, learner engagement, productive struggle orientation, intrinsic motivation, perceived AI adaptiveness, knowledge retention confidence and continued use intention were strongly connected with their respective constructs.

The reliability values also show a very strong measurement model. Cronbach's alpha values ranged from 0.938 to 0.949, while composite reliability values ranged from 0.961 to 0.967. These values are much higher than the recommended threshold of 0.70. Therefore, it can be said that all constructs had strong internal consistency reliability. Moreover, AVE values ranged from 0.890 to 0.907, which means that each construct explained a very high proportion of variance in its indicators. Since all AVE values were higher than 0.50, convergent validity was clearly established.

Discriminant validity was assessed by using both the Fornell-Larcker criterion and the Heterotrait-Monotrait ratio. The Fornell-Larcker criterion explains that the square root of AVE for each construct should be higher than its correlations with other constructs (Fornell & Larcker, 1981). In Table 03, the diagonal values represent the square root of AVE, while the off-diagonal values represent latent variable correlations. From the above result, it can be seen that all diagonal values were higher than the corresponding inter-construct correlations. Therefore, the first condition of discriminant validity was established.

**Table 03:**  
*Fornell-Larcker Criterion*

| Construct | AIA   | CUI   | IM    | KRC   | LE | PDD | PSO |
|-----------|-------|-------|-------|-------|----|-----|-----|
| AIA       | 0.950 |       |       |       |    |     |     |
| CUI       | 0.420 | 0.947 |       |       |    |     |     |
| IM        | 0.150 | 0.518 | 0.952 |       |    |     |     |
| KRC       | 0.162 | 0.689 | 0.658 | 0.943 |    |     |     |

| Construct | AIA   | CUI   | IM    | KRC   | LE    | PDD   | PSO   |
|-----------|-------|-------|-------|-------|-------|-------|-------|
| LE        | 0.140 | 0.566 | 0.407 | 0.663 | 0.950 |       |       |
| PDD       | 0.241 | 0.712 | 0.605 | 0.724 | 0.622 | 0.951 |       |
| PSO       | 0.123 | 0.530 | 0.418 | 0.654 | 0.472 | 0.640 | 0.951 |

Table 03 shows that the square root of AVE values ranged from 0.943 to 0.952. These values were higher than all correlation values in the same rows and columns. For example, the square root of AVE for Perceived Desirable Difficulty was 0.951, while its highest correlation was with Knowledge Retention Confidence at 0.724. Similarly, the square root of AVE for Knowledge Retention Confidence was 0.943, which was also higher than its correlations with Continued Use Intention, Intrinsic Motivation, Learner Engagement and Productive Struggle Orientation. From this result, it can be said that the constructs have adequate discriminant separation under the Fornell-Larcker criterion.

After that, discriminant validity was further assessed by using the Heterotrait-Monotrait ratio. HTMT is considered a stricter criterion for discriminant validity in variance-based SEM because it compares the average correlations across constructs with the average correlations within constructs (Henseler et al., 2015). The HTMT results are presented in Table 04. In this study, all

HTMT values were below the conservative threshold of 0.85. Therefore, the constructs can be considered empirically distinct from each other.

**Table 04:**  
*Heterotrait-Monotrait Ratio*

| Construct | AIA   | CUI   | IM    | KRC   | LE    | PDD   | PSO |
|-----------|-------|-------|-------|-------|-------|-------|-----|
| AIA       |       |       |       |       |       |       |     |
| CUI       | 0.445 |       |       |       |       |       |     |
| IM        | 0.157 | 0.547 |       |       |       |       |     |
| KRC       | 0.171 | 0.733 | 0.697 |       |       |       |     |
| LE        | 0.146 | 0.600 | 0.429 | 0.704 |       |       |     |
| PDD       | 0.253 | 0.754 | 0.637 | 0.768 | 0.657 |       |     |
| PSO       | 0.129 | 0.561 | 0.441 | 0.693 | 0.498 | 0.675 |     |

Table 04 shows that all HTMT values remained below the recommended threshold of 0.85. The highest HTMT value was found between Perceived Desirable Difficulty and Knowledge Retention Confidence at 0.768. This relationship is theoretically reasonable because desirable difficulty is expected to be connected with students' confidence in retaining what they have learned. However, the value is still below the conservative cut-off point. Similarly, the HTMT value between Perceived Desirable Difficulty and Continued Use Intention was 0.754, while the value between Knowledge Retention Confidence and Continued Use Intention was 0.733. These values show that the constructs are related but not overlapping. From the above findings, it can be established that the measurement model has both convergent precision and discriminant separation. Therefore, the structural model can be tested based on a reliable and valid measurement foundation.

