The Potential of Adaptive Learning Systems to Enhance Learning Outcomes: A Meta-Analysis

by

Yizhu Gao

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Measurement, Evaluation, and Data Science

Department of Educational Psychology
University of Alberta

Abstract

Adaptive learning systems (ALSs) serve as a way to personalize learning experiences for students to improve their learning outcomes. Many studies have been conducted to develop and examine ALSs in terms of their effects on student learning outcomes compared to the traditional classroom settings. Comparisons of these studies revealed that different ALSs exhibited varying degrees of success in promoting learning achievement (i.e., discrepant magnitudes of system effectiveness). However, little work has empirically examined factors that impact the effectiveness of ALSs, thus constraining their application in practice. This study performed a meta-analysis of 46 studies on ALSs in order to identify factors significantly accounting for the variation in system effectiveness. The analyses of 77 effect sizes confirmed substantial heterogeneity in system effectiveness (*Mean* = 1.48; *Range* = .09-9.06). Using three-level correlated and hierarchical effects modeling, the heterogeneity was explained by the variability both within publications (i.e., 12.06%) and between publications (i.e., 83.47%). Specifically, the magnitudes of system effectiveness were significantly moderated by learner characteristics and modeling approaches. Moreover, in comparison with other subject areas such as mathematics and computer science, ALSs used to support the learning of the English language were likely associated with higher system effectiveness. No evidence of publication bias was detected in these data. Findings from the present research facilitate the understanding of what and how system components are associated with the effectiveness of ALSs and inform decision making on their design. Implications for pedagogical theories and practice, limitations of this research, and future directions for developing and implementing ALSs in educational settings are discussed.

Acknowledgments

First, I would like to express my deep appreciation to my supervisor, Dr. Ying Cui, for her constant support, guidance, and encouragement throughout my Ph.D. study. She not only mentored me to be an independent researcher but also set an example as a leading researcher in educational measurement. In addition, Dr. Ying Cui always brings fascinating insights into my study, research, and career development.

I would also like to thank my examining committee members, Dr. Carrie Demmans Epp, Dr. Hua-Hua Chang, Dr. Mark Gierl, and Dr. Maria Cutumisu, for providing insightful comments for my dissertation and Dr. Denyse Hayward for hosting my defense. Although I have few opportunities to collaborate with my committee members, their wisdom and immense knowledge greatly improve my dissertation, critical thinking, and beyond.

I also like to thank Dr. Xiaoming Zhai and Dr. Okan Bulut that guide me and collaborate with me to do interesting and innovative research. My first academic article published in 2018 was guided by Dr. Xiaoming Zhai. He taught me a lot such as what valuable research means and how to develop an article. Dr. Okan Bulut always brought cutting-edge research directions and resources. I also want to thank CRAMERs that were always willing to offer help and company.

Finally, I would like to thank my parents and friends (Lijing Wang and Luping Xiao), for their unconditional love and support on this journey.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Is it possible for a computer to mimic a human teacher to provide personalized instruction? The development of adaptive learning systems (ALSs) seems to have the potential to provide a positive answer to this question. However, what one might ask is whether ALSs advance learning outcomes as effectively as one-on-one human instruction. This unsettled question deserves further exploration and motivates the present research.

ALSs, which are also called adaptive learning environments, are often defined as digital learning systems or environments that adapt learning content, presentation styles, or learning paths based on individual learner characteristics (Tseng et al., 2008; Yang et al., 2013). The goal of ALSs is aligned with exemplary instruction (Shute & Towle, 2003): delivering the right content to the right person at the proper time in the most appropriate way – any time, any place, any path, any pace (National Association of State Boards of Education, 2001). That is, ALSs are designed and developed to achieve personalized learning.

The idea of personalized learning can be traced back to John Dewey's enduring work in learner-centered education in the early 20th century (Redding, 2016; Zhang et al., 2020). For example, Dewey (1929) pointed out that effective education must begin with understanding how learner capacities, interests, and habits can be directed to help individuals succeed. After decades of development in learning sciences, educational research, and artificial intelligence (AI), personalized learning has been assimilated into modern education dramatically (Mohd & Shahbodin, 2015; Muñoz et al., 2022; Zhang et al., 2020). Nowadays, personalized learning is used to describe learning environments in which learning objectives and content along with learning methods and pace may all vary depending on individual learners (U.S. Department of Education Office of Educational Technology, 2010).

Traditional classroom settings could be used to offer personalized learning environments. For example, DeMink-Carthew et al. (2017) conducted a qualitative study and explored goal-setting approaches as a whole-school initiative on personalized learning. They found that learning goals that engage students in exploring their personal interests would contribute to the construction of personalized learning environments. However, achieving personalized learning within traditional classroom learning environments (e.g., empowering each student as a co-designer of their own learning) is too costly for most societies to bear on a large scale. In addition, systematic guidelines for adding personalized learning elements to traditional classroom settings are scant (DeMink-Carthew et al., 2017).

Unlike traditional classroom settings, ALSs present a new opportunity for accomplishing personalized learning at a massive scale (Kerr, 2016; Liu et al., 2017). Enabled through technological advances and methodological innovation, various ALSs have been established and used in real education scenarios. For example, one of the well-known ALSs is DreamBox Learning, which teaches mathematics to elementary students. In response to the opportunity that ALSs provide, education departments of many countries have formulated action plans (Li et al., 2021). For example, the U.S. Department of Education released the report “Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics” to realize personalized learning with the help of ALSs (U.S. Department of Education, 2012).

ALSs commonly invoke three stages (Alfonseca et al., 2006; Brusilovsky & Maybury, 2002): retrieving information about learners, processing the information to initialize and update a learner model, and providing adaptive effects based on the learner model. For example, Tseng et al. (2008) designed a two-source adaptive learning system to help junior high school students learn mathematics. Specifically, students’ learning styles (e.g., processing information sequentially) and learning behaviors (e.g., learning achievement, learning effectiveness) were analyzed and recorded in its learner model. Then, the two sources of individual information were exploited to adjust presentation styles and difficulty levels of learning materials. Students who lack sequential processing skills receive subject materials presented in a non-linear manner, at different difficulty levels. In contrast, for students who prefer to process information sequentially and are with high learning effectiveness and test score, the system provides learning materials with increased difficulty levels, in a sequential frame.

Studies on ALSs typically evaluate system effectiveness in terms of learning outcomes improvement. For example, Tseng et al. (2008) examined the effectiveness of their system in enhancing student learning performance in mathematics. After comparing test scores of students under adaptive learning or non-adaptive learning approaches, Tseng et al. (2008) confirmed the high utility of the system in promoting student learning achievement. As Tseng et al. (2008) found, the benefits of ALSs to student learning outcomes have been supported by research published in influential mainstream journals (e.g., Liu et al., 2017; Hwang et al., 2020; Yang et al., 2013). To maximize the benefits of ALSs to student learning process, researchers from both academia and industry attempt to develop new systems by incorporating cutting-edge techniques (e.g., conversational AI). As a result, hundreds of ALSs have been established over the last three decades. Systematic reviews or meta-analyses, as powerful research methods of comparing, analyzing, and combining results from different studies, have enormous potential to enlighten our understanding of established ALSs.

There have been 26 systematic reviews of research on ALSs. Their titles and publication years have been presented in Appendix A. Figure 1 shows the total number of reviews year by year. As indicated by Figure 1, scholarly efforts in reviewing and critically assessing research on adaptive learning have grown significantly since 2015. On closer inspection, most studies focus on qualitative analyses of the current ALSs. For example, a recent study by Martin et al. (2020) reviewed designs, context, strategies, and technologies of ALSs from publications between 2009 and 2018. As the main goal of ALSs is to ensure personalized learning experiences, it is becoming essential to know their effects on learning outcomes improvement (Dziuban et al., 2017, 2018; Essa & Laster, 2017). Although there are different categories in terms of learning outcomes, this study aims to examine learners' advances in cognition such as knowledge achievements and high-order competence. Resulting changes in learners' cognition can be evaluated using an exam specifically designed to evaluate it. However, little work has systematically examined the effectiveness of the existing ALSs in advancing student learning outcomes, constraining their continual improvement and applications in educational scenarios. It is worth noting that the term "effectiveness" may be defined as "learners achieve targeted learning goals using shorter time" or "learners increase their learning motivation" in some studies. However, for simplicity's sake, "effectiveness" specifically denotes the degree of improvement in learning outcomes that learners obtain after using ALSs in this research.

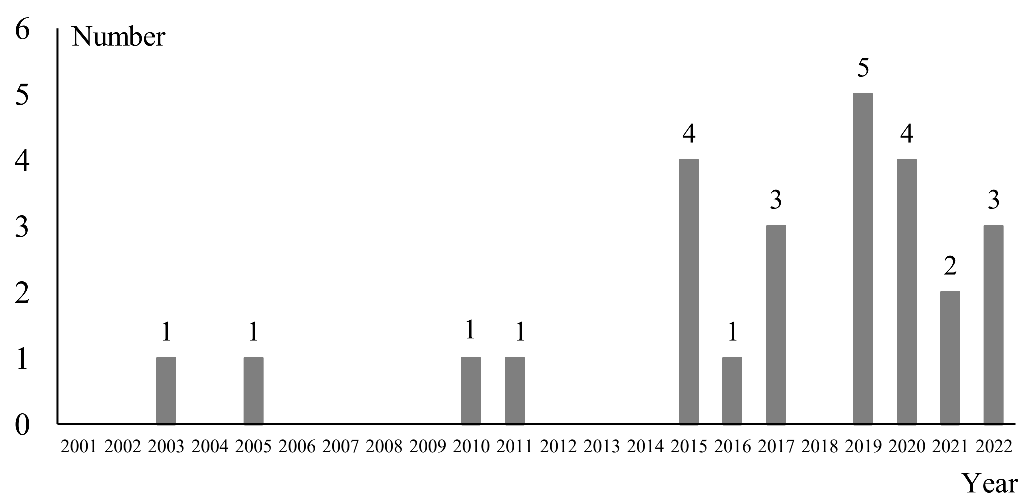

Figure 1. *Number of Literature Reviews on Adaptive Learning from 2000 to 2022*

In addition, as efforts to leverage the power of ALSs to students' learning grow, notable disparities in system effectiveness have emerged, raising the question of what factors account for the heterogeneity. For example, Verdú et al. (2008) conducted a literature review and identified a wide range of the effectiveness of ALSs in improving student learning outcomes: The effect sizes (i.e., Cohen's *d*) of fifteen ALSs from publications between 1997 and 2007 ranged from 0.10 to 3.86. This broad range indicates there are considerable challenges in getting the full benefits of ALSs with respect to promoting student learning outcomes. A better understanding of factors contributing to the heterogeneity is clearly needed. To fill this research gap, this study seeks to identify factors that significantly account for the variability in system effectiveness. However, directly recognizing influential factors is difficult because different ALSs are often equipped with not attuned system contents (Vandewaetere et al., 2011). To overcome this obstacle, this study starts with a system component that is well documented in the literature: learner modeling (Abyaa et al., 2019; Brusilovsky et al., 2004).

Learner modeling denotes the detailed monitoring of learner characteristics within ALSs (Abyaa et al., 2019). The resulting learner models are computational representations of learners' cognitive and non-cognitive characteristics, which inform adaptive effects provided by ALSs. That is, learner modeling acts as the footing of system adaptation. There have been a large number of reviews that specifically focus on learner modeling of ALSs (e.g., Abyaa et al., 2019; Vandewaetere et al., 2011). For example, Abyaa et al. (2019) categorized learner characteristics modeled in studies published between 2013 and 2017 into six groups: learner profile, knowledge, cognitive characteristics, social characteristics, personality traits, and motivation. They also grouped modeling approaches into five categories: clustering and classification, predictive modeling, overlay modeling, uncertainty modeling, and ontology.

Most ALSs model more than one learner characteristic for constructing learner models (e.g., Alshammari et al., 2015), while in some ALSs, only one learner characteristic is modeled in their learner models (e.g., Kabudi et al., 2021; Surjono, 2011). The choice of learner characteristics that are modeled in ALSs depends on system designers and developers. There is a broad consensus on the relative frequency of learner characteristics that have been modeled in the existing ALSs. For example, Nakic et al. (2015), Martin et al. (2020), and Muñoz et al. (2022) found that learning style is the commonly modeled learner characteristic in the present established ALSs. Although learning style is frequently modeled as a source of personalizing learning experience in ALSs, it is unclear how the selection of learner characteristics affects system effectiveness. Particularly, it is necessary to investigate whether ALSs that model learning style are likely to have higher system effectiveness. If they are, learning style indeed deserves more attention than other learner characteristics as it is currently the case in practice.

In addition, there exist various modeling techniques to construct learner models and multiple approaches are available to model a specific learner characteristic. For example, among predictive modeling techniques, item response theory (e.g., Chen et al., 2005; Mohamed et al., 2012) and Bayesian knowledge tracing (e.g., Pardos & Heffernan, 2010) are ready to model learners' prior knowledge. In terms of learning style, there are multiple questionnaires (i.e., Kolb's learning style scale and Jackson's learning styles profiler questionnaire) that can be used to profile learners' learning styles in ALSs. The more precisely learner characteristics are modeled, the more fitly adaptive instructions will be delivered (Abyaa et al., 2019; Truong, 2016). That is, the option of modeling approaches is likely to be related to the degree of success that ALSs customize instruction to suit individual learners. However, it is unknown which modeling approaches are more likely to contribute to more effective ALSs.

To sum up, of particular interest for this study is the debate about whether influences of different learner characteristics and modeling approaches are sufficiently robust to be taken into account, out of consideration for system effectiveness. If yes, identifying learner characteristics and modeling approaches that significantly contribute to system effectiveness will provide practical guidelines for advancing the field of ALSs. The lack of effectiveness evidence and the difficulty in reusing successful design practices (e.g., the combination of a specific learner characteristic and a modeling approach or the assembly of different student characteristics) constitute barriers for ALSs to fulfill their full potential in the context of modern educational practice (Sunitha et al., 2011).

Research Purpose

Three research questions guide this work: (a) How effective are existing ALSs in promoting student learning outcomes? (b) Is the system effectiveness heterogeneous? (c) If so, which factors significantly account for the heterogeneity of system effectiveness? For the third research question, the effect of learner modeling on system effectiveness along with other potential factors was examined. To answer these questions, this research performed a meta-analysis of studies on ALSs.

Organization of the Dissertation

This dissertation is organized into five chapters. An overview of each chapter is described as follows. The current chapter, Chapter 1, is an introduction to the relevant research area as well as a justification for the research questions investigated in this study.

Chapter 2, a literature review, gives an overview of ALSs. Particularly, it presents definitions of adaptive learning, the development history of ALSs, two frameworks that are commonly applicable to different ALSs, and previous reviews (i.e., literature reviews and meta-analyses) on ALSs. In addition, Chapter 2 provides detailed descriptions of learner modeling, which correspond to the common component between the two frameworks of ALSs.

Chapter 3 describes the methods of this meta-analysis study. It starts with detailed procedures to select candidate studies to be included in this meta-analysis and variables to be coded for the selected studies. Then, different types of effect sizes are introduced and compared, including calculations of effect sizes and their variances. Finally, statistical methods to answer each research question as well as publication bias examination are explained.

This is followed, in Chapter 4, with the results of the meta-analysis. Results include a mapping of the main characteristics of included studies, publication bias examination, the heterogeneity and the overall system effectiveness analysis, and moderator effects investigation.

Chapter 5 covers a discussion of the findings, theoretical and practical implications, limitations, and future directions. The results of this study open up discussions on the effectiveness of the existing ALSs and relevant influential factors. Also, they serve as evidence-based recommendations for future developments of ALSs. This chapter also presents implications for pedagogical theories and practice, followed by some limitations facing the present study and possible directions for future studies. This chapter concludes with a summary of the entire research process and significant findings.

Chapter 2 Literature Review

What is Adaptive Learning?

Learning has been widely recognized as a personalized experience in which learners progressively expand their personal knowledge, perspective, skills, and understanding (Shemshack et al., 2021; Shemshack & Spector, 2020). For instructions to be maximally effective for learning, instructors should capitalize on learner characteristics. In other words, enhancing learning performance is a function of adapting instructions to suit individuals best. These statements have been evidenced by the research field of aptitude-treatment interactions (Cronbach & Snow, 1977; Shute, 1993), which regards learner characteristics as important considerations for constructing learning environments and optimizing learning outcomes.

Research on aptitude-treatment interactions in recent years, which serves individual learners with customized instruction within digital learning environments, naturally leads to the emergence of adaptive learning (Shute & Towle, 2003). For example, Mödritscher et al. (2004) suggested limiting the control over learning processes for learners with low-prior knowledge or enhancing such control for learners who have high performance. In addition, Sugawara et al. (2020) examined the relationship between learning effects and learning types using a case of e-learning. They found that e-learning may have little effect when the e-learning system does not match learners' learning types (e.g., short-term learning and long-term learning type).

Importantly, the field of adaptive learning scales up personalized learning experiences at high efficiency. As AI and big data are booming in education, ALSs are able to dynamically facilitate learning processes by automatically monitoring learner characteristics and inferring learner preferences (Shute & Zapata-Rivera, 2008; Vandewaetere et al., 2014). Thus, adaptive learning has been regarded as a critical technology-empowered learning approach and a research hotspot in the field of AI in education (AI-Ed).

Since adaptive learning is sometimes used interchangeably with the term personalized learning (Shemshack & Spector, 2020), a comparison could help to build a better understanding of these terms. Although personalized learning has been used for hundreds of years in the form of apprenticeship and mentoring (Shemshack & Spector, 2020), it has followed different pathways. For example, Jenkins and Keefe (2002) described a couple of basic elements of personalized instruction at school (e.g., teachers maintain both coach and advisor roles, and the schedule of a school is flexible). Following these guidelines, schools create a caring and collaborative environment and value student diversity and individual development. However, small teacher-student ratios in school settings seem to be a major obstacle to making learning experiences personalized for individual students without technology (Lee et al., 2018).

As educational technologies began to mature in the last half of the previous century, personalized learning took the form of intelligent tutoring systems. The rise of hypermedia and the World Wide Web (WWW) was poised to transform personalized learning once again, leading to the advent of adaptive educational hypermedia systems. More recently, personalized learning is increasingly defined as a technology-based instructional model (Bingham et al., 2018; Walkington & Bernacki, 2020). In accordance with this tendency, adaptive learning can be depicted as a scalable personalized learning approach in essence (Li et al., 2021). Compared to approaches not involving technologies, adaptive learning is able to ensure personalized learning experiences for large and heterogeneous groups of learners. More specifically, ALSs, as important carriers of adaptive learning, overlap with technology-based approaches to personalized learning. That is, intelligent tutoring systems and adaptive educational hypermedia systems, which are reviewed in the section that follows, have long been regarded as technologies that stimulate ALSs (Beldagli & Adiguzel, 2010; Surjono, 2011).

Adaptive learning has numerous definitions and various forms of implementation in the literature. For example, Kerr (2016, p. 88) describes adaptive learning as an educational technology that delivers learning materials based on learners' interaction with previous content automatically, dynamically, and interactively. In contrast, Lowendahl et al. (2016) define adaptive learning as a process that "dynamically adjusts the way that instructional content is presented to students based on their comprehension of the material as revealed in their responses to embedded assessments or learner preferences such as visual presentation of materials" (p. 7). Whether defined as a technology or a process, adaptive learning essentially represents unique learning experiences accounting for individual characteristics to ensure personalized learning experiences (Liu et al., 2017; Rosita Cecilia et al., 2016).