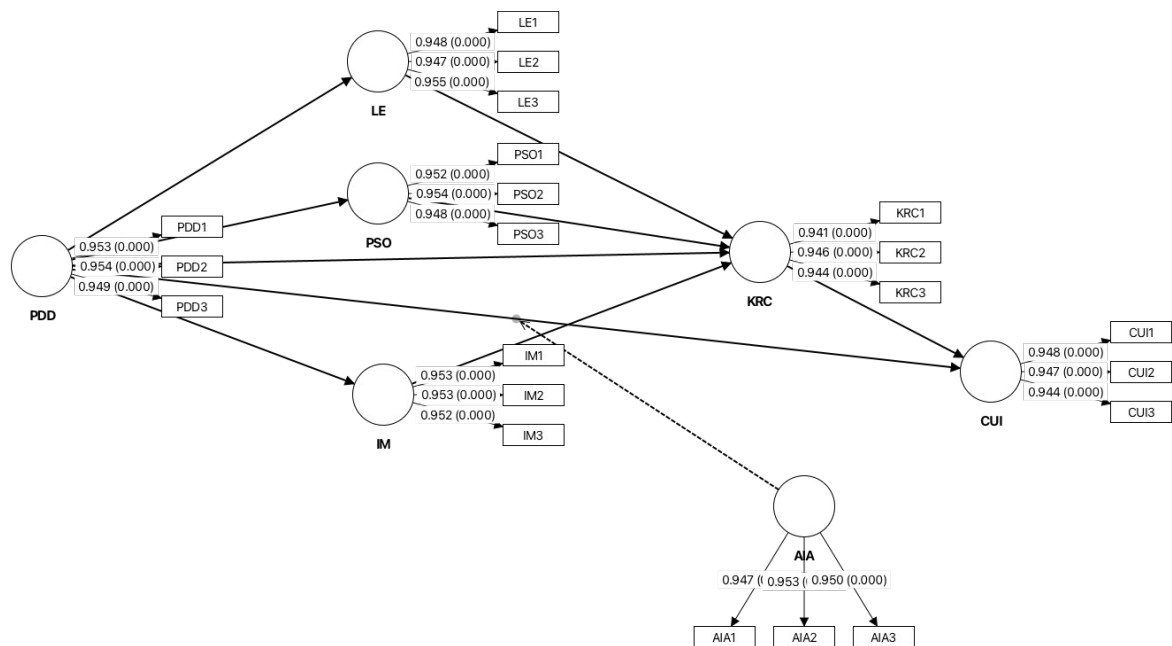


Figure 4.1: Measurement Model Results

### 4.3 Structural Model Assessment and Explanatory Power

After confirming the reliability and validity of the measurement model, the structural model was assessed through the coefficient of determination. The  $R^2$  value

explains how much variance in an endogenous construct is explained by its predictor variables. According to Hair et al. (2019),  $R^2$  values of 0.75, 0.50 and 0.25 can be considered substantial, moderate and weak, although interpretation also depends on the research context. The  $R^2$  results of this study are presented in Table 04.

**Table 04:**  
*Explanatory Power of the Structural Model*

| Endogenous construct                  | $R^2$ | Adjusted $R^2$ | Interpretation             |
|---------------------------------------|-------|----------------|----------------------------|
| Continued Use Intention (CUI)         | 0.703 | 0.700          | Strong explanatory power   |
| Knowledge Retention Confidence (KRC)  | 0.709 | 0.706          | Strong explanatory power   |
| Productive Struggle Orientation (PSO) | 0.410 | 0.408          | Moderate explanatory power |
| Learner Engagement (LE)               | 0.387 | 0.385          | Moderate explanatory power |
| Intrinsic Motivation (IM)             | 0.366 | 0.364          | Moderate explanatory power |

Source: Prepared by author from SmartPLS output.

Table 04 shows that Knowledge Retention Confidence had the highest  $R^2$  value at 0.709. This means that Perceived Desirable Difficulty, Learner Engagement, Productive Struggle Orientation and Intrinsic Motivation jointly explained 70.9% of the variance in

Knowledge Retention Confidence. This is an important result because it shows that students' confidence in retaining AI-supported learning is strongly explained by the psychological and learning mechanisms included in the model. Therefore, retention

confidence is not only a simple outcome of AI use. Rather, it is developed through the combined influence of challenge, engagement, motivation and productive struggle.

Continued Use Intention also showed strong explanatory power, with an  $R^2$  value of 0.703. This indicates that 70.3% of the variance in students' intention to continue using AI tutoring was explained by the model. From this finding, it can be understood that continued use intention is shaped by more than simple usefulness or ease of use. It is also influenced by desirable difficulty, perceived AI adaptiveness and knowledge retention confidence. The  $R^2$

values for Productive Struggle Orientation, Learner Engagement and Intrinsic Motivation were 0.410, 0.387 and 0.366 respectively. These values indicate that Perceived Desirable Difficulty has a meaningful role in explaining the learning responses of students. Therefore, the structural model provides strong support for the argument that productive AI tutoring difficulty can shape both immediate learning mechanisms and longer-term continuance intention.