A common way of carrying out adaptive learning is to develop web-based adaptive learning environments as stand-alone learning platforms. Thus, ALSs are also called adaptive e-learning systems or computer-based adaptive learning environments. With the popularity of learning management systems (e.g., Moodle), researchers offer a new framework that brings learning management systems and ALSs together (Jagadeesan & Subbiah, 2020; Qazdar et al., 2015). A learning management system is a web-based software application that is designed to present learning content, assessment tools, and reports of learning progress and student activities (Kasim & Khalid, 2016). Generally, there is no adaptation component in learning management systems. The resulting adaptive learning management systems are advantageous for both institutes (e.g., universities) and students. Nonetheless, different forms of ALSs are characterized by system adaptivity, even though they may differ in the degree of how advanced adaptation that systems can deliver. System adaptivity refers to the capability of a computer-based system or environment to monitor important learner characteristics, behaviors, and performance and thereupon provide adaptive instruction (Leutner, 2004; Shute & Zapata-Rivera, 2008). Accordingly, ALSs are computerized learning systems that are equipped with system adaptivity for supporting student learning. A relevant but different concept from system adaptivity is system adaptability (also called learner control or user control). Chou et al. (2015) described adaptability as “systems provide an adaptable framework, tools, or choices to enable learners to adapt content sequences, pacing, context, task difficulty, and learning supports per their needs and preferences”. In the present study, system adaptation indicates system-controlled adaptation (i.e., system adaptivity), unless otherwise specified.

Tracing the developmental history of ALSs is beneficial to understand what an ALS is and how it is different from other educational technologies. The following section begins with two intertwined technologies (i.e., intelligent tutoring systems and adaptive educational hypermedia systems) that collaboratively boost the development and prosperity of ALSs. They are presented in chronological order in the section that follows.

Development History of Adaptive Learning Systems

Intelligent Tutoring Systems

Intelligent tutoring systems emerged in the 1980s and aimed to instruct learners in an intelligent way. There is no generally agreed single definition of what it means to tutor “intelligently”. However, a characteristic shared by many intelligent tutoring systems is that they infer models of individual learners’ current understanding of a subject and use this model to provide individualized instructions (Ma et al., 2014). For example, SCHOLAR, which is often regarded as the first intelligent tutoring system (Carbonell, 1970), constructed questions on given topics, carried on contextual dialogues with learners, and provided feedback on the correctness of learners' responses to questions about the geography of South America.

With respect to intelligent tutoring systems, researchers have articulated a four-component conceptual structure: an interface, a domain model, a learner model, and a tutoring model. These four structural components are generally accepted and remarkably resilient, even as intelligent tutoring systems themselves vary significantly in their contents (Dede, 1986; Hartley & Sleeman, 1973). The function of an interface is to communicate with learners by presenting and receiving information. To some degree, an interface determines what operations learners can make in responding to questions and seeking information. A domain model includes knowledge that learners intend to grasp (Nwana, 1990). A set of logical propositions, production rules, or any suitable knowledge representation format can be defined in a domain model. A learner model of intelligent tutoring systems represents relevant aspects of learners' traits, which is determined by learners' responses to questions or their interactions with the interface. A tutoring model represents instructional strategies such as offering a hint or cue when learners are unable to generate a correct response or assigning a problem that requires knowledge only slightly beyond the current knowledge level (Ma et al., 2014).

Instead of emphasizing the structure of intelligent tutoring systems, VanLehn (2006) summarized their behavior. Specifically, there are two loops embedded in intelligent tutoring systems: the outer loop and the inner loop. There are four common types of the outer loop (VanLehn, 2006): (a) the learner selects a task from a menu of all tasks; (b) the system assigns tasks in a predetermined sequence; (c) the system assigns tasks from a unit's pool of tasks until the learner has mastered the knowledge taught by the unit; (d) the system tracks learners' traits such as learning styles and mastered knowledge components and chooses a task based on the match between the task's traits and the learner's traits. The outer loop executes once for each task. In contrast, the inner loop executes once (e.g., give feedback and hints) for each step taken by the learner when solving a task (VanLehn, 2006). The inner loop can also assess the learner's changing competence and update a learner model, which is used by the outer loop to select the next task that is appropriate for the learner.

For example, the Basic Instructional Program (BIP), as an example of earlier intelligent tutoring systems (Barr et al., 1976), offered tutorial assistance to learners solving introductory programming problems. Its domain model presented curriculum structure and also constructed a domain representation that mapped target skills to programming tasks. Learners' performance on tasks supported inferences about their acquisition of skills linked to these tasks. The problem presentation sequence is individualized based on curriculum structure and individual learners' state of knowledge. A recent example of an intelligent tutoring system is a C# language intelligent tutoring system developed by Al-Bastami and Naser (2017). After selecting a lesson that learners intend to learn, they are presented with many questions relevant to the chosen lesson. Based on learners' performance on questions, they are suggested to go back to easier lessons (i.e., less than 50% mark), repeat exercises within the same lesson (i.e., between 50% and 70%), or move to more difficult lessons (i.e., more than 70% mark). Learning materials relevant to C# language topics are stored in its domain model. In its tutoring model, learners can practice questions generated at every difficulty level of each lesson.

As the field of AI techniques develops, intelligent tutoring systems are capable of customizing instructional activities and strategies based on learners' characteristics and needs, which are not limited to knowledge levels (e.g., Cha et al., 2006; Malekzadeh et al., 2015). Although learner knowledge is the most common factor that is modeled in intelligent tutoring systems, other characteristics such as learning styles, activities, and affect states are becoming increasingly prevalent recently. Overall, modeling learner characteristics to build learner models and to provide individualized instruction is the essential feature that distinguishes intelligent tutoring systems from earlier computer-based instructional systems (e.g., computer-assisted instruction systems; Ge et al., 2012; Ma et al., 2014).

The number of systematic reviews and meta-analyses of intelligent tutoring systems is mounting in recent ten years. For example, Mousavinasab et al. (2021) conducted a systematic review of characteristics, applications, and evaluation methods of intelligent tutoring systems from 2007 to 2017. They found that action-condition rule-based reasoning, data mining, and Bayesian network were the most frequent AI techniques used in intelligent tutoring systems. In addition, the majority of intelligent tutoring systems were evaluated based on learners' performance after using the systems. Positive evidence for the effectiveness of intelligent tutoring systems on student academic learning accumulated over the years. For example, Ma et al. (2014) conducted a meta-analysis of intelligent tutoring systems and found that there are positive effect sizes at all levels of education and in almost all subject domains. Kulik and Fletcher (2016) further clarified the median effect of intelligent tutoring systems on learning performance is 0.66 standard deviations higher than conventional teaching.

Adaptive Educational Hypermedia Systems

At the end of the 20th century and the beginning of the 21st century, the combination of AI and hypermedia technology produced a new set of systems, namely adaptive hypermedia systems. According to Brusilovsky (1996), an adaptive hypermedia system should satisfy three criteria: (a) it should be a hypertext or hypermedia system; (b) it should have a user model; (c) it should be able to adapt based on the user model. There are various application areas of adaptive hypermedia systems (e.g., online information systems), among which the area of adaptive educational hypermedia systems is most popular (Brusilovsky, 2001).

Inspired by intelligent tutoring systems, adaptive educational hypermedia systems try to combine adaptive instructional systems and hypermedia-based systems to produce applications in which learning content, link structures, and presentation styles are dynamically adapted to learner characteristics. A typical architecture of adaptive educational hypermedia systems consists of two layers (see Figure 2; Nguyen & Do, 2008): runtime layer and storage layer. Like an interface of an intelligent tutoring system, the runtime layer of an adaptive educational hypermedia system is also responsible for presenting adaptive learning materials to learners and observing learners' actions (Nguyen & Do, 2008).

The storage layer is the main engine that controls the adaptive process with four sections: (a) a learner model describes information and data about an individual learner, such as knowledge status and learning style preferences. Information gathered from the runtime layer can also be used to update the learner model; (b) a media space contains learning resources and associated descriptive information; (c) a domain model describes the structure of domain knowledge; (d) an adaptation model contains concept selection rules and content selection rules. Concept selection rules are applied to select an appropriate concept from the domain model. With regards to content selection rules, they are used to choose a suitable educational resource from the media space.

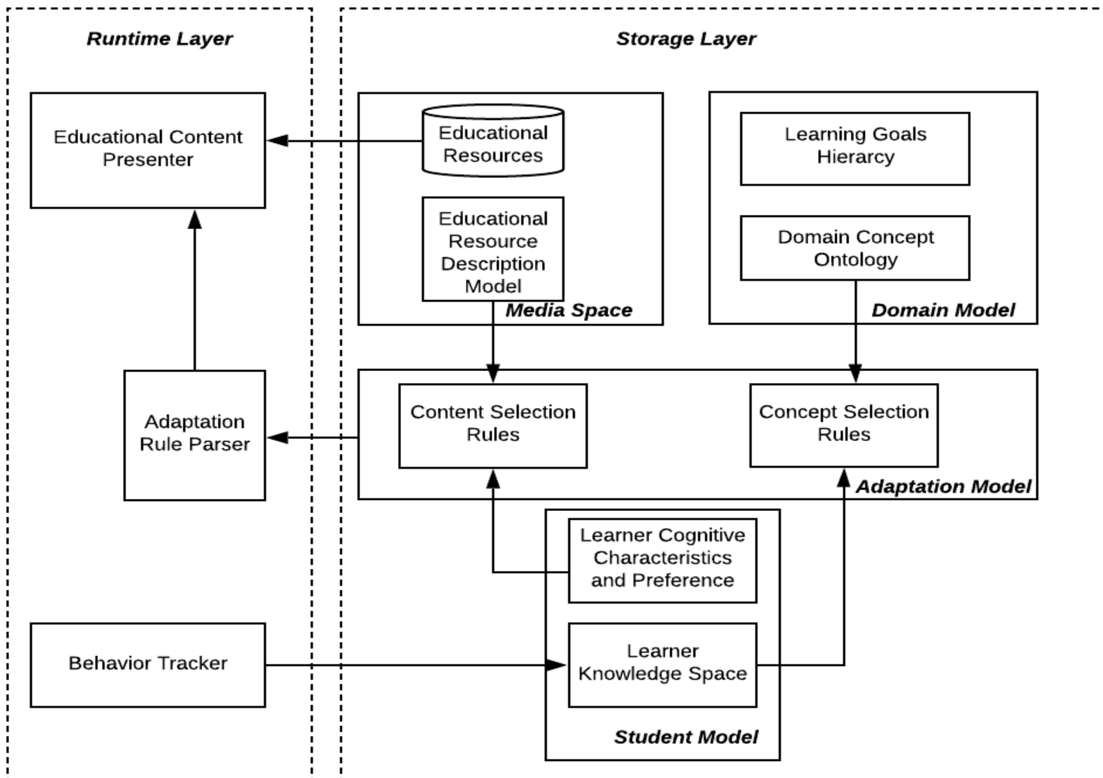

Figure 2. *The General Architecture of Adaptive Educational Hypermedia Systems*

Note: This model was produced by Nguyen and Do in 2008. From "Learner model in adaptive learning", by L. Nguyen and P. Do, 2008, *World Academy of Science, Engineering and Technology, 21*, p. 396.

An adaptive e-learning hypermedia system based on learning styles, which is developed by Mustafa and Sharif (2011), is a good example of an adaptive educational hypermedia system. In the system, learners' learning styles and knowledge levels were recognized and stored in its learner model. Specifically, the system used Fleming's visual, aural, read/write, and kinesthetic learning style model (Fleming, 2001) to classify learners into four learning style categories: learners in a visual category prefer to receive information in maps, diagrams, flow charts, and all symbolic arrows; learners in an auditory category have a preference for information that is heard or spoken; learners in a read/write category tend to receive information displayed as text; learners in a kinesthetic category prefer to use experience and practice to obtain information. In addition, the system used an evaluation quiz to identify learners' knowledge about each domain knowledge concept. Accordingly, the system is typified by two kinds of adaptations: learning style adaptation and prior knowledge adaptation.

The learning style adaptation regulates which object concepts from resource space are chosen for learners with specific types of learning styles. The knowledge adaptation mechanism determines which concepts from knowledge space to be covered based on knowledge attributes in its learner model. Object concepts are indexed to knowledge attributes and are also labeled with text, audio, visual, and kinetic values. It is worth noting that lesson contents appear in its navigation area as a tree-like structure of hyperlinks, whilst learning content is presented with the media matched for learners' preference in its content area.

There are some literature reviews of adaptive educational hypermedia systems (Akbulut & Cardak, 2012; Al-Azawei & Badii, 2014). For example, Akbulut and Cardak (2012) conducted a content analysis of seventy studies on adaptive educational hypermedia systems that accommodate learning styles. They found that the majority of studies proposed a framework or model for adaptation whereas few studies investigated the effectiveness of learning style-based adaptive educational hypermedia systems. The insufficiency of studies with empirical evaluations of systems might explain the absence of meta-analytic studies in this field.

Comparisons Between the Two Types of Systems

Intelligent tutoring systems and adaptive educational hypermedia systems are two major research streams in the field of ALSs. Intelligent tutoring systems provide problem solving support by tracing learners' actions and responses (Desmarais & d Baker, 2012). Certain limitations exist such as they may lack requisite learning material, their regular application is in well-structured domains (e.g., geometry, programming, physics, and algebra; Anderson, 2000), and it is hard to formalize learning processes including metacognition and reflection. Adaptive educational hypermedia systems research was motivated by addressing the issues of intelligent tutoring systems and some problems related to online learning such as cognitive overload and disorientation (Brusilovsky & Peylo, 2003).

In the narrow sense, tutoring models of intelligent tutoring systems may behave differently from adaptation models of adaptive educational hypermedia systems. Intelligent tutoring systems assist learners in solving a problem by offering moment-to-moment hints, cues, or prompts on any step of a solution or a full solution. However, adaptive educational hypermedia systems usually utilize the whole information in learner models and guide learners through link, learning content, or presentation mode adaptation.

VanLehn (2006) disambiguated intelligent tutoring systems from other learning systems with inner-loop and outer-loop adaptations. The inner-loop adaptation means systems help learners while they are working on a given learning task. Typically, the inner-loop adaptation performs step-specific scaffolding strategies, such as error-correction feedback and cues that orient students to essential parts of the current task state. The outer-loop adaptation, in contrast, determines the next learning task that learners will perform, such as a problem to solve or a learning material to read. As stated by Nye (2015), the inner-loop adaptation is a defining characteristic of intelligent tutoring systems, whereas other learning systems (e.g., adaptive educational hypermedia systems) usually employ the outer-loop adaptation. These differences correspond to their diverse functions: intelligent tutoring systems generally assist in the use of concepts to solve problems whereas adaptive educational hypermedia systems are better suited for the instruction of subject concepts (Phobun & Vicheanpanya, 2010).

However, with computational advances, the scopes and objectives of intelligent tutoring systems and adaptive educational hypermedia systems cannot be sharply demarcated. Their advantages in personalized learning are gradually merged to provide a full learning environment (Phobun & Vicheanpanya, 2010). For example, Brusilovsky and Peylo (2003) proposed adaptive and intelligent web-based educational systems (AIWBESs) as a new and exciting stream of work in the AI-Ed field. Both intelligent tutoring systems and adaptive educational hypermedia systems are considered as classic AIWBES technologies. Evolving from AIWEBSs, ALSs frequently introduce new means of adaptive mechanisms (e.g., educational data mining, learning analytics) into commonly adaptive learning facilities (e.g., adaptive educational hypermedia systems or some intelligent tutoring systems). In this regard, adaptive educational hypermedia systems and intelligent tutoring systems collaboratively level up the development of ALSs. Thus, ALSs could be adaptive educational hypermedia systems or intelligent tutoring systems with adaptive functionality (Abyaa et al., 2019; Nguyen & Do, 2008; Vandewaetere et al., 2011).

Common Structures of Adaptive Learning Systems

Constructing an ALS is labor-intensive (Murray, 1999). Although designers and researchers aim to build ALSs that can greatly advance users' learning, notable disparities come up in terms of system effectiveness in improving learner learning outcomes. For example, Yang et al. (2013) found a significantly better learning achievement when learners applied an ALS in a computer science course. However, Liu et al. (2017) did not find significant improvement in learners' learning in biology, chemistry, and information literacy using an ALS. The incongruous performance might be because different ALSs provide varying degrees of adaptations depending on the subjects (Koedinger et al., 2012) and/or knowledge components (Aleven & Koedinger, 2013). Thus, identifying factors that would affect system effectiveness is becoming increasingly critical (Brusilovsky et al., 2004). Nevertheless, studies on ALSs typically examine the effectiveness of one ALS only and sparsely involve comparisons (Vandewaetere et al., 2011). This in turn leads to difficulties in recognizing influential factors of system effectiveness. Common grounds of diverse ALSs, which have been recognized by prior studies (Brusilovsky et al., 2004; Vandewaetere et al., 2011), served as a useful starting point for further investigation of influential factors of system effectiveness. They are Brusilovsky et al.'s (2004) layered evaluation structure and Vandewaetere et al.'s (2011) tripartite structure.

Layered Evaluation Structure

Brusilovsky et al. (2004) pointed out the success of an ALS is addressed at two distinct layers: learner modeling and adaptation decision-making.

Learner Modeling. For an ALS, the process of gathering relevant information in order to infer cognitive and non-cognitive state of a learner is defined as learner modeling (Abyaa et al., 2019; Thomson & Mitrovic, 2009). Learner models are computational representations of learners' characteristics so as to be accessible and useful to ALSs. Imagine a learner model as an avatar of a real learner in the virtual world, the contents of learner models correspond to the characteristics of a real learner (Yang et al., 2013).

Adaptation Decision-Making. Based on learner models derived from the learner modeling phase, adaptation decisions are made to generate specific adaptive effects (i.e., to behave differently for different learners). The logic of adaptation decision-making is often captured into a set of adaptation rules that determine which adaptation constituent(s) should be selected. Adaptive effects could be adaptive presentations, adaptive navigations, and adaptive content aggregations (Esichaikul et al., 2011; Premlatha & Geetha, 2015).

The adaptive presentation indicates adapting the presentation of information to individual learners in accordance with their characteristics. That is, a hypermedia page is individually assembled for each learner. Common technologies of adaptive content presentation include conditional text, stretch text, and page variables (Nguyen & Do, 2008). The adaptive navigation refers to manipulation of a linkage structure to guide individuals to find learning content. Popular technologies of adaptive navigation include direct guidance, adaptive link hiding, adaptive link sorting, adaptive link annotation, adaptive link generation, and map adaptation (Nguyen & Do, 2008). The adaptive content aggregation adapts learning content to learners' knowledge, goals, and other features (Premlatha & Geetha, 2015).

For example, Yang et al. (2013) developed an ALS based on learners' learning styles and cognitive styles. They utilized learning styles to achieve adaptive content aggregation, whereas cognitive styles are used to deal with adaptive navigation modes. In particular, the ALS provides learners with a visual learning style with more visual learning materials (e.g., diagrams, sketches, photographs); the ALS serves learners with a sensing learning style with more specific examples of concepts and illustrations of how concepts can be applied to practical applications. In addition, learners with a field-dependent cognitive style are supplied with links to the current learning content. This separate presentation form aims to avoid distractions from learning content. On the contrary, learners with a field-independent cognitive style are served with other relevant information at the same time to help them make a comprehensive inspection of learning content.

To summarize, learner modeling generates and updates learner models, based on which adaptation decision-making process brings about adaptive effects (e.g., adaptive presentation or adaptive navigation, see Figure 3). In this study, the learner profile is defined differently from the learner model. This is also the view held by Nguyen & Do (2008): learner profile describes basic information about a learner (e.g., gender, class, questionnaire response); in contrast, learner model depicts references about learner characteristics based on information stored in learner profile and/or learner interactions with learning systems. Keeping track of previous ALSs, it can be noted that the same learner model can produce varied adaptation decisions. For example, Yang et al. (2013) provided adaptation navigation based on learners' cognitive styles (i.e., field dependence/independence). However, Triantafillou et al. (2004) adapt presentation modes to learners' cognitive styles (i.e., field dependence/independence). The independence of the two layers enables us to evaluate them independently (Brusilovsky et al., 2004).

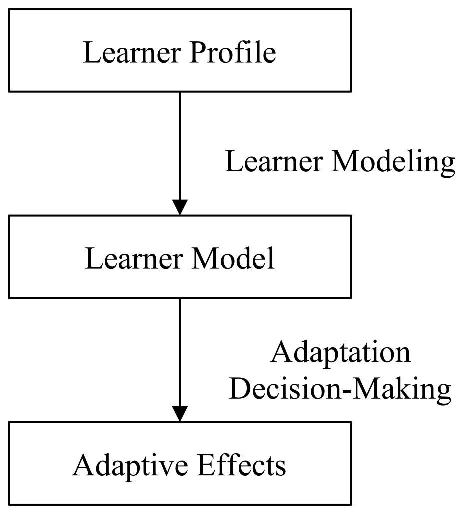

Figure 3. *Two Layers Decomposed from Adaptive Learning Systems*

Note: This structure was produced by Brusilovsky et al. in 2004. From "Layered evaluation of adaptive learning systems", by P. Brusilovsky, C. Karagiannidis, and D. Sampson, 2004, *International Journal of Continuing Engineering Education and Life Long Learning, 14*(4–5), p. 406.

Tripartite Structure

After reviewing a large number of studies on ALSs, Vandewaetere et al. (2011) extracted underlying building blocks across different ALSs, which constitute a tripartite structure (see Figure 4). The first component is the source of adaptive instruction, the second component refers to the target of adaptive instruction, and they are connected to each other by the third component—pathway.

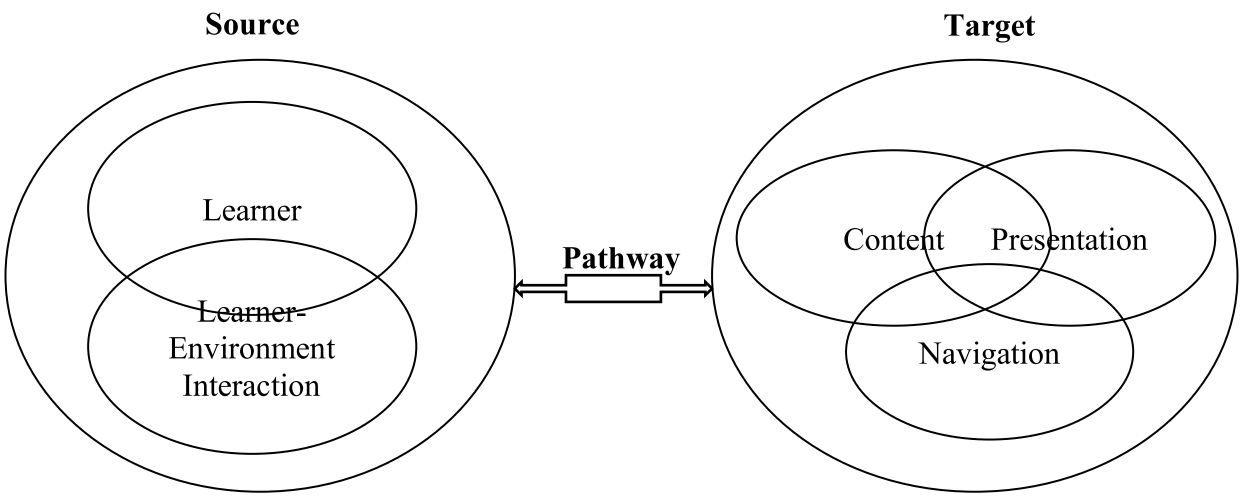

Figure 4. *Tripartite Structure of Adaptive Learning Systems*

Note: This structure was produced by Vandewaetere et al. in 2011. From "The contribution of learner characteristics in the development of computer-based adaptive learning environments", by M. Vandewaeter, P. Desmet, and G. Clarebout, 2011, *Computers in Human Behavior, 27*, p. 122.

The Source and Target of Adaptive Instruction. The source of adaptive instruction, as a starting point from where adaptation occurs, determines what type of adaptation an ALS can provide (Siddique et al., 2019). It answers the question "To what will be adapted in an ALS?" In contrast, the target of adaptive instruction responds to the question "What will be adapted in an ALS?" Specifically, the target comes in many forms such as adaptive content, adaptive presentation, and adaptive navigation.

The Pathway of Adaptive Instruction. The pathway of adaptive instruction targets to address the question "How to translate source into target?" Vandewaetere et al. (2011) called the route from the source to the target as a modeling process. In addition, they summarized modeling approaches in computer-based ALSs and categorized them into four types: (a) stereotype modeling; (b) feature-based modeling; (c) the combination of stereotype modeling and feature-based modeling; and (d) constraint-based modeling.

Stereotype modeling clusters learners into different groups. Learners in the same group share common characteristics and receive the same instruction adapted to the group feature. Instead of modeling learners at a group level, *feature-based modeling* is a more fine-grained approach and focuses on modeling specific features of individual learners such as their prior knowledge, interests, and learning goals. The advantage of feature-based modeling is the ability to dynamically track changes in learner characteristics. *Constraint-based modeling* mainly focuses on erroneous knowledge (e.g., learners' errors or misconceptions). This method follows Ohlsson's learning theory of performance errors, in which he argues that learning is demonstrated through the correction of errors (Ohlsson, 1996).

Among these modeling categories, feature-based modeling is currently the dominant approach in web-based adaptive systems (Brusilovsky & Millán, 2007). The combination of feature-based modeling with stereotype modeling is a promising direction. It means that a learner is classified into a group firstly where after an individual feature-based model is initiated. This method allows for alleviating the effect of the cold start problem in adaptive learning environments, where no or very limited information about new learners is available when they enter an ALS (Brusilovsky & Millán, 2007; Pliakos et al., 2019). The four categories described by Vandewaetere et al. (2011) are coarse-grained classifications of modeling techniques in ALSs. Specific modeling approaches to individual characteristics are not touched upon in their study.

For example, Tseng et al. (2008) developed an ALS, in which student learning behavior and learning styles were adapted. Specifically, the system can provide subject materials (adaptation target: adaptive content) at different difficulty levels after monitoring student test scores and learning time on subject units. In addition, the system presented learning materials in the sequential form or using non-linear hypermedia based on student learning styles (adaptation target: adaptive presentation). By comparison, Yang et al. (2013) established an ALS by including student cognitive styles and learning styles as adaptation sources. Its layout (adaptation target: adaptive presentation) was based on student cognitive styles (i.e., field-dependent, field-independent) and its instructional strategy (adaptation target: adaptive content) was in harmony with student learning styles (e.g., active, reflective).

Comparisons of the Two Common Structures

Since Brusilovsky et al. (2004) proposed the layered evaluation structure, it has greatly facilitated studies of developing and evaluating new ALSs (e.g., Chrysafiadi & Virvou, 2013; Ounaies et al., 2012). Likewise, the tripartite structure of ALSs proposed by Vandewaetere et al. (2011) has acted as an important impetus in carrying out new literature reviews on adaptation source (e.g., Normadhi et al., 2019; Truong, 2016), adaptation pathway (e.g., Almohammadi et al., 2017; Mavroudi et al., 2018), and adaptation target (e.g., Premlatha & Geetha, 2015). Their broad quotations, which are supported by the number of times cited by other studies (i.e., 224 and 242), demonstrate the two structures are generally applicable to ALSs, if not all-inclusive.

Although outwardly different, the layered structure and the tripartite structure are connected internally. Learner characteristics, learner modeling, and adaptive effects in the layered structure correspond to the source, the pathway, and the target of adaptive instruction in the tripartite structure, respectively (Martin et al., 2020). However, the tripartite structure does not touch upon the adaptation decision-making process. Nonetheless, both structures emphasize the importance of learner modeling in ALSs (Martin et al., 2020). In fact, learner modeling could be regarded as a construct with three layers (Brusilovsky & Millán, 2007): (a) What is being modeled (the nature of learner characteristics)? (b) How is the information represented (structures of modeled learner characteristics)? And (c) how different kinds of models are maintained (learner modeling approaches)? In summary, learner modeling, involving learner characteristics (i.e., nature and structure) and modeling approaches, is a crucial component of ALSs. However, it remains a topic of much debate (Abyaa, 2019; Martin et al., 2018). The literature on learner modeling of ALSs is particularly reviewed in the following section.

Learner Modeling

Rarely do ALSs from different authors share an identical learner modeling component. The lack of coherence stems from various modeling technologies being used to establish learner models as well as different pedagogical theories being applied to the creation of the models. This section examined existing literature reviews on learner modeling, followed by three commonly involved learner characteristics in ALSs and their modeling approaches.

Existing Literature Reviews on Learner Modeling

To configure the learner model of an ALS, researchers and designers need to carefully consider what information and data about learners should be gathered and how it can be accurately modeled (Nguyen & Do, 2008). The more appropriate and precise the learner model is, the more advanced adaptive effects the ALS can provide (Abyaa, 2019; Vandewaetere et al., 2011). Given the significance of learner modeling in ALSs, five studies in the past decade reviewed learner modeling (i.e., Abyaa et al., 2019; Chrysafiadi & Virvou, 2013; Nakic et al., 2015; Normadhi et al., 2019; Vandewaetere et al., 2011).

As shown in Table 1, a wide variety of learner characteristics has been included in the development of ALSs. Nakic et al. (2015) enumerated 16 different variables that act as the source of adaptive instruction in ALSs. The other four studies categorized learner characteristics into groups (e.g., cognition, affect, and behavior) and each group contains several fine-grained features. However, none of the existing ALSs model all these features to provide adaptive effects in practice. Most ALSs model only a small number of learner characteristics. For example, Nakic et al. (2015) found that the most frequently used variable for adaptation is learning style. The second most commonly used variable is background knowledge, while cognitive styles and preferences are following. Martin et al. (2020) also found that learning style is the most used learner characteristic, followed by cognitive style and thinking style, and learner prior knowledge. In summary, current literature reviews reach a consensus that learning style, cognitive style, and prior knowledge are the most commonly modeled learner characteristics in ALSs. The remainder of this section elaborates on the nature, structure, and modeling approaches of learning style, cognitive style, and prior knowledge, separately.

Learning Style and Modeling Approach

Within the last three decades, the proposition that learner study in different ways has emerged as a prominent pedagogical issue (Hawk & Shah, 2007). Learning style is typically defined as the way people learn and prefer to learn (Honey & Mumford, 1992; Jonassen & Grabowski, 1993; Truong, 2016). Different learners have different favored methods to acquire knowledge. For example, some learners may understand quickly through graphical representations of learning materials; in contrast, some may prefer audio materials (Popescu,

Table 1

Literature Reviews on Learner Modeling of ALSs

Authors	Year	Title	Venue	Learner Characteristics
Vandewaetere et al. (2011)	No year limitation	The contribution of learner characteristics in the development of computer-based adaptive learning environments	*Computers in Human Behavior*	(1) Cognition (2) Affect (3) Behavior
Chrysafiadi & Virvou (2013)	2002 to 2012	Student modeling approaches: A literature review for the last decade	*Expert Systems with Applications*	(1) Knowledge (2) Errors/Misconceptions (3) Learning styles & Preferences (4) Other cognitive aspects (5) Affective features (6) Motivation (7) Meta-cognitive features
Nakic et al. (2015)	2001 to 2013	Anatomy of student models in adaptive learning systems: A systematic literature review of individual differences from 2001 to 2013	*Journal of Educational Computing Research*	(1) Age (2) Gender (3) Cognitive abilities (4) Meta-cognitive abilities (5) Psychomotor skills (6) Personality (7) Anxiety (8) Emotions and affect (9) Cognitive styles (10) Learning styles (11) Experience (12) Background knowledge (13) Motivation (14) Expectations (15) Preferences (16) Interaction styles
Normadhi et al. (2019)	2010 to 2017	Identification of personal traits in adaptive	*Computers & Education*	(1) Cognition (2) Affective (3) Behavior

		learning environment: Systematic literature review		
Abyaa et al. (2019)	2013 to 2017	Learner modelling: Systematic review of the literature from the last 5 years	*Educational Technology Research and Development*	(1) Student profile (2) Knowledge (3) Cognitive characteristics (4) Social characteristics (5) Personality traits (6) Motivation

2008). Keefe (1991) indicated that learning style is not only a characteristic that indicates how a learner studys and prefers to study but also an instructional strategy informing the cognition, context, and content of learning. Learning style has been included in a considerable number of ALSs. Most of these ALSs explore content-level adaptation attempting to match a learner with a specific learning style to content that should be the most appropriate for his/her learning style. Reviewing previous learning style-based ALSs revealed two main approaches to modeling learners' learning styles.

First, several inventories or questionnaires are applied to label learners with learning styles. One learning style model is the Felder-Silverman learning style (FSLS), which has been adopted by many ALSs (e.g., Hong & Kinshuk, 2004; Paredes & Rodriguez, 2004). The FSLS model classifies learners into four dimensions: active/reflective, sensing/intuitive, visual/verbal, and sequential/global (Felder & Silverman, 1988). Another learning style model is the visual, aural, read/write, and kinesthetic, which divides learners according to sensory preferences for learning (Fleming & Mills, 1992). As concluded by Truong (2016), the FSLS model is by far the most widely used learning style model in ALSs. For example, Hong and Kinshuk (2004) use the FSLS model to categorize learners into different learning styles, based on which their ALS presents course contents to individual learners. Specifically, if learners are recognized as with the sensing style, the system provides examples first, followed by exposition; in contrast, learners with the intuitive style receive exposition first, followed by examples.

Although questionnaires are easy to use, results from questionnaires may be biased as they depend on self-assessments of learners. In addition, results from questionnaires are not easy to update when learning styles change over time (Truong, 2016). For example, the questionnaire corresponding to the FSLS model can reach over 40-question long, thus updating learning styles would be time extensive. These drawbacks of questionnaires have encouraged a growing number of researchers to explore other alternative approaches. Recently, studies have been more concerned with automatically detecting learners' learning styles, with the aid of machine learning algorithms. The rationale is simple: learners' actions and interactions with ALSs' interfaces are analyzed to acquire learners' learning styles.

For example, Graf et al. (2010) introduced a framework that monitors learners' actions in online courses and uses this information to frequently update learners' learning styles. Using the FSLS model, action information (e.g., the number of times a learner visited or skipped a content page) serves as indications of a specific type of learning style (e.g., global style). In comparison with questionnaires, modeling learners' learning styles with machine learning algorithms is unobtrusive and likely to produce more user-friendly ALSs (Abyaa et al., 2019). However, one challenge of implementing machine learning algorithms to model learners' learning styles is to figure out connections between learning styles and observations of learners' actions and behaviors in ALSs. Most of the research done in this area requires the usage of questionnaires to validate the accuracy of machine learning classifiers (Rasheed & Wahid, 2021).

Although learning style is the most widely modeled learner characteristic in ALSs, there are no proven recipes for its usage. First, the extent to which such learning style-based ALSs improve learners' learning performance is unclear at present. In the background of classroom learning, a number of experimental studies (e.g., Papanagnou et al., 2016), which aim to evaluate the value of treating learners with different learning styles differently, concluded findings without any significant differences. Aslaksen and Lorås (2018) reviewed studies that examined the impact of tailoring instruction toward learning style preferences on learning outcomes. They did not find supportive statistical evidence for enhanced learning outcomes by aligning instruction to specific learning styles. However, some studies have reported that students' learning performance could be improved if proper learning style dimensions could be taken into consideration when developing ALSs (e.g., Hauptman & Cohen, 2011; Hwang et al., 2013). Given that learning style is quite a prevalent feature in ALSs, it is necessary to examine the benefits that learning style-based ALSs bring to learning outcomes.

Second, diverse structures of learning styles are available when researchers and designers establish ALSs. As claimed by Tseng et al. (2008), Keefe's learning style, which is used to identify sequential processing, discrimination, analytic, and spatial of learners, is the most suitable model for a web-based ALS. However, there is no supportive evidence for this claim. In practice, ALSs typically involve a specific learning style structure; they rarely incorporate more than one learning style structure and compare their effects on learning outcomes improvement. Because of this, it seems arbitrary to conclude that Keefe's learning style structure is the best option for ALSs. A meta-analysis allows us to compare the effects of different learning style structures and modeling approaches on learners' learning outcomes.

To sum up, questions such as "the extent to which such learning style-based ALSs improve learners' learning performance" and "whether system effectiveness of ALSs with different combinations of learning style structure and modeling approaches is sufficiently different" deserve further exploration.

Cognitive Style and Modeling Approach

Researchers typically define cognitive style as an individually preferred and habitual approach to organizing and representing information (Messick, 1984; Riding & Rayner, 1998). Learners' differences in cognitive style relate to different browsing strategies and learning preferences. For example, field-dependent learners generally perceive things in the entire perceptual field, whereas field-independent learners tend to see things individually (Witkin et al., 1977). In computer-based learning environments, cognitive style has been regarded as one of the key learner characteristics to system configuration. Likewise, some ALSs are typified by presentation- or navigation-level adaptation attempting to match a learner typified with a cognitive style to a specific interface condition (Mampadi et al., 2011; Uruchrutu et al., 2005).

Questionnaires relevant to different cognitive style models are used to measure learners' cognitive styles in ALSs. The most exploited constructs of cognitive styles are field dependence/field independence (Witkin et al., 1977), holist-serialist cognitive style (Pask, 1976), and verbalizer/imager cognitive style (Riding & Buckle, 1990). These cognitive styles have been successfully employed in the implementation of different instructional strategies in ALSs (e.g., Stash & De Bra, 2004; Triantafillou et al., 2004). For example, Triantafillou et al. (2004) developed an adaptive educational system based on cognitive styles (AES-CS). In the AES-CS, the Group Embedded Figures Test (GEFT) is the tool to determine learners' cognitive styles (i.e., field-dependent/independent). The system adapts system presentations to learners' cognitive styles. Specifically, field-dependent learners are provided with learning materials from general to specific. In contrast, learners with a field-independent cognitive style are served with learning contents from specific to general.