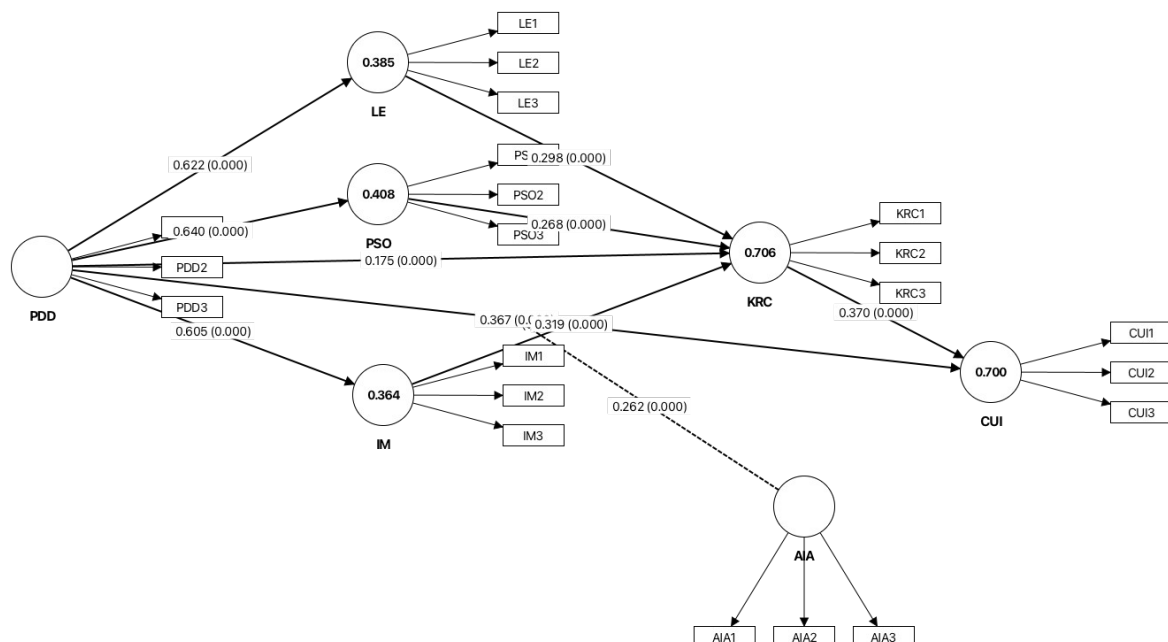


Figure 03: Structural Model Results

#### 4.4 Hypothesis Testing: Direct Effects

After confirming the explanatory power of the structural model, the direct hypotheses were tested through the bootstrapping procedure with 5,000 resamples. The path coefficient, t-value and p-value were used to

identify whether the proposed relationships were statistically significant. The direct effect results are presented in Table 05. From the findings, all direct hypothesised paths were found significant at the level of  $p < 0.001$ . Therefore, the direct relationship structure of the model is supported.

**Table 05**

shows that *Perceived Desirable Difficulty* had a strong positive influence on *Learner Engagement*, *Productive Struggle Orientation* and *Intrinsic Motivation*. The strongest direct effect was found for  $PDD \rightarrow PSO$ , where the coefficient value was 0.640. This indicates that when students perceive AI tutoring difficulty as useful and manageable, they become more willing to work through challenging learning tasks. The second strongest effect was  $PDD \rightarrow LE$  with  $\beta = 0.622$ , followed by  $PDD \rightarrow IM$  with  $\beta = 0.605$ . From these findings, it can be said that perceived desirable difficulty works as an important learning condition because it strengthens engagement, struggle orientation and motivation at the same time.

| Hypothesis | Relationship          | Path coefficient $\beta$ | T-statistic | P-value | Decision    |
|------------|-----------------------|--------------------------|-------------|---------|-------------|
| H1         | $PDD \rightarrow LE$  | 0.622                    | 18.059      | 0.000   | Supported   |
| H2         | $PDD \rightarrow PSO$ | 0.640                    | 19.485      | 0.000   | Supported   |
| H3         | $PDD \rightarrow IM$  | 0.605                    | 16.351      | 0.000   | Supported   |
| H4         | $PDD \rightarrow KRC$ | 0.175                    | 3.608       | 0.000   | Supported   |
| H5         | $PDD \rightarrow CUI$ | 0.367                    | 8.677       | 0.000   | Supported   |
| H6         | $PSO \rightarrow KRC$ | 0.268                    | 7.016       | 0.000   | Supported   |
| H7         | $IM \rightarrow KRC$  | 0.319                    | 8.903       | 0.000   | Supported   |
| H8         | $LE \rightarrow KRC$  | 0.298                    | 8.434       | 0.000   | Supported   |
| H9         | $KRC \rightarrow CUI$ | 0.370                    | 8.937       | 0.000   | Supported   |
| H10        | $AIA \rightarrow CUI$ | 0.272                    | 8.549       | 0.000   | Significant |

The result also shows that *Perceived Desirable Difficulty* had a significant direct effect on *Knowledge Retention Confidence* and *Continued Use Intention*. Although the direct effect of *PDD* on *KRC* was weaker than its effects on the learning mechanism variables, it was still statistically significant with  $\beta = 0.175$  and  $p = 0.000$ . This means that productive difficulty directly contributes to students' confidence in retaining AI-supported learning. Moreover, *PDD* also influenced *CUI* with  $\beta = 0.367$ . Therefore, students are more likely to continue using AI tutoring when the system creates a difficulty level that is perceived as meaningful rather than unnecessary.