Research results about the impact of incorporating cognitive style as an adaptation source on the effectiveness of ALSs is a story of mixed success. On one side, a number of studies confirmed that learners with different cognitive styles search and browse information differently. On the other side, few successful stories on using cognitive styles to increase system effectiveness were reported. For example, Uruchrutu et al. (2005) found that most participants show their preference for the image and whole interface or the verbal and analytic interface. However, there is no significant improvement in participants' learning performance when they are in a matched interface condition. The possible reason is that learning strategies developed by learners (e.g., learning skills, strategies, and study orientations) prevail over their cognitive styles in improving learning. As a whole, the situation of modeling cognitive styles in ALSs is similar to the case of including learning styles in ALSs: Adaptation to cognitive styles may hold potential to improve the effectiveness of ALSs, but existing research offers almost no practical suggestions on what structures of cognitive style deserve modeling in ALSs.

Prior Knowledge and Modeling Approach

Learners' prior knowledge of the subject or the domain being taught is often used to adapt learning content in ALSs (Brusilovsky & Millan, 2007; Chrysafiadi & Virvou, 2013). The majority of ALSs focused on two types of domain knowledge: conceptual knowledge (i.e., facts and relationships) and procedural knowledge (i.e., problem-solving skills). Conceptual knowledge is typically represented in the form of a network of concepts. Procedural knowledge is most frequently represented as a set of problem-solving rules. Systems that focus on helping learners solve educational problems usually rely on procedural knowledge. In contrast, systems that assist learners in selecting educational content always rely on conceptual knowledge. The process of modeling learners' prior knowledge is also called knowledge modeling. There are multiple knowledge modeling approaches available for ALSs.

The simplest form of knowledge modeling is the scalar model approach, which estimates the level of learner domain knowledge by a single value on a quantitative scale (e.g., a number ranging from 0 to 5) or qualitative scale (e.g., good, average, poor). A number of ALSs use the scalar model to support an adaptive presentation. That is, systems divide learners into distinct groups according to their knowledge level and serve different versions of learning content to learners with different levels of knowledge.

Another form of knowledge modeling is the overlay model. The purpose of the overlay model is to represent an individual learner's knowledge as a subset of a domain model, which reflects expert-level knowledge of a subject. In order to use the overlay model, domain knowledge must be able to be broken down into generic items such as rules, concepts, and facts. Basically, the overlay model estimates the mastery status of each item in the domain model (i.e., yes or no). Thereby, the complexity of the overlay model depends on the granularity of the domain model structure. As concluded by Chrysafiadi and Virvou (2013), the overlay approach is the most preferred technique for representing learners' mastery of knowledge.

With advances in educational data mining algorithms, learners' mastery status represents the degree to which students know a domain fragment. That is, educational data mining algorithms are combined with the overlay model to infer learners' knowledge states. Unlike yes or no resulting from the overlay model, educational data mining algorithms provide probabilistic diagnoses of mastering state. For example, Chrysafiadi and Virvou (2013) developed the fuzzy knowledge state definer, which is able to dynamically identify and update a learner's knowledge level of all the concepts of the domain knowledge. The functionality is achieved by combining fuzzy theory with the overlay model.

Research Gaps and the Need for a Meta-Analysis

As a new generation of technology-enhanced learning systems, ALSs have attracted high interests of researchers over the previous three decades. There have been dozens of reviews that summarized critical components of ALSs (e.g., Vandewaetere et al., 2011). Additionally, some studies specifically reviewed learner modeling of ALSs (e.g., Abyaa et al., 2019; Martins et al., 2008). However, upon closer inspection, several questions concerning system effectiveness in promoting student learning performance are still unresolved.

First, the overall effectiveness of ALSs in improving student learning outcomes is unclear. Basically, ALSs are established to achieve personalized learning on a large scale. As an exemplary method of personalized learning, one-to-one tutoring was found to be with a two-sigma effect size compared with group instruction (Bloom, 1984). More recently, the effect size of intelligent tutoring systems in promoting learning outcomes has been broadly investigated (e.g., Kulik & Fletcher, 2016; Ma et al., 2014). For example, Ma et al. (2014) found that the use of intelligent tutoring systems was associated with greater student achievement (Hedges's g = 0.42) in comparison with teacher-led, large-group instruction. Although intelligent tutoring systems lay the groundwork for the development of ALSs, they are not equivalent.

In addition to intelligent tutoring systems which are typified by system adaptation, adaptive educational hypermedia systems also account for a large proportion of ALSs. However, there is a dearth of meta-analyses of empirical studies on adaptive educational hypermedia systems. The conclusion of a medium-to-large effect size of ALSs in Verdú et al.'s (2008) study requires further verification by meta-analyses. The effect size of ALSs concluded by Fontaine et al. (2019) is only suitable for health professionals and students. This dissertation intends to provide a more accurate effect size estimate by conducting a systematic meta-analysis with a comprehensive review of existing ALSs studies that focused on the evaluation of ALSs' effects on learner learning outcomes. Thereupon, the estimated effect size would provide an orientation on how far existing ALSs are away from other instructions (e.g., one-to-one tutoring, intelligent tutoring systems) in advancing learning achievements.

Second, although the diversity of existing ALSs is widely accepted, no research is available on the difference in their effectiveness in promoting learning achievement. For example, the importance of learner modeling to ALSs has been broadly recognized. Constructing a learner model requires choosing learner characteristics (e.g., learning style), structures of learner characteristics (e.g., the FSLS model), and learner modeling approaches (e.g., questionnaire). With regard to each layer, there are various options open to researchers. As a result, there is a wide variability of learner modeling for existing ALSs (Abyaa et al., 2019; Normadhi et al., 2019). In addition, both Verdú et al. (2008) and Fontaine et al. (2019) found significantly discrepant magnitudes of system effectiveness. Thus, it is necessary to investigate the heterogeneity of system effectiveness in improving student learning outcomes.

Third, there is no research on what critical factors could contribute to the effectiveness of ALSs. For example, to date, the selection of learner characteristics and modeling approaches often depends on researchers' or designers' personal preferences and experiences. Given that a multitude of learner characteristics and modeling approaches have been identified in the literature, it is therefore important to know which of them are directly associated with student learning process when they are used in ALSs. Otherwise, the design of a learner model will become unnecessarily complex or arduous. In this study, the most frequently used learner characteristics (i.e., learning style, cognitive style, and prior knowledge) are considered. Although previous research described implementations of these characteristics in ALSs, there is sparse data related to the empirical effectiveness of adding them to learner model in increasing student learning outcomes (Vandewaetere et al., 2011). Therefore, a meta-analysis is warranted to identify the effectiveness of the three learner characteristics and their corresponding modeling approaches, thus avoiding a larger proliferation of ad hoc constructed learner model in the field of ALSs.

To fill research gaps surrounding system effectiveness, it is important to emphasize the need for a meta-analysis of prior efforts in developing ALSs in order to enhance the effectiveness of adaptive learning environments.

Objectives of the Present Research

The objective of this meta-analysis is to gain quantitative insight into the effectiveness of ALSs in advancing student learning outcomes. For the meta-analysis to produce meaningful results, the present study needs to focus on a specific topic in order to include studies that are sufficiently comparable. With reference to the layered evaluation structure (Brusilovsky et al., 2004) and the tripartite structure (Vandewaetere et al., 2011), this study focuses on the common ground across different ALSs: learner modeling. The primary questions that this meta-analysis seeks to answer are as follows: Is the system effectiveness heterogeneous? If so, does the variability of learner modeling (i.e., learner characteristics and modeling approaches) significantly account for the heterogeneity of system effectiveness? Specifically, we are interested in identifying important contributors to system effectiveness, under consideration of promoting student learning outcomes.

The present research makes the following contributions. Theoretically, the overall effect size of ALSs not only provides a comprehensive picture of the existing efforts on the implementation of ALSs, but also acts as a landmark that instructs the continued development of ALSs. For example, ALSs and other types of instruction (e.g., one-to-one tutoring and intelligent tutoring systems) could be compared in terms of their effectiveness (i.e., effect size) in promoting student learning outcomes. Practically, identifications of critical factors such as appropriate learner characteristics provide a mapping between influential factors and system effectiveness for system designers and developers. Specifically, the results of this research are useful for the decision-making process when system designers and developers only model one or a few student characteristics in their ALSs.

Chapter 3 Methodology

Glass and his colleagues (1981) proposed four steps to conduct a meta-analysis: (a) finding studies, (b) coding study features, (c) measuring study effects, and (d) statistically analyzing and combining findings. This four-step procedure has been widely used in meta-analysis studies (e.g., Cheng et al., 2019; Garzón & Acevedo, 2019; Kulik & Fletcher, 2016) and was adopted in the present study.

Finding Studies

A thorough and systematic search was employed to find studies for this meta-analysis, which consists of three steps: (a) assembling a large pool of candidate studies through computer searches of electronic library databases; (b) developing inclusion criteria to capture evidence relevant to research questions; (c) examining the candidate studies individually to determine whether they are suitable for this meta-analysis.

Candidate Studies

Computer searches were carried out in six databases: Educational Resources Information Clearinghouse (ERIC), Web of Science, Google Scholar, ScienceDirect, Scopus, and ProQuest Dissertations & Theses Global. Different from the other five databases, the ProQuest Dissertations & Theses Global database includes unpublished dissertations, which are commonly examined in meta-analysis studies (e.g., Ma et al., 2014; Steenbergen-Hu & Cooper, 2013). The search string was: (intellige* OR adapt* OR customi*) AND (learning OR instruction OR education OR tutoring OR mentoring) AND (system OR environment OR software OR application OR program). The asterisk is a commonly used wildcard symbol that broadens a search by finding words that start with the same letters. The following key terms were searched in returned titles: adapt* e-learning system*, adapt* e-learning environment*, adapt* e-learning hypermedia system*, adapt* learning system*, adapt* scaffolding e-learning system*, online adapt* learning tool*, computer-based adapt* learning environment*, computer-based adapt* learning system*, adapt* and intellige* education system*, and adapt* intellige* tutor* system*. In addition, the word “adaptive” or its variants (i.e., “adaptation” or “adaptivity”) had to appear in abstracts. No restrictions were imposed regarding the publication year and publication type during the searches to prevent leaving out relevant studies. References retrieved were exported to Endnote version X9 (The Endnote Team, 2013), resulting in a pool of candidate studies.

Additional studies were searched by branching from reference sections of existing literature review studies on ALSs (e.g., Erümit & Çetin, 2020; Fontaine et al., 2019; Jando et al., 2017; Kumar et al., 2017; Martin et al., 2020; Normadhi et al., 2019; Özyurt & Özyurt, 2015; Verdú et al., 2008; Xie et al., 2019) and manually added to the candidate pool. Three literature reviews and one meta-analysis were especially helpful in developing a more comprehensive candidate pool: Normadhi et al. (2019) reviewed 78 adaptive learning environments that were established spanning the year of 2010 to 2017; Erümit and Çetin (2020) reviewed 32 studies on adaptive intelligent tutoring systems; Martin et al. (2020) systematically reviewed adaptive learning research (*n* = 61) ranging from 2009 to 2018; Fontaine et al. (2019) investigated 21 studies to calculate the effectiveness of ALSs in improving knowledge, skills, and clinical behavior of health professionals and students.

It is important to note that the term “adaptive learning system” rather than the term “intelligent tutoring system” was centered when databases were searched to assemble the candidate studies pool for two reasons. First, the usage of the keyword “adaptive learning systems” is consistent with how existing systematic reviews search for studies on ALSs. For example, Kabudi et al. (2021) performed a systematic mapping of the literature on AI-enabled

ALSs. They used two main terms to perform database searches: adaptive learning system and AI. In contrast, Xu et al. (2019) used the keyword “intelligent tutoring systems” to search databases, given that their research focused on intelligent tutoring systems. Second, there have been several meta-analyses of intelligent tutoring systems in the literature (e.g., Kulik & Fletcher, 2016; Xu et al., 2019), but few meta-analytic studies focused on ALSs in the literature. For example, Xu et al. (2019) conducted a meta-analytic study on the effectiveness of intelligent tutoring systems on K-12 students' reading comprehension. They included studies published in peer-reviewed journals from 2000 to 2017.

Inclusion Criteria

In order to select target studies from the candidate studies pool for the current meta-analysis, the following six inclusion criteria were developed: (a) the study is published in a journal article, scientific book, book chapter, conference proceeding, or dissertation; (b) the study is published in the English language; (c) the study compares the effectiveness of its ALS on learning outcomes of individual learners to those under non-adaptive instructions; (d) the study adopts a randomized experimental or quasi-experimental design with an independent comparison group; (e) the study provides sufficient quantitative information for the calculation or estimation of effect sizes; and (f) the study provided detailed information regarding the architecture of ALSs (i.e., learner model, domain model, and adaptation model).

Criterion (a) and criterion (b) are used to include high-quality and accessible publications. Meeting abstracts are excluded in this selection step because it is unclear whether these types of documents had been subjected to peer review, which is a generally accepted criterion for ensuring scientific quality.

Criterion (c) is used to select studies relevant to the purpose of this meta-analysis. In this selection step, studies that do not compare the effectiveness of ALSs to non-adaptive instructional methods are excluded. Studies that focus on providing general information about the interface design of an ALS or proposing a new framework for an ALS but do not provide quantitative evidence for their systems' effectiveness are excluded.

With regard to criterion (d), selected studies could be either field evaluations or laboratory investigations, but a randomized experimental or quasi-experimental design is required. Their control group(s) need to receive instructions without adaptations, which could be either a group that got a traditional teacher-led classroom instruction or a group that received computer-based non-adaptive instructions. If a quasi-experimental design is used, evidence has to be provided that the treatment and comparison groups were equivalent at baseline (What Works Clearinghouse, 2013). Studies with a significant pre-existing difference between the treatment and comparison groups are excluded unless the information is available to calculate effect sizes that would consider the prior difference. In this case, the method of estimating effect sizes from a pretest-posttest-control group design was used. That is, effect sizes are calculated based on the mean pre-post change in the treatment group minus the mean pre-post change in the control group, divided by the pooled pretest standard deviation (Morris, 2008).

Criterion (e) is used to select studies that reported sufficient data to calculate effect sizes. In this respect, common learning outcome measurements including course grades or scores on either locally developed tests or standardized tests are considered acceptable for this meta-analysis. However, studies that do not measure learners' learning outcomes in a quantitative way were excluded from this meta-analysis.

Criterion (f) is also relevant to the purpose of this meta-analysis. That is, studies that clarify the architecture of ALSs are chosen. Detailed information on the architecture of ALSs, especially for learner models, is necessary to answer the third research question of this study. The common article structure of candidate studies is shown in Figure 5. That is, the authors of these studies first introduce the architecture of ALSs, which includes learner models, domain models, and adaptation models. Information about learner characteristics and modeling approaches is usually provided in learner models, followed by their experimental design (e.g., quasi-experimental design) and experimental results (e.g., mean score of experimental/control group) presented in their studies.

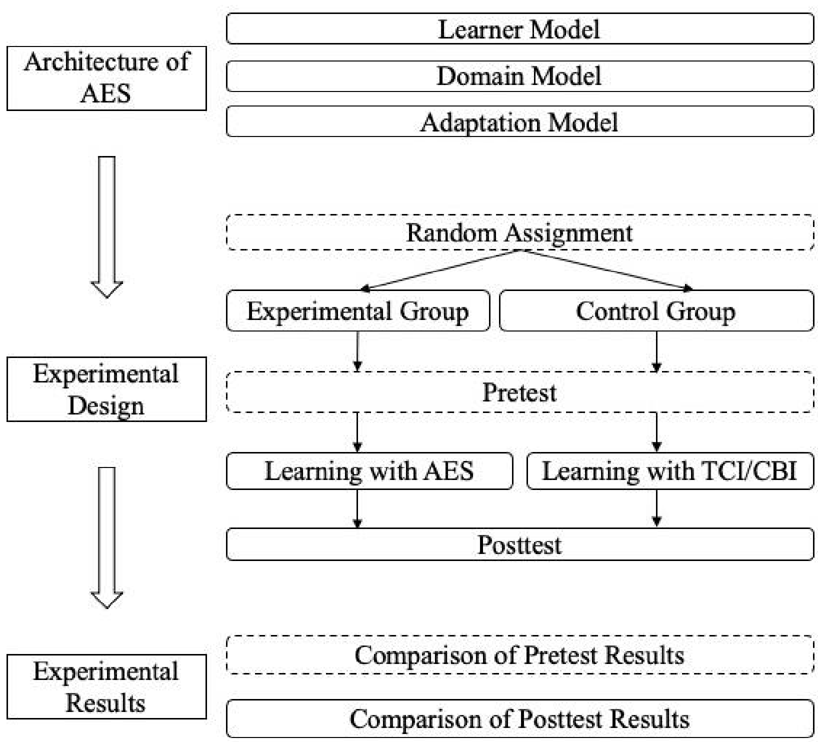

Figure 5. *Common Structure of Candidate Studies*

Note*:* TCI: Traditional Classroom Instruction; CBI: Computer-Based Instruction. Components within dotted lines could be absent for some candidate studies.

Final Data Set

According to the six inclusion criteria, selections from candidate studies were performed. The process of selecting studies followed the procedure of the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA; Moher et al., 2009). An adapted PRISMA flow diagram for the process of literature search and the number of studies identified, screened, ultimately eligible, and included in the meta-analysis can be viewed in Figure 6.

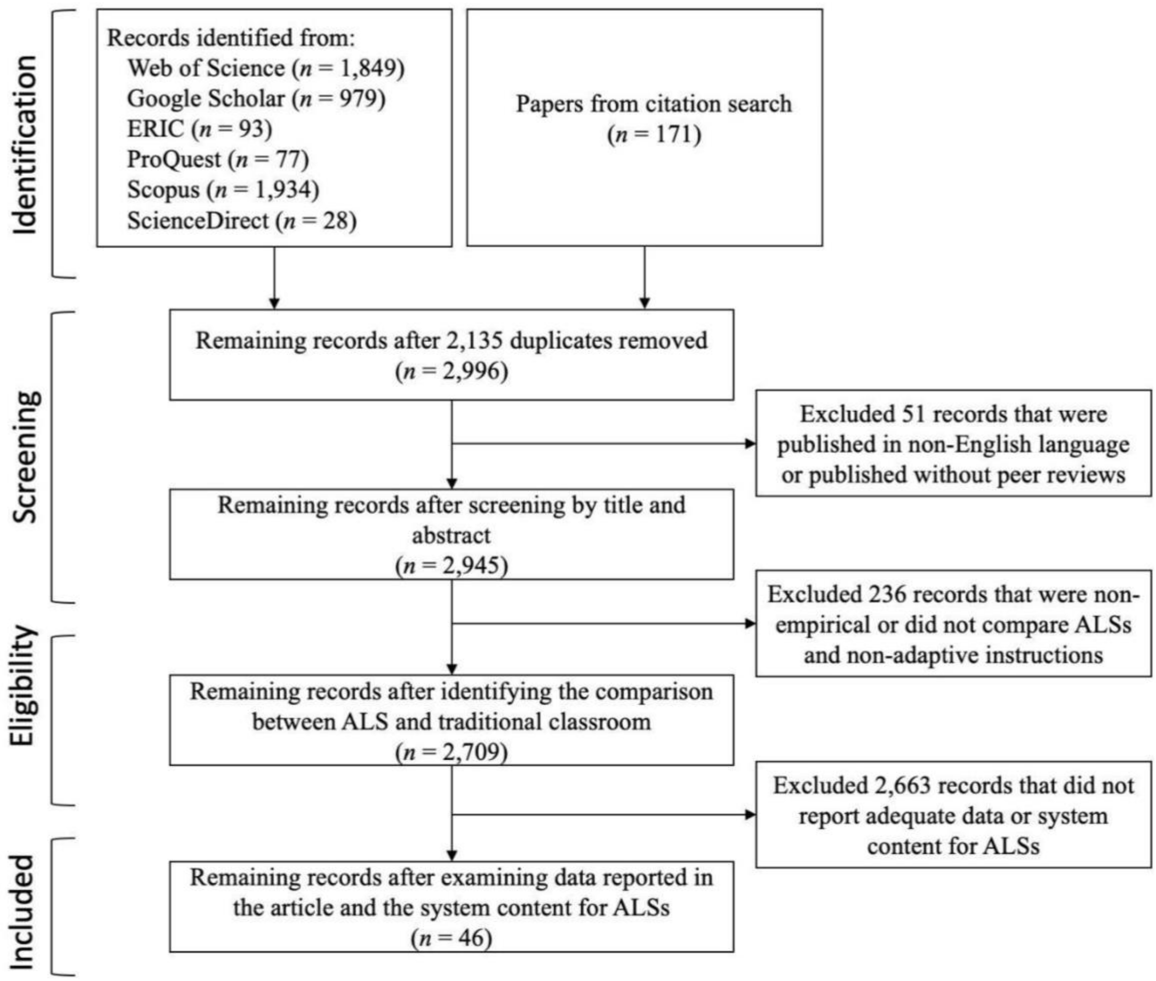

Figure 6. *A Step-By-Step PRISMA Flowchart for Selecting the ALS Articles*

As shown in Figure 6, the phase of identification found 5131 studies that were imported into Endnote. During the first step of the screening phase, all the duplicates were removed, and

2996 publications were further examined. In the second step of publication selection, studies that did not use the English language or were published as workshop introductions were identified and excluded. Included publications could be peer-reviewed journal articles, conference proceedings, and dissertations/theses. During the third step of the screening phase, the titles and the abstracts of 2945 studies were read to make sure these publications were relevant to ALSs. After the execution of phase two, the number of publications was reduced to n = 2709 for consideration in the study.

During the eligibility phase, the methodological designs and the data generated by the studies were examined. Studies that did not provide adequate data for calculating effect sizes were excluded. Studies that did not depict architectures of ALSs were also excluded. Studies could have reported multiple student outcome measures, such as a midterm score or a final exam score. These data were used to calculate effect sizes using the appropriate equations for the outcomes. The statistical analysis section provides more details about the specific statistical models employed. To summarize, this study consists of 46 studies and a total of 77 effect size estimates; individual studies contributed between 1 and 12 effect sizes. A complete list of these studies can be found in Appendix B.

Coding Procedures

A detailed coding protocol to guide information retrieval and coding was designed (see Table 2). The protocol covered the major characteristics of the studies, which included (a) study design features (e.g., sample sizes; whether the study used a randomized or quasi-experimental design; whether the study compared an ALS with regular classroom instruction or computer-based non-adaptive instruction), (b) contexts of instruction (e.g., subject domain; education level; the duration of ALS instruction), (c) architectures of ALSs, including learner characteristics (i.e., learning style, cognitive style, and prior knowledge) and modeling approach (i.e., questionnaire, machine learning algorithms, scalar modeling, and overlay modeling), and (d) study outcomes (e.g., the magnitude and the direction of effect sizes).

Table 2

Study Features and Associated Coding Categories in the Meta-Analysis

Features and Coding Categories
Category 1: Study Design Features
Study Design
Randomized-experimental design (RED)
Quasi-experimental design (QED)
Control Group
Regular classroom instruction (RCI)
Computer-based instruction (CBI)
Sample Sizes
Sample size for treatment/control group
Category 2: Contexts of Instruction
Subject Domains
Mathematics
Physics
Computer science
Language and literacy
Biology and physiology
Humanities and social science
Not reported
Education Levels
Elementary school
Middle school
High school
Postsecondary
Mixed grades
Not reported
Duration of Treatment: A number that describes the instruction duration in weeks

Category 3: Adaptation Content
  Student Characteristics
    Learning style
    Cognitive style
    Prior knowledge
    Mixed characteristics
  Modeling Approaches
    Questionnaire
    Machine learning algorithms
    Scalar
    Overlay
    Mixed approaches

About 35% of the studies (i.e., randomly selected 16 from the 46 studies) were coded independently by the author and another graduate student, both with more than four years of research experience in education. Both coders understood what and how features should be coded, and their codes were checked and compared. The inter-coder reliability as percentage agreement reached 94.5%, with a Cohen's Kappa of 0.89, indicating that the level of agreement was strong (McHugh, 2012; Stemler & Tsai, 2008). Inconsistent codes were discussed and resolved. The rest of the studies (i.e., 30 studies) was reviewed and coded by the author.

Calculating Size of Effects

An effect size is a metric introduced by Glass (1977) representing the difference between the means of an experimental group and a control group expressed in standardized units (i.e., divided by a standard deviation). Three most common methods for calculating an effect size are (a) Glass's *delta* (Glass, 1976), which uses the standard deviation of the control group, (b) Cohen's *d* (Cohen, 1962, 1988), which makes use of the pooled standard deviation of the control and experimental groups, and (c) Hedges's *g* (Hedges, 1982), which applies a correlation to correct the problem of the overestimation of the effect size based on small samples.

Researchers usually choose one of the three methods to calculate effect sizes. For example, Tamim et al. (2011) conducted a second-order meta-analysis to summarize 40 years of research on the impact of instructional technology on student learning. In their study, four of the 25 meta-analyses, covering a total of 128 studies, used Cohen's *d* exclusively to report size of effects; six meta-analyses, covering 505 studies, used Glass's *delta* exclusively to calculate effect sizes; and ten meta-analyses, covering 239 studies, used Hedges's *g* exclusively. Glass's *delta* and Hedges's *g* are the two estimators of size of effects that are used more often than Cohen's *d*. A review of the literature on ALSs was conducted and six relevant meta-analyses were identified (Kulik & Fletcher, 2016; Ma et al., 2014; Nesbit et al., 2014; Steenbergen-Hu & Cooper, 2013; 2014; Xu et al., 2019). All of the six studies used Hedges's *g* to calculate effect sizes. The preference for Hedges's *g* over other standardized-difference indices such as Cohen's *d* and Glass's *delta*, is due to the fact that Hedges's *g* can be corrected to reduce the bias that may arise when the sample size is small (i.e., $n < 40$; Glass et al., 1981). It should be noted that study samples in the field of ALSs are usually small.

To avoid potential influences of different effect sizes over analysis results, Cohen's *d*, Hedges's *g*, and Glass's *delta* were reported in this study. However, we gave primary emphasis to Hedges's *g* and treat Cohen's *d* and Glass's *delta* as important supplementary measures. Our preference for Hedges's *g* is based primarily on our preference to make a correction of effect sizes from studies with small samples. Their calculations were shown below.

The standardized mean differences effect size statistic (i.e., Cohen's *d*) is calculated as

$$d = \frac{\bar{X}_1 - \bar{X}_2}{S_{pooled}}, \tag{1}$$

where $\bar{X}_1$ and $\bar{X}_2$ are means of treatment and control groups, and the S_{pooled} is pooled standard deviation, defined as

$$S_{pooled} = \sqrt{\frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 2}}, \tag{2}$$

where n_1 and n_2 are the numbers of subjects in treatment and control groups, and S_1 and S_2 are the standard deviations for treatment and control groups. The variance of Cohen's d is derived by

$$V_d = \frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2(n_1 + n_2)}, \tag{3}$$

and the standard error of Cohen's d is the square root of its variance (as shown in Equation 4)

$$SE_d = \sqrt{\frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2(n_1 + n_2)}}. \tag{4}$$

An unbiased alternative estimator (i.e., Hedges's *g*) developed by Hedges (1981) was used to correct for bias from small sample sizes as follows:

$$g = \left[1 - \frac{3}{4(n_1 + n_2 - 9)}\right] d, \tag{5}$$

where n_1 and n_2 are the numbers of subjects in treatment and control groups, and d is the biased standardized mean difference shown in Equation 1. The variance and standard error of this estimator is calculated as

$$V_g = \left[1 - \frac{3}{4(n_1 + n_2 - 9)}\right]^2 \times V_d, \tag{6}$$

and

$$SE_g = \sqrt{\left[1 - \frac{3}{4(n_1 + n_2 - 9)}\right]^2 \times V_d}, \tag{7}$$

where V_d is calculated by Equation 4.

Glass's *delta* is defined as the mean difference between the experimental and control group divided by the standard deviation of the control group, which is calculated using Equation 10

$$D = \frac{\bar{X}_1 - \bar{X}_2}{S_2}. \tag{8}$$

In the absence of reported means or standard deviations, other procedures prescribed by Lipsey and Wilson (2001) using test statistics (*F* or *t*) were used to estimate unbiased size of effects. The *F* value can be used to calculate effect size as

$$d = 2\sqrt{\frac{F}{N}}, \tag{9}$$

where *N* denotes the total number of participants. In addition, the effect size can be calculated by *t* value as follows:

$$d = t\sqrt{\frac{(n_1 + n_2)^2}{(n_1 n_2)(n_1 + n_2 - 2)}}. \tag{10}$$

The calculations of effect sizes shown above are suitable for studies without significant pre-existing differences between the treatment and control groups. For studies that have significant difference in pretests of different groups, the method of estimating effect sizes from a pretest-posttest-control group design is used (Morris, 2008). That is, pretest results of the treatment and control group are also used to provide a more precise estimate of the treatment effect,

$$d = c_P\left[\frac{(\bar{X}_1 - \bar{X}_{pre,1}) - (\bar{X}_2 - \bar{X}_{pre,2})}{S_{pre}}\right], \tag{11}$$

where the pooled standard deviation is defines as

$$S_{pre} = \sqrt{\frac{(n_1 - 1)S_{pre,1}^2 + (n_2 - 1)S_{pre,2}^2}{n_1 + n_2 - 2}} \tag{12}$$

and

$$c_P = 1 - \frac{3}{4(n_1 + n_2 - 2) - 1}. \tag{13}$$

The $\bar{X}_{pre,1}$ and $S_{pre,1}$ represent the group mean and the standard deviation of the treatment group at pretest. Likewise, the $\bar{X}_{pre,2}$ and $S_{pre,2}$ describe the group mean and the standard deviation of the control group at pretest. After getting effect sizes of Cohen's *d*, Equations 5 can be used to calculate Hedges's *g*.

Statistical Analysis

This study extracts and integrates empirical research data for performing a meta-analysis to obtain the overall effect size between different quantitative works and to contrast the average effect of research groups with different characteristics. The software R (R Core Team, 2022) was adopted to conduct the meta-analysis, including the estimation of the potential publication bias, the overall effect size, and the moderator effects among the 46 collected articles.

Publication Bias

Publication bias is defined as "an editorial predilection for publishing particular findings, which leads to the failure of authors to submit negative findings for publication" (Porta, 2014). Publication bias is a widespread problem in systematic reviews and meta-analyses, which can affect the validity and generalization of conclusions. The current study assessed publication bias through funnel plots (Bartolucci & Hillegass, 2010), trim and fill method (Egger et al., 1997), and Fail-Safe *N* approach (Orwin, 1983).

A visual tool—funnel plot—is often used to examine publication bias (Bartolucci & Hillegass, 2010). It is a simple scatterplot of the treatment effects estimated from individual studies (horizontal axis) against a measure of study size (vertical axis). In the absence of bias, results from small studies will scatter widely at the bottom of the graph, with the spread narrowing among larger studies (Sterne & Harbord, 2004). However, publication bias may lead to an asymmetrical funnel plot. In this case, a nonparametric "trim and fill" method can be employed (Duval & Tweedie, 2000a, 2000b) to adjust potential publication biases.

Rosenthal's (1979) Fail-Safe *N* is an estimate of the number of unpublished studies that would be needed to reverse a conclusion that an effect does indeed exist. According to Rothstein (2008), a Fail-Safe *N* value, which equals to or is greater than five times the number of studies plus ten studies, would indicate that the meta-analytic results are robust to the threat of publication bias. The "*meta*" package together with its "*funnel.meta*" and "*trimfill*" functions was used to print a funnel plot and adjust publication bias. The "*fsn*" function from the "*metafor*" package was used to conduct a Fail-Safe *N* analysis.

Models for Meta-Analysis

Meta-analysis is a method that enables the combination and summary of quantitative information from different studies. Using data from different studies creates the opportunity to provide a structured summary of a specific research topic and to find relationships between variables that otherwise would not be detected. To synthesize effect size estimates from selected studies and to investigate potential moderators for the heterogeneity across different studies, correlated and hierarchical effects (CHE) models (Pustejovsky & Tipton, 2022) were used.

Correlated and Hierarchical Effects Model. Statistical independence is one of the core assumptions when effect sizes are analyzed in a meta-analysis. If there is a dependency between effect sizes, this can artificially reduce heterogeneity and thus lead to false-positive results. Hierarchical effects models and correlated effects models are commonly used models to handle effect sizes' dependency that is regularly encountered in meta-analytic research (Fisher & Tipton, 2015). Hierarchical effects models assume that dependence arises solely through common features of a research group, while within a group, each effect size is estimated on an independent sample. In contrast, correlated effects models assume that dependence emerges because effect sizes are estimated based on the same sample (e.g., multiple measures of a common outcome construct or one outcome assessed over multiple time points). Like hierarchical effects models and correlated effects models, the CHE model provides an approach to include dependent effect sizes in meta-regression analyses, even when the nature of the dependence structure is unknown. Moreover, by combining features of hierarchical effects models and correlated effects models, the CHE model allows for both between-study heterogeneity and within-study heterogeneity in true effect sizes (Pustejovsky & Tipton, 2022).

In this study, effect sizes are not independent; instead, some are nested within the same studies. Furthermore, effect sizes within the same studies are dependent such as their calculations depending on the same participants. Thus, the CHE model, which not only considers nested structure but also accounts for relations of effect sizes from the same studies, is appropriate for the present study. Given the CHE model can be regarded as an extended multi-level model, a multi-level model (i.e., three-level model) is shown below, followed by the CHE model. In terms of a three-level meta-analysis model, effect size dependencies are accounted for by three levels.

First, in each of the selected publications, the author(s) "pooled" the results of individual participants in study j and report the aggregated effect size i. The first level model is denoted as

$$\hat{\theta}_{ij} = \theta_{ij} + \epsilon_{ij}, \tag{14}$$

where $\hat{\theta}_{ij}$ is an estimate of the true effect size θ_{ij}, the term ij can be understood as "effect size i nested in study j", and ϵ_{ij} is the random sampling error.

Then, on level 2, these effect sizes are nested within different clusters. These clusters can either be individual studies or subgroups of studies. The second level model can be represented as

$$\theta_{ij} = \kappa_j + \zeta_{(2)ij}, \tag{15}$$

where parameter κ_j is the average effect size of study j.

Lastly, on level 3, pooling the aggregated cluster effects leads to the overall true effect size. The third level model can be represented as

$$\kappa_j = \mu + \zeta_{(3)j}, \tag{16}$$

where parameter μ is the overall average population effect. The overall formula is

$$\hat{\theta}_{ij} = \mu + \zeta_{(2)ij} + \zeta_{(3)j} + \epsilon_{ij}. \tag{17}$$

There are two heterogeneity terms in the overall formula: (a) $\zeta_{(2)ij}$ stands for the within-cluster heterogeneity on level 2, and (b) $\zeta_{(3)j}$ represents the between-cluster heterogeneity on level 3. It is worth noting that the three-level model assumes effect sizes within one study are independent (i.e., $\mathrm{Cov}(\epsilon_{hj}, \epsilon_{ij}) = 0$). Thus, the only source of dependence between effect sizes in the same study is about the true effect sizes, not estimation error.

The structure of the CHE model is the same as the multi-level model, but the CHE model assumes that effect size estimates within studies are dependent. That is, when several effect sizes are from one study, their sampling errors are expected to be correlated. A single, known correlation ρ between pairs of effect sizes from the same study, which is the same across all studies, is assumed and $\mathrm{Cov}(\epsilon_{hj}, \epsilon_{ij}) = \rho s_j^2$. When the CHE model is performed, the Robust Variance Estimation (RVE; Hedges et al., 2010; Tanner-Smith et al., 2016) is firstly used to handle dependent effect sizes. The RVE approach is used to output a sandwich estimator (White, 1982), which is also known as the empirical variance-covariance matrix estimator and a useful tool for variance estimation (Ma et al., 2014). The resulting sandwich estimator can be used in combination with the CHE model to obtain robust confidence intervals and *p*-values.

Given there is a broad variety of dependent effect sizes in the present study, the CHE model was selected to synthesize effect sizes from different studies and investigate potential moderator effects for the heterogeneity of effect sizes across studies. Specifically, a three-level CHE model (i.e., participant, effect size, and publication) was compared with a two-level CHE model (i.e., participant and effect size). The Akaike Information Criterion (AIC; Akaike, 1974), the Bayesian Information Criterion (BIC; Schwarz, 1978), and a likelihood-ratio test were used to compare the two models. The more robust model was used to examine moderator effects.

Moderator Effect Analyses with the Correlated and Hierarchical Effects Model. As shown above, the CHE model allows to examine differences in outcomes within studies (i.e., within-study heterogeneity) as well as differences between studies (i.e., between-study heterogeneity). If there is evidence for heterogeneity in effect sizes, moderator analyses can be conducted to test variables that may explain within-study or between-study heterogeneity. For these analyses, the CHE model can be extended with different characteristics. In this study, factors (i.e., learner characteristic, modeling approach, subject area, publication source, and publication year) were added to equations to examine their moderation effects. For example, predictors can be entered into the three-level CHE model

$$\hat{\theta}_{ij} = \theta + \beta x_i + \zeta_{(2)ij} + \zeta_{(3)j} + \epsilon_{ij}, \quad (18)$$

where θ is the intercept and β is the regression weight of a predictor variable x. When x is a dummy variable, the model performs a subgroup analysis. When x is continuous, the formula represents a meta-regression model. A $p < 0.05$ significance level was adopted for the CHE model analyses. Given the complexity of the CHE model as well as the large number of the moderators, the moderators were entered to the CHE model one by one. As suggested by Pustejovsky and Tipton (2022), the "*rma.mv*" function of the "*metafor*" package was used to construct a series of CHE models.

Chapter 4 Results

This study intends to figure out how successfully the existing adaptive learning environments advance student learning performance in comparison with non-adaptive instruction. This purpose was achieved by a meta-analysis of quantitative evidence from 46 studies on ALSs. The meta-analysis also identified factors contributing to the varying success of adaptive learning environments in promoting student learning outcomes (i.e., different magnitudes of system effectiveness). This chapter arranges the findings in four sections: The mapping of the main characteristics of the 46 studies and the examination of publication bias of the effect sizes are presented in the first two sections, followed by the analyses of the overall system effectiveness and moderator effects.