*Knowledge Retention Confidence* was also found as a strong predictor of *Continued Use Intention*, where the coefficient value was 0.370. This finding indicates that students' intention to continue using AI tutoring is strongly connected with their

belief that the system helps them retain learning. In the case of predictors of *KRC*, *Intrinsic Motivation* had the strongest effect with  $\beta = 0.319$ , followed by *Learner Engagement* with  $\beta = 0.298$  and *Productive Struggle Orientation* with  $\beta = 0.268$ . From the above findings, it can be established that retention confidence is developed through motivation, engagement and productive struggle. Therefore, the direct effect results support the main argument that desirable difficulty works through both psychological and learning-based channels.

#### 4.5 Mediation Analysis

The mediation analysis was conducted by examining the specific indirect effects generated through the bootstrapping procedure. In this study, mediation is important because the model does not only assume that perceived desirable difficulty directly influences retention confidence and

continued use. Rather, the model argues that desirable difficulty becomes useful when it activates learner engagement, intrinsic

motivation and productive struggle orientation. The results of the specific indirect effects are presented in Table 06.

**Table 06:**  
*Specific Indirect Effects*

| Indirect relationship                                     | Indirect effect $\beta$ | T-statistic | P-value | Decision  |
|-----------------------------------------------------------|-------------------------|-------------|---------|-----------|
| PDD $\rightarrow$ PSO $\rightarrow$ KRC                   | 0.171                   | 6.551       | 0.000   | Supported |
| PDD $\rightarrow$ IM $\rightarrow$ KRC                    | 0.193                   | 7.726       | 0.000   | Supported |
| PDD $\rightarrow$ LE $\rightarrow$ KRC                    | 0.186                   | 7.444       | 0.000   | Supported |
| PDD $\rightarrow$ KRC $\rightarrow$ CUI                   | 0.064                   | 3.269       | 0.001   | Supported |
| PDD $\rightarrow$ PSO $\rightarrow$ KRC $\rightarrow$ CUI | 0.063                   | 5.109       | 0.000   | Supported |
| PDD $\rightarrow$ IM $\rightarrow$ KRC $\rightarrow$ CUI  | 0.071                   | 5.961       | 0.000   | Supported |
| PDD $\rightarrow$ LE $\rightarrow$ KRC $\rightarrow$ CUI  | 0.069                   | 5.735       | 0.000   | Supported |
| PSO $\rightarrow$ KRC $\rightarrow$ CUI                   | 0.099                   | 5.312       | 0.000   | Supported |
| IM $\rightarrow$ KRC $\rightarrow$ CUI                    | 0.118                   | 6.586       | 0.000   | Supported |
| LE $\rightarrow$ KRC $\rightarrow$ CUI                    | 0.110                   | 6.196       | 0.000   | Supported |

Table 06 shows that all specific indirect effects were statistically significant. The strongest indirect path from Perceived Desirable Difficulty to Knowledge Retention Confidence was found through Intrinsic Motivation, where the indirect effect was 0.193. This indicates that perceived desirable difficulty increases intrinsic motivation, and this motivation then improves students' confidence in retaining knowledge. The second strongest indirect path was PDD  $\rightarrow$  LE  $\rightarrow$  KRC with  $\beta = 0.186$ , followed by PDD  $\rightarrow$  PSO  $\rightarrow$  KRC with  $\beta = 0.171$ . From these findings, it can be understood that desirable difficulty does not work through one single psychological route. Rather, it works through a combined mechanism of motivation, engagement and productive struggle.

The sequential mediation results also support the deeper relationship nexus of the model. The indirect effects from PDD to CUI through IM and KRC, LE and KRC, and PSO

and KRC were all significant. Among these, the strongest sequential indirect effect was PDD  $\rightarrow$  IM  $\rightarrow$  KRC  $\rightarrow$  CUI with  $\beta = 0.071$ . This was followed by PDD  $\rightarrow$  LE  $\rightarrow$  KRC  $\rightarrow$

CUI with  $\beta = 0.069$  and  $PDD \rightarrow PSO \rightarrow KRC \rightarrow CUI$  with  $\beta = 0.063$ . Therefore, it can be said that students are more likely to continue using AI tutoring when desirable difficulty first increases motivation, engagement or productive struggle, then strengthens retention confidence, and finally develops continuance intention.