Mapping of Study Characteristics

Following the procedure of the PRISMA (Moher et al., 2009), this study found 46 studies that were eligible for this meta-analysis. The mapping of the main characteristics of the 46 studies is shown in Table 3. A total of 77 effect sizes were reported in the included studies, which involved 3699 participants. In terms of publication channels, 69.57% (n = 32) of the included papers were published in scientific journals, and 21.74% (n = 10) were conference papers published in conference proceedings. The articles were categorized based on the educational levels of their included participants. Most of the articles (n = 34; 73.91%) established ALSs that targeted students at the postsecondary education level, and 21.74% (n = 10) adopted participants at the elementary and secondary education levels. Correspondingly, there were different subject areas where ALSs aimed to improve student learning: The subject area of computer science represents the highest proportion of the included articles (n = 23; 50.00%), followed by English language learning (n = 8; 17.39%). Some articles developed ALSs for

Table 3

Overview of the Background Characteristics of the Reviewed Studies

Overview of the Studies Included in this Meta-Analysis
No. of Studies: 46
No. of Effect Sizes: 77
No. of Participants: 3699
Publication Channel Journal: 32 (69.57%) Conference Proceeding: 10 (21.74%) Dissertation: 3 (6.52%) Book: 1 (2.17%)
Educational Level Elementary: 4 (8.70%) Secondary: 6 (13.04%) Postsecondary: 34 (73.91%) Unclear: 2 (4.35%)
Subject Area Computer Science: 23 (50.00%) Mathematics: 5 (10.87%) Technology/Science: 4 (8.70%) English Language: 8 (17.39%) Others (e.g., botanical): 6 (13.04%)
Treatment Duration: From one week to 24 weeks, *Mean* = 9 weeks, *SD* = 6.95 for studies providing numbers of weeks
Control Group Condition Traditional Classroom Instruction: 8 (17.39%) Computer-Based Instruction: 36 (78.26%) Unclear: 2 (4.35%)
Publication Year 2000–2010: 6 (13.04%) 2011–2015: 19 (41.30%) 2016–2022: 21 (45.65%)
Learner Characteristic Learning Style: 14 (30.43%) Prior Knowledge: 23 (50.00%) Cognitive Style: 2 (4.35%)

Note: Each study can contribute multiple ESs; thus, the total number of ESs exceeds the total number of studies. The learner characteristics that were modeled in the 46 ALSs were not limited to learning style, prior knowledge, and cognitive style.

subject areas of technology/science (*n* = 4; 8.70%) and mathematics (*n* = 5; 10.87%). The rest of the articles were categorized into one group (*n* = 6; 13.04%), which covered subject areas such as botanical learning and grammar of the Chinese language.

The duration that researchers conducted randomized experiments or quasi-experiments ranged from one week to 24 weeks. More specifically, the average duration was approximately nine weeks (*SD* = 6.95). However, it was difficult to get accurate measures of time on tasks because most of the articles did not report the accurate duration that participants used ALSs. It seems arbitrary to regard the duration of experiments to be equivalent to the duration of system usage. Unlike other meta-analyses that investigated the effectiveness of technology-enhanced learning systems on student learning (e.g., Xu et al., 2019), this study did not investigate the moderator effect of interaction duration on the magnitude of system effectiveness in the moderator effects section. In terms of control conditions, most of the included articles (*n* = 36; 78.26%) compared students' learning performance under ALSs with those under non-adaptive computer-based instruction. Participants of control groups in eight studies (17.39%) received traditional classroom instruction (i.e., large-group instruction). Although there was no restriction on publication years when studies were searched and screened, most of the included studies were published between 2010 and 2022 (*n* = 40; 86.96%).

In terms of learner characteristics, 50% (*n* = 23) of the included articles modeled learners' prior knowledge in their established ALSs. There were 30.43% (*n* = 14) of the proposed ALSs that provided adaptive instruction based on learners' learning styles. However, only two articles (4.35%) modeled learners' cognitive styles in their developed ALSs. Learner characteristics that the 46 studies modeled in their ALSs were not limited to prior knowledge, learning style, and cognitive style. Other learner characteristics such as motivation and goals were also modeled, which served as the source of adaptive instruction. It is worth noting that 58.70% (n = 27) of the included studies modeled one learner characteristic (e.g., modeling prior knowledge), but 26.09% (n = 12) modeled more than one learner characteristic (e.g., modeling both prior knowledge and learning style).

There were 23 studies that limited learner characteristics to knowledge, learning style, and cognitive style. As shown in Figure 7, 17 studies modeled prior knowledge, and eight studies modeled learning style, among which three studies modeled both of them. In terms of modeling approaches of prior knowledge, the most commonly used technique among the 17 studies was overlay modeling (n = 8; 47.06%). The questionnaire of the Index of Learning Styles (Felder & Spurlin, 2005) was the most frequently used tool to model learners' learning styles (n = 6; 75%). Among the 23 articles, there was only one study that used cognitive style as the source of adaptation instruction via the Study Preference Questionnaire (Ford, 1985).

Risk of Bias within Studies

This study employed a funnel plot to test for potential publication bias of effect sizes reported by the selected studies (Bartolucci & Hillegass, 2010). In a funnel plot, publications with larger sample sizes appear at the top of the funnel, while those with smaller sample sizes appear at the bottom. As shown in Figure 8(a), most of the reported Hedges's *g* appeared at the top, indicating large sample sizes. However, the funnel plot was asymmetric (t = 7.498, p < .01), which indicates potential missing values and publication biases. To adjust the potential publication biases, the method of trim and fill was used. Results indicated that 18 potential Hedges's *g* were missed (shown as white dots in Figure 8[b]), which is significantly lower than the number of Hedges's *g* (i.e., 77) in this study. Also, the difference between the adjusted

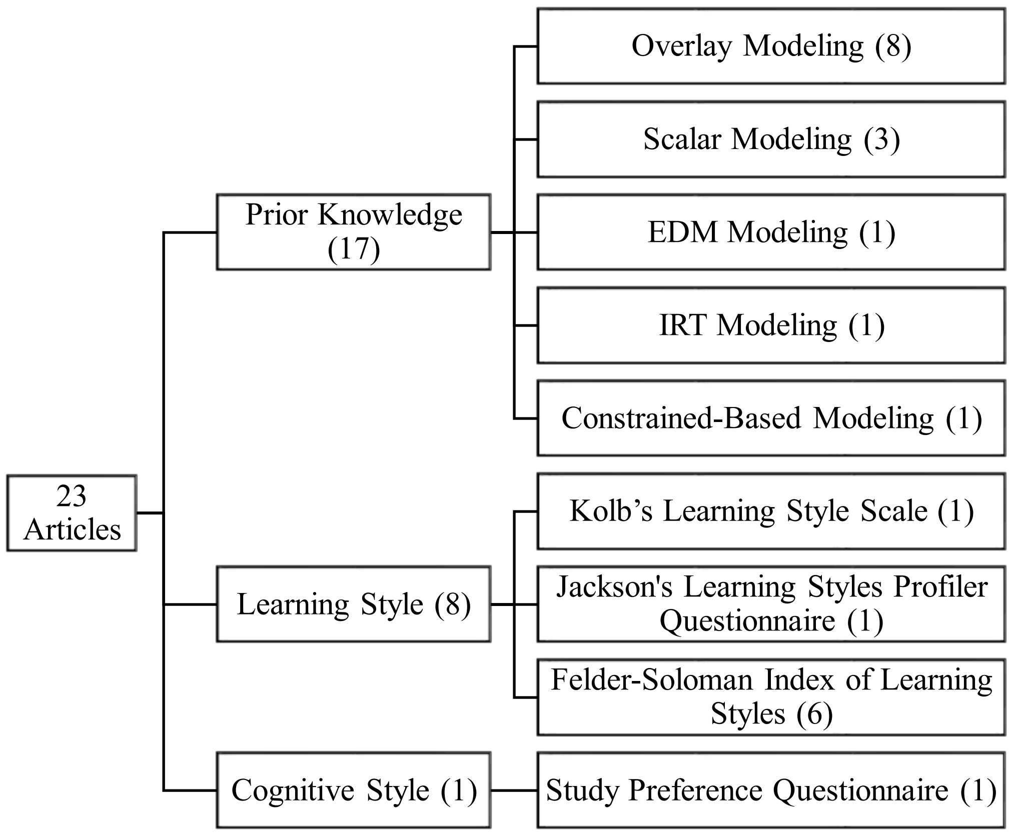

Figure 7. *Distribution of Studies across Learner Characteristics and Modeling Approaches*

estimate suggested by the trim and fill method and the unadjusted estimate was calculated. The difference is less than 20%, suggesting the bias can be classified as absent or negligible (Kepes et al., 2012; Vevea et al., 2019).

To account for the possibility that the current meta-analysis overlooked non-significant results, the Fail-Safe *N* analysis was performed (Rosenthal, 1979). Fail-Safe *N* suggested the number of Hedges's *g* that is needed to obtain no effect to completely nullify the observed mean effect size. Using R, a Fail-Safe *N* value of 31037 was calculated. According to Rothstein (2008), a Fail-Safe *N* value which equals to or is greater than five times the number of studies (i.e., Number of Hedges's *g*) in the original meta-analysis plus ten studies (i.e., [5×77 + 10] would indicate that the meta-analytic results were robust to the threat of publication bias. According to this guideline, the cut-off value for missing Hedges's *g* in this study was 395.

Given that the Fail-Safe N value was significantly larger than the cut-off value, there appears to be limited publication bias. Thus, the consensus among the inspection of the funnel plot of studies and results of the trim and fill method as well as the Fail-Safe N test suggest that the publication bias is unlikely to be a problem in the meta-analysis for examining the effectiveness of ALSs.

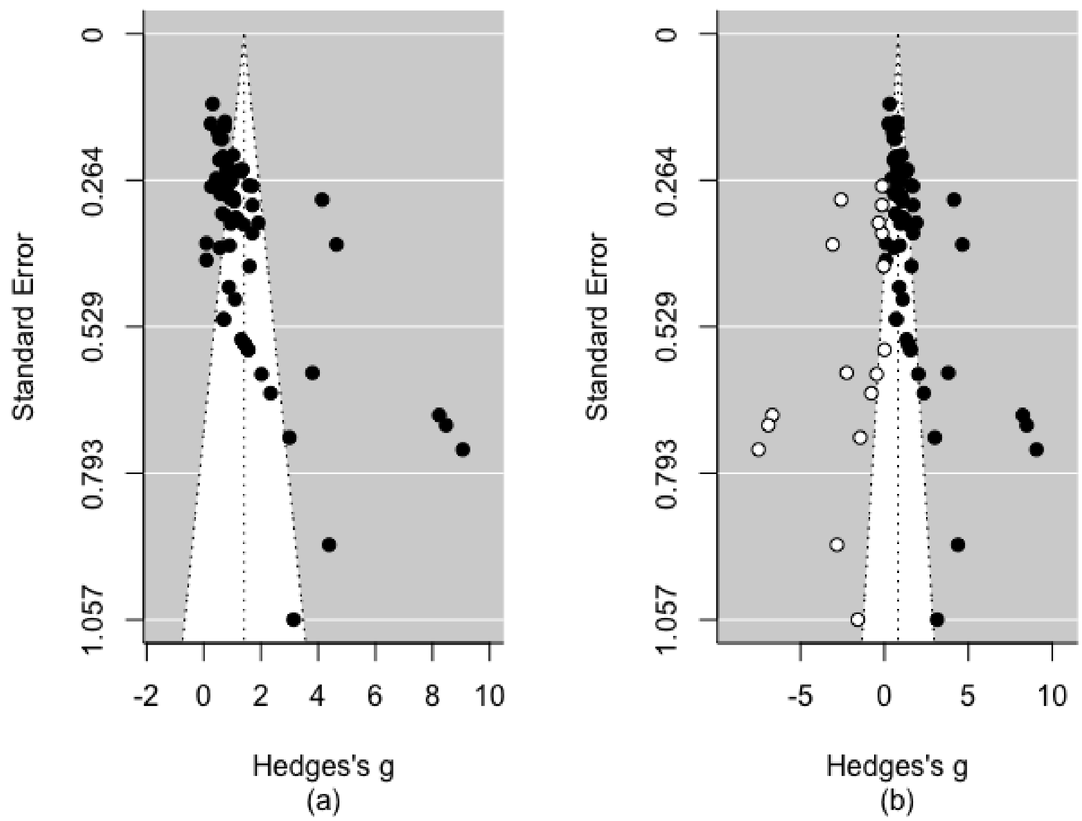

Figure 8. *Funnel Plot and the Adjusted Funnel Plot using Trim and Fill Approach*

Overall Effectiveness and Heterogeneity

Models Selection

Forty-six publications met the inclusion criteria, and some of them reported more than one Hedges's *g*. That is, multiple effect sizes might be nested within one publication in this study. Thus, a two-level (i.e., participant and effect size) CHE model was compared to a three-level (i.e., participant, effect size, and publication) CHE model. As shown in Table 4, the likelihood-ratio test indicated that the difference between the two models was significant and the three-level CHE model had a better model fit, with lower AIC and BIC values ($p < .001$). That is, the three-level CHE model was found to be more robust than the two-level CHE model for the present study. Thus, the three-level CHE model was used in the following data analyses.

Table 4

Model Comparison between the Two-Level and Three-Level CHE Models

	Two-Level CHE Model	**Three-Level CHE Model**
% of Total Variance		
I_0^2 (%)	5.655	4.462
I_T^2 (%)	94.345	83.471
I_A^2 (%)	–	12.061
$Total\ I_A^2$ (%)	94.345	95.532
Heterogeneity		
$Q_M(df)$	621.329 (76)	621.329 (76)
p	< .001	< .001
Model Fit of		
AIC	277.771	235.142
BIC	282.433	242.134
Between Two Models	$p < .001$	

The Effectiveness of ALSs

According to model comparison results, the three-level CHE model was used to analyze the effectiveness of ALSs reported in the 46 publications. Results revealed that the overall estimated Hedges's g was 1.227 (95% CI = 0.832, 1.622; $p < .001$). Thus, the overall summary estimate of the 77 effect sizes demonstrated a substantial level of effects that ALSs had on student learning performance.

In addition to the overall effectiveness of ALSs on student learning outcomes, the individual effectiveness was calculated in terms of different control group conditions: regular classroom instruction and non-adaptive computer-based instruction. Results showed that the estimated Hedges's g was 1.025 (95% CI = 0.065, 1.986; $p < .001$) when compared to large-group classroom instruction, and 1.206 (95% CI = 0.786, 1.627; $p < .001$) when compared to non-adaptive computer-based instruction. Furthermore, there was no significant difference between learning from large-group classroom instruction and non-adaptive computer-based instruction (Q_M = 0.114, *df* = 1, p = .735).

Although the average estimates of effect sizes were large, an examination of the heterogeneity in effect sizes contributed to a better understanding of ALSs' effects on learning outcomes. The forest plot (see Figure 9) displays the heterogeneity of Hedges's g estimates across the 77 effect sizes. Model results indicated that the variance of Hedges's g was high (Q_M = 621.329, *df* = 76, $p < .001$). The I^2 (95.54%) suggested that the total variability in Hedges's g estimates could be attributed to the true within- and between-study heterogeneity and not a sampling error (τ^2 = 4.46%). Furthermore, the within-publication (i.e., effect size, level 2) and between-publication (level 3) variability accounted for 12.06% and 83.47% of the total

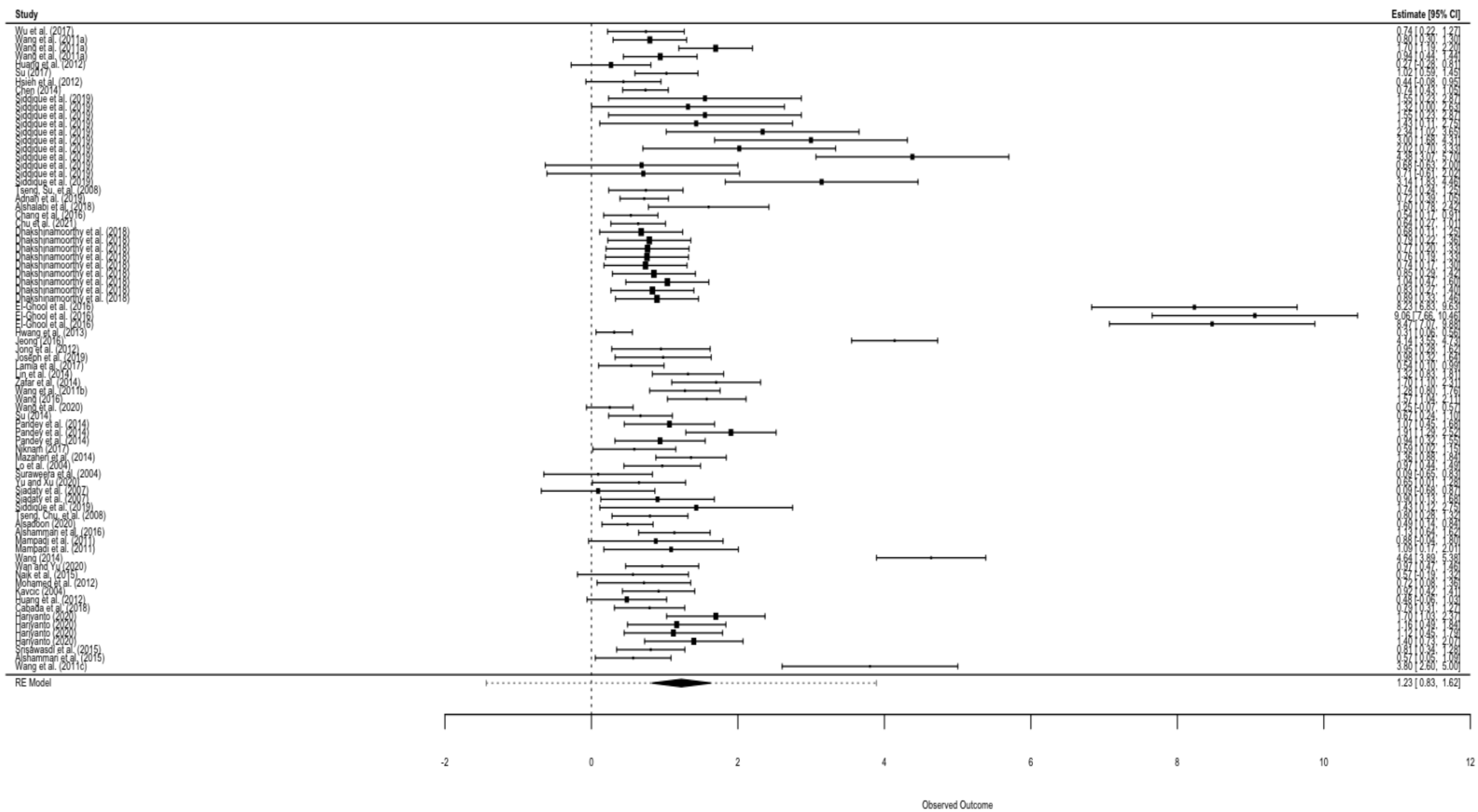

Figure 9. *A Forest Plot of the Hedges's g Estimates and the Overall Estimate in the Reviewed Studies*

heterogeneity, respectively (see Table 4). Thus, the between-study variability was the major source of the heterogeneity in ALS effectiveness. Moderator analyses are necessary to test variables that can explain the between-study heterogeneity. The section below presented the effects of five moderators (i.e., learner characteristic, modeling approach, subject area, publication source, and publication period) on the heterogeneity in system effectiveness.

Moderator Analyses of System Effectiveness

Learner Characteristic

Given that only two of the 46 publications modeled cognitive styles, we compared the effectiveness of ALSs that modeled prior knowledge and/or learning style. Publications that exclusively modeled prior knowledge and/or learning style were classified into three categories: only learning style (6 *effect sizes*), only prior knowledge (17 *effect sizes*), and both of them (14 *effect sizes*). We performed a subgroup analysis by entering the learner characteristic indicator into the three-level CHE model. The subgroup analysis results are summarized in Table 5.

A test of the three-level CHE model showed that the model was able to explain a significant portion of the variation in the 77 effect sizes ($Q = 9.155$, $df = 2$, $p < .05$), although it was not able to explain all the heterogeneity ($Q = 72.149$, $df = 34$, $p < .001$). This result indicates that learner characteristic is a variable that can significantly moderate the effectiveness of ALSs. In addition, the average effect sizes are 1.067 (95% CI = 0.689, 1.444) for ALSs modeling both prior knowledge and learning style, 0.737 (95% CI = 0.552, 0.921) for ALSs merely modeling prior knowledge, and 0.572 (95% CI = 0.275, 0.869) for ALSs only modeling learning style. All these effect sizes are significantly different from zero. Moreover, the expected average effect size of ALSs modeling only learning style is 0.495 points ($Z = -2.019$, $p < .05$) lower than that of

ALSs modeling both prior knowledge and learning style, with the standard error 0.245 and a confidence interval of -0.975 and -0.014.

Table 5

Effects of the Learner Characteristic Modeling on System Effectiveness

Moderator	Subgroup Analysis						Heterogeneity		
	ES (*g*)	**Standard Error**	**95% Lower**	**95% Upper**	**Z Value**	***p* Value**	***Q* Value**	***df***	***p* Value**
PK and LS	1.067	0.193	0.689	1.444	5.539	$< .001$			
PK only	0.737	0.094	0.552	0.921	7.831	$< .001$			
LS only	0.572	0.151	0.275	0.869	3.779	$< .001$			
Total Between							9.155	2	$< .05$

Note: ES = effect size; *g* = Hedges's *g*; PK = prior knowledge; LS = learning style.

The moderator analysis results presented above showed that the current ALSs that model both prior knowledge and learning style are likely to have significantly higher system effectiveness than those only modeling learning style. Although the current ALSs that model both prior knowledge and learning style have higher system effectiveness in comparison with those only modeling prior knowledge, their difference is insignificant. Altogether, these results indicate that merely modeling learning style is likely to relate to the lower effectiveness of ALSs, in comparison with systems modeling both prior knowledge and learning style.

Modeling Approach

Considering that only two of the 46 publications modeled cognitive styles, this section focused on the investigation of modeling approaches of prior knowledge and learning style. There are 14 studies (17 *effect sizes*) that specifically modeled prior knowledge and five studies that exclusively modeled learning styles (6 *effect sizes*). We performed a subgroup analysis to compare the effectiveness of ALSs that used different approaches to modeling prior knowledge. In terms of learning style, we presented descriptive statistics of the effectiveness of ALSs that employed different modeling approaches.

To investigate moderating effects of modeling approaches of prior knowledge in ALSs, publications that exclusively modeled prior knowledge were classified into four groups: overlay modeling (7 *effect sizes*), scalar modeling (2 *effect sizes*), educational data mining modeling (3 *effect sizes*), and item response theory modeling (2 *effect sizes*). We performed a subgroup analysis by entering the modeling approaches indicator into the three-level CHE model. The subgroup analysis results are summarized in Table 6.

As shown in Table 6, ALSs that employed educational data mining approach to modeling prior knowledge were associated with higher effectiveness in improving student learning outcomes (1.144; 95% CI = 0.608, 1.681); in contrast, ALSs that used item response theory modeling approach generated lower effectiveness (0.375; 95% CI = -0.250, 1.000). In addition, the difference in effect sizes between ALSs using the educational data mining approach and those employing the item response theory modeling approach is marginally significant (Z = -1.830, p = .067). Furthermore, the average effect size estimate of ALSs using the item response theory modeling approach is insignificantly different from zero (Z = 1.176, p = .240). There is no significant difference in effect sizes among ALSs using the educational data mining modeling, overlay modeling, and scalar modeling approach.

Table 6

Effects of the Modeling Approach of Prior Knowledge on System Effectiveness

Moderator	Subgroup Analysis						Heterogeneity		
	ES (*g*)	Standard Error	95% Lower	95% Upper	Z Value	*p* Value	*Q* Value	*df*	*p* Value
EDM	1.144	0.274	0.608	1.681	4.179	< .001			
IRT	0.375	0.319	-0.250	1.000	1.176	.240			
Overlay	0.912	0.145	0.629	1.196	6.304	< .001			
Scalar	0.624	0.253	0.128	1.121	2.463	< .05			
Total Between							4.332	3	.228

Note: ES = effect size; *g* = Hedges's *g*; EDM = educational data mining; IRT = item response theory.

Regarding the comparison of modeling approaches of learning style in ALSs, it is unfeasible to perform statistical modeling analyses because of the small sample size of available publications (i.e., $n = 5$). Rather than presenting statistical differences in the effectiveness of ALSs with different modeling approaches, we showed an overview of the effectiveness of the five ALSs in Table 7. Among the five ALSs, one ALS modeled learning style with Kolb's Learning Style Scale (Kolb, 1985), and one ALS used Jackson's Learning Styles Profiler Questionnaire (Jackson, 2005) to model learning style. The other three ALSs used the Index of Learning Styles Scale (Felder & Spurlin, 2005) to profile learners' learning styles.

The average effect size of ALSs that modeled learning style using the Index of Learning Styles scale (0.473) is lower than those using Kolb's Learning Style Scale (1.020) or Jackson's Learning Styles Profiler Questionnaire (0.497). In fact, all three ALSs that modeled both learning style and prior knowledge utilized the Index of Learning Styles Scale to model learning styles. Although the Index of Learning Styles Scale is commonly used in ALSs to model learning styles, ALSs might be not as effective in enhancing learning outcomes as others using Kolb's Learning

Style Scale or Jackson's Learning Styles Profiler Questionnaire. However, these results from descriptive statistical analysis require further investigation.

Table 7

Overview of ALSs Modeling Learning Style

	No. of Studies	**No. of ES**	***M***	***SD***
Kolb’s Learning Style Scale	1	1	1.020	NA
Jackson's Learning Styles Profiler Questionnaire	1	2	0.497	0.574
Index of Learning Styles Scale	3	3	0.473	0.141
Index of Learning Styles Scale*	3	14	0.990	0.290

Note: NA = not applicable; * denotes studies modeled both learning style and prior knowledge.

Subject Area

In this section, we examined the effect of the subject area variable on the effectiveness of ALSs. Publications were classified into five groups based on their subject areas: computer science (38 *effect sizes*), the English language (23 *effect sizes*), mathematics (5 *effect sizes*), and technology/science (4 *effect sizes*). We performed a subgroup analysis by entering the subject area indicator into the three-level CHE model. As shown in Table 8, the test of the three-level CHE model indicated that the variable of subject area was a significant moderator for system effectiveness (Q = 9.598, df = 3, p < .05).

More specifically, the expected mean effect sizes of ALSs aiming to improve the English language learning (2.483; 95% CI = 1.571, 3.395) and the learning of the computer science area (0.981; 95% CI = 0.453, 1.508) are significantly higher than zero. In terms of English language learning, adaptive systems aim to improve learners’ vocabulary, grammar, and reading. In contrast, the estimated mean effect sizes of ALSs designed for subject areas of mathematics

(0.684; 95% CI = -0.451, 1.819) and technology/science (0.801; 95% CI = -0.461, 2.063) are insignificantly different from zero. In addition, comparisons among effect sizes of the four subgroups show that the mean effect size for studies improving the English language learning is significantly higher than those targeting other subject areas (i.e., computer science, mathematics, and technology/science). That is, ALSs included in this study and targeted to improve the English language learning are likely to have more positive effects on improving student learning outcomes. However, there is no significant difference in the average effect sizes of ALSs for the other three subject areas (i.e., computer science, mathematics, and technology/science).

Table 8

Effects of the Subject Area on System Effectiveness

Moderator	**Subgroup Analysis**						**Heterogeneity**		
	ES (*g*)	**Standard Error**	**95% Lower**	**95% Upper**	**Z Value**	***p* Value**	***Q* Value**	***df***	***p* Value**
CS	0.981	0.269	0.453	1.508	3.644	< .001			
English Language	2.480	0.459	1.580	3.381	5.399	< .001			
Mathematics	0.684	0.579	-0.451	1.819	1.181	.237			
Technology/ Science	0.801	0.644	-0.461	2.063	1.244	.214			
Total Between							9.598	3	< .05

Note: ES = effect size; *g* = Hedges's *g*; CS = computer science.

Publication Source

This section examined the effect of the publication source variable on the effectiveness of ALSs. The 46 publications were categorized into four groups based on their publication source: journal article (*n* = 32), conference proceeding article (*n* = 10), dissertation/thesis (*n* = 3), and book chapter (n = 1). Given that only one article was published as a book chapter, the moderator analysis focused on comparisons of ALSs within the other three categories. We performed a subgroup analysis by adding the publication source indicator into the three-level CHE model. As shown in Table 9, the test of the three-level CHE model indicated that the variable of publication source is an insignificant moderator of system effectiveness (Q = 1.067, df = 2, p < .587).

Table 9

Effects of the Publication Source on System Effectiveness

Moderator	**Subgroup Analysis**						**Heterogeneity**		
	ES (*g*)	**Standard Error**	**95% Lower**	**95% Upper**	**Z Value**	***p* Value**	***Q* Value**	***df***	***p* Value**
Journal	1.371	0.244	0.894	1.849	5.625	<.001			
Dissertation/ Thesis	1.176	0.794	-0.381	2.732	1.481	.139			
Conference Proceeding	0.858	0.436	0.004	1.712	1.969	<.05			
Total Between							1.067	2	.587

Note: ES = effect size; g = Hedges's g.

Specifically, the journal article group that represents the vast majority of the 46 publications has an overall effect size at d = 1.371 (95% CI = 0.894, 1.849), p < .001. The average effect sizes for the conference proceeding article group and the dissertation/thesis group are 0.858 (95% CI = 0.004, 1.712) and 1.176 (95% CI = -0.381, 2.732). Although there is no significant difference in the effectiveness of ALSs published in the form of a journal, conference proceeding, and dissertation/thesis, the average effect size of ALSs published as dissertation/thesis is not significantly different from zero, which is different from those published as journal articles or conference proceeding articles.

Publication Period

The effect of the publication period variable on the effectiveness of ALSs was also examined in this study. Publication years of the 46 articles were grouped into three publication periods: 2004 to 2010 (n = 6), 2011 to 2015 (n = 19), and 2016 to 2021 (n = 21). We performed a subgroup analysis by entering the publication period indicator into the three-level CHE model. The subgroup analysis results are summarized in Table 10.

Table 10

Effects of the Publication Period on System Effectiveness

Moderator	Subgroup Analysis						Heterogeneity		
	ES (g)	Standard Error	95% Lower	95% Upper	Z Value	p Value	Q Value	df	p Value
2004 to 2010	0.672	0.556	-0.419	1.763	1.208	.227			
2011 to 2015	1.216	0.312	0.605	1.828	3.897	< .001			
2016 to 2021	1.396	0.296	0.816	1.977	4.713	< .001			
Total Between							1.322	2	.516

Note: ES = effect size; g = Hedges's g.

As shown in Table 10, the statistics of Q = 1.322, df = 2, p = .516 suggest that the between-level difference was not statistically significant for the publication year. However, the estimated effect sizes of ALSs have increased over the last twenty years. Specifically, the estimated effect size of ALSs in publications between 2004 to 2010 is 0.672 (95% CI = -0.419, 1.763), which was not significantly different from zero; in contrast, the estimated effect sizes of ALSs in publications after 2010 are significantly higher than zero (1.216 for publications between 2011 to 2015 and 1.396 for publications between 2016 to 2021).

Summary

By performing the meta-analysis of the effectiveness of ALSs on student learning outcomes, we found that the existing adaptive learning environments certainly enhance student learning outcomes, no matter whether compared with large-group traditional classroom instruction or non-adaptive computer-based instruction. In addition, we found that the effectiveness of ALSs is moderated by variables of learner characteristics, modeling approaches, and subject areas. The next chapter discusses these major findings and expands them into the field of ALSs to provide insights into educational implications and future research directions.

Chapter 5 Discussion

Chapter 5 restates the purpose of the study, discusses the major findings, and expands into theoretical and practical implications. It also points out the limitations of the study and presents suggestions for future research on establishing adaptive learning environments and on investigating the effectiveness of adaptive learning environments.

Purpose of the Study

In response to the COVID-19 pandemic, lots of learners have shifted from face-to-face learning environments to e-learning. As an extension of e-learning, the field of ALSs has gained its importance and prevalence in the present education scenario. However, the effectiveness of ALSs in promoting student learning outcomes is unclear and little work has examined variables that impact the effectiveness of ALSs in this growing field. To fill these research gaps, the twofold purpose of this study was: (a) to estimate the effectiveness of ALSs in enhancing student learning outcomes and (b) to identify moderators of the effectiveness of ALSs. Following the PRISMA statement, six databases were searched and a total of 46 publications with 77 effect sizes were included in the meta-analysis. With the three-level CHE modeling, the overall effect size for the effectiveness of ALSs was estimated. The estimated effect size provides evidence that ALSs certainly have more effects on improving student learning outcomes in comparison with large-group traditional classroom instruction and non-adaptive computer-based instruction. However, we should be careful in interpreting the overall estimated effect size because the effectiveness of ALSs in enhancing student learning performance differed considerably across the 46 publications. Analyses of moderator effects revealed that the effectiveness of ALSs was moderated by learner characteristics, modeling approaches, and subject areas.

Discussion of the Findings

The last four decades have witnessed important developments in the area of technology-enhanced learning systems, among which ALSs are of great interest to researchers given their potential to personalize learning. The theoretical structures and the practical implementations of ALSs have been well documented in the literature. As noted in Chapter 2, the layered evaluation structure (Brusilovsky et al., 2004) and the tripartite structure (Vandewaetere et al., 2011) depict common components of diverse ALSs. Starting with learner modeling that both structures highlight across different ALSs, this study investigates how the design of learner modeling (i.e., learner characteristics and modeling approaches) affects the effectiveness of ALSs in practice. This section discusses the findings, which add to the growing theory and empirical knowledge base concerning the continued development and implementation of ALSs.

Characteristics of Included Studies

This meta-analysis included 46 publications, most of which investigate participants from the postsecondary level of institutions of higher education. The distribution of grade level is consistent with other meta-analyses that investigate the effectiveness of technology-enhanced learning systems (e.g., Kulik & Fletcher, 2016; Ma et al., 2014). This finding is unsurprising as researchers are housed in institutions of higher education and may find accessing student populations easier than collecting relevant data in K-12 schools. In addition, the estimated effect sizes showed positive results in favor of ALSs' effectiveness in helping both K-12 and postsecondary students. Furthermore, this study did not detect a significant difference in the estimated effect sizes between K-12 and postsecondary levels.

The subject of computer science is the most common learning topic within the 46 publications, followed by the English language. It was found that ALSs focusing on the English language instruction overperformed ALSs concentrating on other disciplines (i.e., mathematics, computer science, science/technology, and database). It is interesting to note that ALSs designed for the computer science learning appear not effective as those established for the English language learning, although the computer science community has been a vocal advocate of ALSs. Because many studies on ALSs lack the requisite information for this meta-analysis, it is likely that the 46 articles included in the present study are not representative of the full spectrum of the existing ALSs. As shown in Chapter 3, 2663 of 2709 studies were excluded because of the absence of mean scores, standard deviations, the number of observations, or descriptions of their architecture. Although the overall estimated effect sizes of ALSs designed for the five disciplines were positive, ALSs that target to help students learn mathematics, science/technology, and database did not significantly perform better than non-adaptive instruction.

Unlike previous systematic reviews (Nakic et al., 2015, Martin et al., 2020, Muñoz et al., 2022), prior knowledge is the most frequently modeled learner characteristic of the 46 publications rather than learning style. This inconsistency may be a result of different inclusion criteria involved in previous systematic reviews and this study. The systematic reviews performed by Nakic et al. (2015), Martin et al. (2020), and Muñoz et al. (2022) did not require eligible studies to report quantitative information for the calculation of effect sizes. However, reporting quantitative information about empirical experiments for the calculation of effect sizes is an important selection criterion in this study. In addition, researchers typically used the overlay modeling approach to profiling learners' knowledge status and questionnaires to model learning styles in ALSs. Although modern computing power and interaction capabilities have led to a broadening in learner modeling techniques (e.g., educational data mining and learning analytics), the results of this study showed that these emerging learner modeling techniques have not been the mainstream in the field of ALSs.

The Effectiveness of ALSs

The overall estimated effect size (i.e., Hedges's *g*) of the 46 publications was found to be substantial (1.227; 95% CI = 0.832, 1.622). The other recently published meta-analysis on ALSs (Fontaine et al., 2019) also showed positive, statistically significant overall effect sizes ranging from .70 (for knowledge) to 1.19 (for skill). The inconsistency in effect sizes between the present study and Fontaine et al.'s (2019) study may be a function of the choice of databases for their search strategies: The present research did not include academic medical sources (e.g., PubMed); in contrast, Fontaine et al. (2019) limited their search to medical databases. Although their effect sizes are different, both of them confirmed the effect of ALSs in promoting student learning outcomes is considerable. In comparison with the earlier study conducted by Verdú et al. (2008), the effectiveness of ALSs seems to be improved from a medium level (e.g., around 0.5 of Cohen's *d*) to a large level (e.g., higher than 0.5 of Cohen's *d*) over the last few years.

This review also compared ALSs to other instructional strategies in terms of their effectiveness. First, this study compared the estimated effect size of ALSs in this study and the effect sizes of intelligent tutoring systems reported by previous meta-analyses. Specifically, Ma et al. (2014) found that the use of intelligent tutoring systems was associated with greater achievement in comparison with teacher-led large-group instruction (Hedges's $g = 0.42$) and non-intelligent tutoring system computer-based instruction (Hedges's $g = 0.57$). Kulik and Fletcher (2016) found that the median effect size in 50 studies on intelligent tutoring systems was .66 (Glass's *delta*). In this study, the use of ALSs was associated with greater achievement in comparison with traditional large-group instruction (Hedges's $g = 1.025$) and non-adaptive computer-based instruction (Hedges's $g = 1.206$). Based on these findings, ALSs appear to have more effects on learning achievement than intelligent tutoring systems.

In addition, this study compared the estimated effect size of ALSs in this study to the effect size of one-on-one human instruction reported by Bloom (1984). Specifically, Bloom (1984) reported a two-sigma effect size (i.e., Glass's *delta* = 2.00) of one-on-one human instruction in replacement of classroom teaching. In this study, the use of ALSs was associated with about one-sigma effect size (i.e., Glass's *delta* = 1.30) in comparison with teacher-led large-group instruction. Thus, ALSs appear to have fewer effects on student learning achievement than one-on-one human instruction. Nevertheless, it is interesting to observe that the effectiveness of ALSs has been stably improved in the recent five years. Specifically, the publications within the year 2016 to 2022 yielded an average effect size of 1.396 (Glass's *delta*) as a result of the continued development in the field of ALSs.