#### 4.6 Moderation Analysis

The moderating effect of Perceived AI Adaptiveness was tested by using the

interaction term  $AIA \times PDD$ . The result shows that the interaction effect on Continued Use Intention was positive and statistically significant, where  $\beta = 0.262$ ,  $t = 9.531$  and  $p = 0.000$ . Therefore, H10 is supported. This means that perceived AI adaptiveness strengthens the relationship between perceived desirable difficulty and continued use intention. In other words, students are more willing to continue using AI tutoring when the learning difficulty feels adaptive and matched with their learning condition.

**Table 07:**  
*Moderation and Conditional Direct Effects*

| Relationship                     | Condition          | Effect $\beta$ | 95% CI lower | 95% CI upper | Decision        |
|----------------------------------|--------------------|----------------|--------------|--------------|-----------------|
| $AIA \times PDD \rightarrow CUI$ | Interaction effect | 0.262          | 0.209        | 0.316        | Supported       |
| $PDD \rightarrow CUI$            | AIA at +1 SD       | 0.629          | 0.540        | 0.716        | Significant     |
| $PDD \rightarrow CUI$            | AIA at mean        | 0.367          | 0.281        | 0.445        | Significant     |
| $PDD \rightarrow CUI$            | AIA at -1 SD       | 0.105          | -0.004       | 0.210        | Not significant |

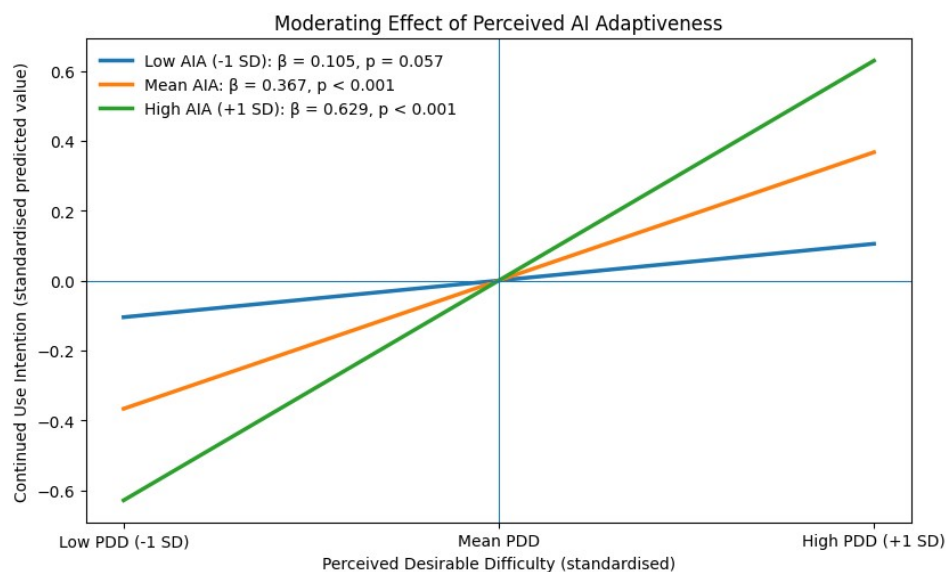


Figure 04: Moderating Effect of Perceived AI Adaptiveness

Table 07 shows that the effect of PDD on CUI was strongest when AIA was high. At +1 standard deviation of AIA, the effect of PDD on CUI was 0.629, and the confidence interval did not include zero. At the mean level of AIA, the effect was also significant

with  $\beta = 0.367$ . However, when AIA was low, the effect became weak with  $\beta = 0.105$ , and the confidence interval included zero. From this result, it can be said that desirable difficulty alone is not always enough to create continued use intention. It becomes more

acceptable when students believe that the AI tutoring system is adaptive, responsive and properly adjusted to their level of understanding. Therefore, perceived AI adaptiveness works as an important boundary condition in this study.

#### 4.7 Summary of Hypothesis Decisions

The overall hypothesis testing result is presented in Table 08. From the above findings, it can be seen that all direct, mediating and moderating hypotheses were supported. Therefore, the proposed conceptual framework is empirically supported.

**Table 08:**  
*Summary of Hypothesis Decisions*

| Hypothesis | Relationship    | Decision  |
|------------|-----------------|-----------|
| H1         | PDD → LE        | Supported |
| H2         | PDD → PSO       | Supported |
| H3         | PDD → IM        | Supported |
| H4         | PDD → KRC       | Supported |
| H5         | PDD → CUI       | Supported |
| H6         | PSO → KRC       | Supported |
| H7         | IM → KRC        | Supported |
| H8         | LE → KRC        | Supported |
| H9         | KRC → CUI       | Supported |
| H10        | AIA × PDD → CUI | Supported |

## 5. Discussion and Conclusion

### 5.1 Summary of Major Findings

The findings of this study show that Perceived Desirable Difficulty is the main psychological and learning driver in the AI tutoring model. The result shows that PDD had strong direct effects on Learner Engagement, Productive Struggle Orientation and Intrinsic Motivation. This means that when students perceive AI tutoring difficulty as useful, manageable and learning-supportive, they are more likely to become engaged, motivated and willing to work through challenging learning tasks. Therefore, difficulty in AI tutoring should not be understood only as a barrier. Rather, it can work as an important learning condition when the difficulty is perceived as desirable.

The result also shows that PDD directly influenced Knowledge Retention Confidence and Continued Use Intention. This indicates that students who experience productive difficulty in AI tutoring are more likely to believe that they can remember what they

have learned and are also more likely to continue using the AI tutoring system. From the above findings, it can be understood that knowledge retention confidence works as an important outcome in this model because it connects the learning experience with future technology use.

The mediation findings further confirm that PDD works through several psychological mechanisms. Intrinsic Motivation, Learner Engagement and Productive Struggle Orientation significantly explained the relationship between PDD and Knowledge Retention Confidence. The sequential indirect effects also showed that PDD influences Continued Use Intention through these mechanisms and Knowledge Retention Confidence. Finally, the moderation result shows that Perceived AI Adaptiveness strengthens the relationship between PDD and CUI. Therefore, desirable difficulty becomes more useful when the AI system is perceived as adaptive and responsive to student learning needs.

### 5.2 Discussion of Theoretical Implications

The findings of this study provide important theoretical implications for desirable difficulties, self-determination theory and technology continuance research. First, the result extends the Desirable Difficulties Framework into the AI tutoring context. Bjork (1994) and Bjork and Bjork (2020) argued that some learning difficulties can improve long-term learning when they create deeper processing and stronger retrieval. The present study supports this argument by showing that perceived desirable difficulty has a strong effect on engagement, productive struggle and retention confidence. Therefore, the result indicates that difficulty is not harmful by nature. It becomes harmful only when it is unmanageable, poorly supported or disconnected from learning value.

Second, the findings provide support for Self-Determination Theory. Ryan and Deci (2000, 2020) explained that motivation becomes stronger when learners experience competence, autonomy and meaningful involvement. In this study, PDD had a strong influence on Intrinsic Motivation, and Intrinsic Motivation also had a significant effect on Knowledge Retention Confidence. From this finding, it can be said that productive AI tutoring difficulty may support students' sense of competence when the challenge is not excessive. This is also consistent with Galindo-Domínguez et al. (2025), Li et al. (2025), Yang et al. (2025) and Annamalai et al. (2025), where AI-supported learning and motivation were shown to depend on how learners psychologically experience the AI system.

Third, the findings contribute to productive struggle literature. Young et al. (2024), Bellwether (2025) and Baule (2025) argued that struggle can become educationally useful when it is purposeful and supported. In this study, PDD strongly influenced Productive Struggle Orientation, and PSO also improved Knowledge Retention Confidence. Therefore, the result shows that AI tutoring should not remove all struggle from the learning process. Rather, it should help students work through struggle in a guided and meaningful way.

Finally, the result contributes to technology continuance logic. Davis (1989) explained that technology acceptance is influenced by perceived usefulness, while recent studies on AI-based learning also show that continued use depends on usefulness, motivation, satisfaction and self-determination-related factors (Dahri et al., 2024; Fang & Feng, 2026; Xu et al., 2025). This study adds that continued use may also be shaped by productive difficulty and retention confidence. Therefore, AI tutoring continuance is not only a matter of ease or convenience. It is also a matter of whether the system creates learning value through

adaptive and psychologically meaningful challenge.

### 5.3 Practical Implications

The findings of this study provide practical implications for AI tutoring designers, universities and educators. First, AI tutoring designers should not develop systems that only make learning easy or provide immediate answers. The result shows that perceived desirable difficulty strongly improves learner engagement, intrinsic motivation and productive struggle orientation. Therefore, AI tutoring systems should be designed to create manageable challenge, step-by-step feedback and opportunities for students to think before receiving the final answer. This is consistent with the argument of Kestin et al. (2025) and Chung et al. (2026), where AI tutoring becomes more valuable when it supports learning through structured and personalised guidance.

Second, universities should use AI tutoring as part of learning support rather than only as a convenience tool. If students use AI only to complete tasks quickly, the deeper learning value may become limited. For this reason, institutions should train students to use AI tutoring for explanation, revision, retrieval practice and problem-solving. This can improve engagement and retention confidence. From the above findings, it can be understood that AI-supported learning should be connected with active learning behaviour, not passive answer collection.

Third, educators should design AI-integrated activities where difficulty is purposeful and properly scaffolded. The moderation result shows that students are more likely to accept difficulty when the AI system feels adaptive. Therefore, teachers should guide students to use AI tools that adjust explanation, provide feedback and respond to the learner's level. This implication is consistent with Contrino et al. (2024), Demartini et al. (2024), Jamali et al.

(2025) and Hidayat-ur-Rehman (2024), where adaptive support and student engagement were important for effective AI-based learning. Therefore, the practical message is clear. AI tutoring should create challenge, but it should also create support.

#### **5.4 Limitations and Future Research**

Although this study provides strong empirical support for the proposed model, some limitations are present. First, the study used a cross-sectional research design. Due to this reason, the findings explain the relationship among students' perceptions at one point in time, but they do not show how perceptions of AI tutoring difficulty may change over a longer period. Future researchers can apply longitudinal designs to examine whether desirable difficulty continues to influence motivation, retention confidence and continued use after repeated AI tutoring experience.

Second, the study used self-reported survey data. This is appropriate for measuring perception-based constructs such as motivation, engagement and continued use intention. However, self-reported data may contain response bias because students may overstate their engagement or confidence. Future studies can combine survey responses with actual learning records, platform usage logs, performance scores or experimental learning tasks. This would provide a stronger understanding of how perceived desirable difficulty connects with actual learning behaviour.

Third, the study included students who used different types of AI tutoring and AI-assisted learning tools. This reflects the real educational environment because students use different AI systems for different purposes. However, tool variation may also influence the way students experience difficulty, adaptiveness and feedback. Future studies can compare different AI tutoring platforms, such as conversational AI tutors,

adaptive learning systems and discipline-specific intelligent tutoring systems.

Finally, the study focused on university students. Therefore, the findings may not fully represent school learners, adult learners or professional learners. Future research can test the model in different educational levels, cultural settings and subject areas. Experimental studies may also manipulate difficulty and adaptiveness directly to examine whether productive AI difficulty improves real retention outcomes.

#### **5.5 Conclusion**

This study examined how perceived desirable difficulty influences engagement, motivation, productive struggle, retention confidence and continued use intention in AI tutoring. From the above findings, it can be concluded that desirable difficulty is a central learning condition in AI-supported education. The result shows that PDD improves learner engagement, intrinsic motivation and productive struggle orientation, and these mechanisms strengthen knowledge retention confidence. Knowledge retention confidence then supports continued use intention.

The study also confirms that perceived AI adaptiveness is an important boundary condition. When students feel that the AI system adjusts to their learning level, difficulty becomes more acceptable and more useful. Therefore, AI tutoring should not only reduce effort or make learning easier. It should create a balanced learning environment where challenge, feedback and adaptive support work together. From this study, it can be said that productive difficulty can support sustained AI tutoring use when students experience the system as personalised, meaningful and learning-focused.

#### **References**

Almufarreah, A. (2024). Determinants of students' satisfaction with AI tools in

- p>education: A PLS-SEM-ANN approach. Sustainability, 16(13), Article 5354.
- 
- <https://doi.org/10.3390/su16135354>
- Annamalai, N., Bervell, B., Mireku, D. O., & Andoh, R. P. K. (2025). Artificial intelligence in higher education: Modelling students' motivation for continuous use of ChatGPT based on a modified self-determination theory. Computers and Education: Artificial Intelligence, 8, Article 100346.  
<https://doi.org/10.1016/j.caeai.2024.100346>
- Baule, S. M. (2025, April 1). Opportunities in failing: Why K-12 education needs more productive struggle. eSchool News.  
<https://www.eschoolnews.com/innovative-teaching/2025/04/02/failing-k-12-education-productive-struggle/>
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185–205). MIT Press.  
<https://doi.org/10.7551/mitpress/4561.003.0011>
- Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. Journal of Applied Research in Memory and Cognition, 9(4), 475–479.  
<https://doi.org/10.1016/j.jarmac.2020.09.003>
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. Psychological Bulletin, 132(3), 354–380. <https://doi.org/10.1037/0033-2909.132.3.354>
- Chen Kulesa, A., Mission, M., Croft, M., & Wells, M. K. (2025, June). Productive struggle: How artificial intelligence is changing learning, effort, and youth development in education. Bellwether.  
<https://bellwether.org/publications/productive-struggle/>
- Chung, A. T.-H., Zhang, B., Kung, L.-C., Bastani, H., & Bastani, O. (2026). Effective personalized AI tutors via LLM-guided reinforcement learning. SSRN.  
<https://doi.org/10.2139/ssrn.6423358>
- Contrino, M. F., Reyes-Millán, M., Vázquez-Villegas, P., & Membrillo-Hernández, J. (2024). Using an adaptive learning tool to improve student performance and satisfaction in online and face-to-face education for a more personalized approach. Smart Learning Environments, 11, Article 6.  
<https://doi.org/10.1186/s40561-024-00292-y>
- Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE.
- Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almotairi, S., Shutaleva, A., & Soomro, R. B. (2024). Extended TAM based acceptance of AI-powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study. Heliyon, 10(8), Article e29317.  
<https://doi.org/10.1016/j.heliyon.2024.e29317>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.  
<https://doi.org/10.2307/249008>
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human

- needs and the self-determination of behaviour. *Psychological Inquiry*, 11(4), 227–268.  
[https://doi.org/10.1207/S15327965PLI1104\\_01](https://doi.org/10.1207/S15327965PLI1104_01)
- Demartini, C. G., Sciascia, L., Bosso, A., & Manuri, F. (2024). Artificial intelligence bringing improvements to adaptive learning in education: A case study. *Sustainability*, 16(3), Article 1347.  
<https://doi.org/10.3390/su16031347>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). SAGE.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method* (4th ed.). Wiley.
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1–4.  
<https://doi.org/10.11648/j.ajtas.20160501.11>
- Fang, X., & Feng, J. (2026). Research on the application behavior of generative artificial intelligence learning of college students based on self-determination theory. *Frontiers in Psychology*, 17, Article 1805498.  
<https://doi.org/10.3389/fpsyg.2026.1805498>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50.  
<https://doi.org/10.2307/3151312>
- Galindo-Domínguez, H., Delgado, N., Urruzola, M.-V., Etxabe, J.-M., & Campo, L. (2025). Using artificial intelligence to promote adolescents' learning motivation: A longitudinal intervention from the self-determination theory. *Journal of Computer Assisted Learning*, 41(2), Article e70020.  
<https://doi.org/10.1111/jcal.70020>
- Goldstone, G. (2026, January 21). Designing the 2026 classroom: Emerging learning trends in an AI-powered education system. *Faculty Focus*.  
<https://www.facultyfocus.com/articles/teaching-with-technology-articles/designing-the-2026-classroom-emerging-learning-trends-in-an-ai-powered-education-system/>
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24.  
<https://doi.org/10.1108/EBR-11-2018-0203>
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modelling. *Journal of the Academy of Marketing Science*, 43(1), 115–135.  
<https://doi.org/10.1007/s11747-014-0403-8>
- Hidayat-ur-Rehman, I. (2024). Examining AI competence, chatbot use and perceived autonomy as drivers of students' engagement in informal digital learning. *Journal of Research in Innovative Teaching & Learning*, 17(2), 196–212.  
<https://doi.org/10.1108/JRIT-05-2024-0136>
- Jamali, H., Dascalu, S. M., Harris, F. C., Jr., & Wu, R. (2025). AI-powered adaptive learning interfaces: A user experience study in education platforms. *Frontiers in Computer Science*, 7, Article

1672081.  
<https://doi.org/10.3389/fcomp.2025.1672081>
- Katalinic, A., Slavuj, V., & Jaksic, D. (2026). Artificial intelligence in online education: A systematic review of its impact on learner engagement and satisfaction. *Education Sciences*, 16(3), Article 389.  
<https://doi.org/10.3390/educsci16030389>
- Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports*, 15, Article 17458. <https://doi.org/10.1038/s41598-025-97652-6>
- Koo, M., & Yang, S.-W. (2025). Likert-type scale. *Encyclopedia*, 5(1), Article 18.  
<https://doi.org/10.3390/encyclopedia5010018>
- Li, J., King, R. B., Chai, C. S., Zhai, X., & Lee, V. W. Y. (2025). The AI Motivation Scale (AIMS): A self-determination theory perspective. *Journal of Research on Technology in Education*. Advance online publication.  
<https://doi.org/10.1080/15391523.2025.2478424>
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 1–55.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioural research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903.  
<https://doi.org/10.1037/0021-9010.88.5.879>
- Ringle, C. M., Wende, S., & Becker, J.-M. (2024). *SmartPLS 4*. SmartPLS.  
<https://www.smartpls.com>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78.  
<https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, Article 101860.  
<https://doi.org/10.1016/j.cedpsych.2020.101860>
- Structural Learning. (2025, August 29). Desirable difficulties: Bjork’s 5 principles for teachers.  
<https://www.structural-learning.com/post/desirable-difficulties>
- Strielkowski, W., Grebennikova, V., Lisovskiy, A., Rakhimova, G., & Vasileva, T. (2025). AI-driven adaptive learning for sustainable educational transformation. *Sustainable Development*, 33(2), 1921–1947. <https://doi.org/10.1002/sd.3221>
- Wang, J. (2025). Students’ continuance intention to use generative AI in higher education: Integrating the Technology Acceptance Model and Information Success Model. In *Proceedings of the 2025 3rd International Conference on Educational Knowledge and Informatization* (pp. 115–124). Association for Computing Machinery.  
<https://doi.org/10.1145/3765325.3765346>
- Yang, Z., Deng, H., & Jiang, N. (2025). The impact mechanism of artificial

intelligence dependence on college students' innovation capability: An empirical study from China. *Frontiers in Psychology*, 16, Article 1732837. <https://doi.org/10.3389/fpsyg.2025.1732837>

Young, J. R., Bevan, D., & Sanders, M. (2024). How productive is the productive struggle? Lessons learned from a scoping review. *International Journal of Education in Mathematics, Science and Technology*, 12(2), 470–495. <https://doi.org/10.46328/ijemst.3364>

Zhao, X., Lynch, J. G., Jr., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37(2), 197–206. <https://doi.org/10.1086/651257>