We should be cautious about interpreting the overall estimated effect size of ALSs in this study. This is because the 46 publications did not produce homogeneous effect sizes about the use of ALSs. Instead, the effectiveness of ALSs varied significantly among different publications. The three-level CHE modeling results, which indicated the effect sizes between publications (level 3; 83.47%) contributed a larger percentage to the overall heterogeneity than that found within publications (level 2; 12.06%), confirmed the heterogeneity in effect sizes from the 46 publications. This finding highlights the wide range of ALSs in promoting student learning and indicates that a better understanding of influential factors would be helpful to advance this area of work (Brusilovsky et al., 2004; Vandewaetere et al., 2011). As Bernacki et al. (2021) claimed, significant opportunities exist to improve ALSs before systems can be applied in education. To achieve this goal, it is crucial to inspect factors that account for the variability in system effectiveness. Given that most effect sizes within one publication shared many attributes (e.g., common system architecture) and the main source of heterogeneity is between publications, it is important to investigate publication factors.

Learner Modeling Accounting for the Heterogeneity

Previous literature reviews consistently suggested learner modeling to be a crucial component of ALSs. For example, Vandewaetere et al. (2011) pointed out that the difference among various ALSs centered on learner characteristics and modeling approaches. This theory-guide emphasis on the importance of learner modeling to ALSs was confirmed by empirical findings obtained in this study. Analyses of design features of learner modeling within the 46 publications revealed that the choice of learner characteristics and modeling approaches significantly accounted for the heterogeneity in system effectiveness. Specifically, ALSs that model both learning style and prior knowledge are likely to enhance student learning achievement more than those only model learning style. This finding is consistent with the statement that the effectiveness of ALSs improves with an increased number of learner characteristics considered to build systems (Dhakshinamoorthy & Dhakshinamoorthy, 2019).

The observations from the existing adaptive systems reveal that the development of systems that adapt to more than one learner characteristic is more challenging than systems that adapt to a single learner characteristic (Brusilovsky et al., 2004). If only one learner characteristic can be included in the learner model of an ALS, our results suggested that prior knowledge seems to be associated with better learning outcomes in comparison with learning style. In addition, educational data mining methods are associated with more effects of ALSs on student learning outcomes, in comparison with other modeling approaches to prior knowledge (i.e., overlay modeling, scalar modeling, and item response theory modeling).

However, the comparison results of system effectiveness should be interpreted with caution because studies included in this meta-analysis typically used questionnaires to model learning styles. In the literature, there have been some studies (e.g., Botsios et al., 2008; Dorça et al., 2013) that applied machine learning algorithms to automatically detect learners' learning styles. The effectiveness of ALSs which model learning styles using machine learning algorithms and those modeling prior knowledge deserves further comparisons.

Taken together, these findings highlight the importance of system feature specification when embarking on ALSs work in education.

Educational Implications

Instead of inventing a new ALS, this study systematically evaluated established ALSs. The findings on the system effectiveness and the influential factors of the system effectiveness contribute to the increasing theory and empirical knowledge concerning ALSs.

Theoretical Implications

The synthesis of the effectiveness of different ALSs facilitates a comprehensive evaluation of the current ALSs. In addition to the meta-analysis conducted by Fontaine et al. (2019), to my knowledge, this is the second meta-analysis of empirical studies on ALSs. Fontaine et al. (2019) limited their focus to health professionals and students. In contrast, this study investigated the effects of ALSs on K-12 and postsecondary students' learning achievements. Thus, this study acts as an important supplement to Fontaine et al.'s (2019) study concerning the evaluation of ALSs.

The evaluation of ALSs revealed that ALSs appear to be associated with more effects in improving student learning achievement than those of intelligent tutoring systems. However, the current ALSs do not seem to be comparable to one-on-one human instruction in promoting student learning outcomes. As the exemplary method of personalized learning, one-on-one human instruction is expected to provide "the right content at the proper time in the most appropriate way" (National Association of State Boards of Education, 2001). ALSs may be not capable to deliver adaptive instruction to the degree of human instructors. There have been studies that investigated important components that make one-to-one instruction one of the most effective tutoring methods (e.g., Graesser et al., 1995; Zhang et al., 2021). For example, one prominent component of effective human tutoring is the collaborative dialogue between learners and tutors (Graesser et al., 1995). However, few ALSs carry out the function of dialogue. Thus, continued development of ALSs, especially in the improvement of system adaptation, is warranted to be a parallel alternative to one-on-one human instruction. The improved effectiveness of ALSs in the last seven years indicates a promising trend for the future of ALSs.

The comparison result between ALSs and intelligent tutoring systems needs to be interpreted carefully because these systems cannot be clearly separated in the literature. ALSs and ITSs are two overlapping fields and the boundary between them is blurred. In comparison with Ma et al.'s (2014) results, the effect sizes in this study were higher when compared to non-adaptive computer instruction and traditional large-group instruction. However, Ma et al. (2014) reviewed evaluative studies on intelligent tutoring systems published prior to 2013. The different publication year ranges might be the reason for the inconsistency in system effectiveness. Although the comparison between ALSs and intelligent tutoring systems is impracticable, a feasible research direction is to investigate the effects of the inner loop and the outer loop feedback on learning process and performance. It is uncommon for ALSs to offer the inner loop feedback to learners. However, Tacoma et al. (2020) suggested that providing both the inner loop and the outer loop feedback would benefit students more than merely delivering a single type of feedback. A great deal is still unknown and requires further investigations about which design options (i.e., inner loop and outer loop) are effective under which circumstances (e.g., knowledge level, knowledge component).

Analysis results of factors that are associated with the effectiveness of ALSs add to the growing theory concerning the architecture of ALSs. Specifically, the discussion on the similarities and differences between the layered evaluation structure (Brusilovsky et al., 2004) and the tripartite structure (Vandewaetere et al., 2011) in Chapter 2 provides insights into the core components of ALSs. That is, the importance of learner modeling to ALSs is emphasized theoretically. This study performed a series of three-level CHE models and found the significant effects of learner characteristics and modeling approaches on the effectiveness of ALSs. This is the first study that supported the significance of learner modeling to ALSs empirically. These results account for the breadth of challenges when approaching technology-enhanced learning in education (Cha et al., 2006; Xie et al., 2019): not only implementations but also designs of adaptive learning require deliberate efforts in the research communities.

Practical Implications

The findings on the estimated effectiveness of ALSs offer practical implications for practitioners and policymakers. Along with the prevalence of computers and improved Internet connectivity in schools, ALSs present a viable option for providing accessible personalized instruction for students. Although the use of ALSs is associated with fewer effects in promoting student learning achievement in comparison with one-on-one human instruction, ALSs certainly promote student learning performance in comparison with non-adaptive computer instruction and traditional large-group instruction. In addition, like other technology-enhanced learning systems, ALSs may provide a good opportunity for under-served and geographically dispersed populations who may otherwise not have easy access to well-trained teachers.

In addition, the mapping between system effectiveness and influential factors is useful for designers and developers of ALSs. To establish an ALS, system designers and developers need to select suitable learner characteristics and modeling approaches since learning materials delivered by ALSs are to be adapted to the modeled learner characteristics with artificial intelligent mechanisms. The lack of empirical research on the association between system effectiveness and learner modeling is not only a key issue but also a challenging research area in the development of ALSs (Abyaa et al., 2019). However, the present study performed a meta-analysis of different ALSs, which makes it feasible to compare the option of different learner characteristics and modeling approaches in terms of system effectiveness. The findings of this study facilitate researchers and educators to better incorporate learner characteristics and modeling approaches in the development of ALSs.

Specifically, ALSs that model both learning style and prior knowledge are associated with more effects on student learning achievement, especially when compared with ALSs merely modeling learning style. In addition, prior knowledge has more potential in comparison to learning style profiled by questionnaires if only one learner characteristic is included in the learner model of an ALS. Moreover, in comparison with other modeling approaches to prior knowledge, educational data mining methods tend to be associated with more effects of ALSs on student learning outcomes. In terms of learning style questionnaires, Kolb's Learning Style Scale (Kolb, 1985) and Jackson's Learning Styles Profiler Questionnaire (Jackson, 2005) appear to be associated with larger effect sizes in comparison with the Index of Learning Styles Scale (Felder & Spurlin, 2005). These findings are useful for the decision-making process when system designers and developers only model one or two student characteristics from prior knowledge and learning style in their ALSs.

Limitations

The current research is characterized by limitations that must be considered when interpreting its findings. The limitations are categorized into two groups in this section: issues with theoretical framework and issues with quantitative evaluation.

Issues Related to Theoretical Framework

This study distinguished between adaptive learning systems and intelligent tutoring systems in Chapter 2. The term "adaptive learning system" was centered when databases were searched to assemble the candidate studies pool in Chapter 3. However, in the literature, it is still debatable whether intelligent tutoring systems belong to adaptive learning systems. For example, Xie et al. (2019) stated that one important stream of adaptive learning systems is intelligent tutoring systems. In contrast, when Kabudi et al. (2021) performed a systematic mapping of the literature on AI-enabled learning systems, they identified adaptive learning systems and intelligent tutoring systems as separate streams of AI-enabled learning systems.

As stated by Psotka et al. (1988), intelligent tutoring systems employ computational algorithms or models to deliver immediate feedback and learning instructions to learners. That is, intelligent tutoring systems are characterized by step-specific scaffolding strategies (e.g., error-correction feedback and cues; Nye, 2015). However, from the standpoint of both the layered evaluation structure (Brusilovsky et al., 2004) and the tripartite structure (Vandewaetere et al., 2011), adaptive effects mainly include adaptive presentations, adaptive navigations, and adaptive content aggregations (Esichaikul et al., 2011; Premlatha & Geetha, 2015). Because intelligent tutoring systems also deliver step-specific scaffolding strategies, which are not specified by the two structures, this study specifically focused on ALSs.

However, we cannot draw a clear boundary between intelligent tutoring systems and adaptive learning systems in fact. From the perspective of adaptation, intelligent tutoring systems typically provide both inner-loop and outer-loop adaptation; in contrast, ALSs usually serve outer-loop adaptation. Their advantages in personalized learning are gradually merged to provide a full learning environment (Phobun & Vicheanpanya, 2010). Future studies might include intelligent tutoring systems and adaptive learning systems and compare the effectiveness of inner-loop, outer-loop, and/or a combination of inner-loop and outer-loop adaptation.

Issues Related to Quantitative Evaluation

One of the limitations related to the meta-analysis is that this study merely examined the effectiveness of ALSs in promoting student learning outcomes. However, there are other aspects of ALSs' evaluation. For example, learning satisfaction is found to be an important factor that is able to predict learners' decision to drop out or persist in e-learning (Park & Choi, 2009). In real education scenarios, the high dropout rate in e-learning has been of concern to many educational institutions and organizations (Njenga & Fourie, 2010). Thus, exploring the association between learning satisfaction and learner modeling has enormous potential for the development of ALSs.

Another limitation related to the meta-analysis is that this study did not investigate the association between adaptive effects and the effectiveness of ALSs. Like learner modeling, which was investigated in the present study, adaptive effects are important components of an ALS (Brusilovsky et al., 2004; Vandewaetere et al., 2011). However, because researchers do not report adaptive effects explicitly, it is difficult to code adaptive effects of an ALS. Specifically, studies on ALSs always mix together adaptive presentations, adaptive navigations, and adaptive content aggregations. However, to advance our understanding of ALSs, it appears necessary to examine the impacts of different types of adaptive effects on system effectiveness.

Finally, this study examined the effects of multiple moderator one by one. However, interactions among moderators are possible. Although interaction effects between moderators might be meaningful to understanding ALSs' effectiveness, Dusseldorp et al. (2016) pointed out that "when several study features are available, regression in meta-analysis lacks sufficient power to detect interactions between them (p. 1)". As the body of work continues to grow, future studies can explore feature interactions. For example, interactions between learner characteristics and modeling approaches are possible, making further explorations of these factors essential for improving system effectiveness and building more effective ALSs needed for the changing landscape of education efforts. In addition, although researchers are suggested to give priority to prior knowledge when selecting modeled learner characteristics in ALSs, it is unclear whether learning style is a more important learner characteristic in specific subject areas.

Recommendations for Future Research

There are at least four directions for future research. First, future studies might focus on comparisons between intelligent tutoring systems and adaptive learning environments. As discussed above, some researchers consider intelligent tutoring systems as a stream of adaptive learning environments because intelligent tutoring systems are characterized by system adaptation (Xie et al., 2019); in contrast, some researchers regard them as separate AI-enabled learning systems because of different adaptive effects provided by them (Kabudi et al., 2021). A key factor that researchers consider whether to include studies on intelligent tutoring systems might be what research questions they would like to answer. For example, studies should incorporate intelligent tutoring systems when they aim to provide a comprehensive review of the present AI techniques of system adaptation (Imhof et al., 2020; Kabudi et al., 2021). In contrast, studies are suggested to leave intelligent tutoring systems out when they refer to the layered evaluation structure (Brusilovsky et al., 2004) and/or the tripartite structure (Vandewaetere et al., 2011). Future studies should conduct a review of adaptive learning environments and intelligent tutoring systems and clarify their connections, based on which propose guidelines on how to delimit the scope of research on adaptive learning environments.

Second, future studies might focus on comparing the effectiveness of ALSs using other modeling approaches. Specifically, this study found ALSs using an artificial neural network to model prior knowledge were associated with higher system effectiveness in comparison with other modeling approaches (i.e., overlay modeling, scalar modeling, and item response theory modeling). However, this finding cannot be generalized to other data mining techniques. Regarding learning style, publications included in this study typically employ different types of questionnaires. The effectiveness of ALSs modeling prior knowledge was found to be higher than those modeling learning styles with questionnaires. However, the effectiveness of ALSs that used machine learning algorithms to model learning styles has not yet been compared with the effectiveness of ALSs that modeled prior knowledge. It is possible that modeling approaches of learning styles rather than learning styles themselves should be taken seriously in the design of ALSs. That is, only specific modeling approaches could accurately capture learners' learning styles when they use ALSs. Future studies might conduct experiments to examine the effectiveness of ALSs, which model learning styles with machine learning algorithms, in comparison with non-adaptive instruction. These studies will contribute to profiling a comparative complete picture of learner modeling comparison (i.e., prior knowledge and learning style) in ALSs.

Third, future studies might focus on students' learning satisfaction with using ALSs. With improved internet connectivity, ALSs have gained prevalence in real educational settings. Along with the wide application of ALSs, students' learning satisfaction is becoming an important aspect of ALS evaluation. Low learning satisfaction is likely to increase the dropout rate of ALSs, thus deteriorating learning improvement (Park & Choi, 2009; Tan & Shao, 2015). In this study, the lack of sufficient information and the inconsistency of reporting across publications constitute barriers to the evaluation of learners' learning satisfaction from different ALSs. Future studies on the development of new ALSs might survey the learning satisfaction of students who use ALSs and report students' learning satisfaction as an indicator of system effectiveness. When there are more studies available on students' learning satisfaction with using ALSs, researchers may conduct a more comprehensive meta-analysis to investigate the relationship between different system characteristics and learning satisfaction.

Potential of Measurement Model for Adaptive Learning Systems

Psychometrics can be an important building block of ALSs, given its dominance in the measurement of proficiency levels. Chang (2015) claimed that the field of computerized adaptive testing (CAT) has enormous potential to greatly facilitate individualized learning. Specifically, providing more efficient latent trait estimates with fewer items by CAT (e.g., Weiss, 1982) would put ALSs at an advantage in terms of knowledge modeling. In addition, cognitive profiles identified by cognitive diagnostic computerized adaptive testing (CD-CAT) can be used to customize exercises and learning materials in ALSs. However, methodologies in the field of psychometrics for measuring underlying individuals' abilities are not commonly used in the present ALSs. This might be because the ALS community's participants are usually associated with computer science (Nakic et al., 2015).

Zhang and Chang (2016) indicated that it is necessary for adaptive learning to find its niche in the current educational institution; otherwise, both schools and students are challenged to take full advantage of adaptive learning. This warning corresponds to the basic question of "how to integrate technology into a traditional learning environment". Chang (2015) holds the perspective that technology should help classroom instructors rather than completely replace their roles. A good integration of ALSs and traditional learning scenarios will promote the acceptance of ALSs among schools, teachers, and students (Zhang & Chang, 2016). Wang et al. (2013) worked on making paper-and-pencil tests adaptive, which sets a good example of how to integrate technology into the classroom. Specifically, their design includes a PC server and a smart printer-scanner. Students' answers to paper-and-pencil tests are scanned into the system by the printer-scanner. Then, the system automatically scores students' answers and generates individualized diagnostic reports. Based on the diagnostic results, the system can also generate a stapled booklet that provides an assignment with personalized instructions to each student.

Conclusion

ALSs serve as a way to personalize learning experiences for students to improve their learning outcomes. The study performed a meta-analysis to investigate the potential of ALSs to promote student learning outcomes. Findings indicate that the current ALSs can advance student learning achievement greatly. However, they are not as effective as one-on-one human instruction. The high level of heterogeneity in the success of existing efforts for ALSs indicates many opportunities for continually improving their effectiveness. As found in the present research, the choice of learner modeling (i.e., learner characteristics and modeling approaches) significantly accounts for the wide variation in system effectiveness. Regarding learner characteristics, ALSs that model both learning style and prior knowledge are associated with higher student learning achievement than those that only model learning style. In addition, prior knowledge has more potential in enhancing student learning outcomes in comparison with learning styles modeled by questionnaires if only one learner characteristic is included in the learner model of an ALS. Regarding modeling approaches of prior knowledge, the educational data mining method (i.e., artificial neural network) was found to be associated with larger effect sizes of ALSs, in comparison with other modeling approaches (e.g., overlay modeling and item response theory modeling). These findings can act as evidence-based recommendations for ongoing and flourishing developments in ALSs.

References

Abyaa, A., Khalidi Idrissi, M., & Bennani, S. (2019). Learner modelling: Systematic review of the literature from the last 5 years. *Educational Technology Research and Development, 67*, 1105–1143. https://doi.org/10.1007/s11423-018-09644-1

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Akbulut, Y., & Cardak, C. S. (2012). Adaptive educational hypermedia accommodating learning styles: A content analysis of publications from 2000 to 2011. *Computers & Education, 58*(2), 835–842. https://doi.org/10.1016/j.compedu.2011.10.008

Al-Azawei, A., & Badii, A. (2014). State of the art of learning styles-based adaptive educational hypermedia systems (LS-BAEHSs). *International Journal of Computer Science & Information Technology, 6*(3), 1–19. https://doi.org/10.5121/ijcsit.2014.6301

Al-Bastami, B. G. H., & Naser, S. S. A. (2017). Design and development of an intelligent tutoring system for C# language. *European Academic Research, 4*(10), 8795–8809.

Aleven, V., & Koedinger, K. R. (2013). Knowledge component (KC) approaches to learner modeling. *Design Recommendations for Intelligent Tutoring Systems, 1*, 165–182.

Alfonseca, E., Carro, R. M., Martin, E., Ortigosa, A., & Paredes, P. (2006). The impact of learning styles on student grouping for collaborative learning: A case study. User *Modeling and User-Adapted Interaction, 16*(3), 377–401. https://doi.org/10.1007/s11257-006-9012-7

Almohammadi, K., Hagras, H., Alghazzawi, D., & Aldabbagh, G. (2017). A survey of artificial intelligence techniques employed for adaptive educational systems within e-learning platforms. *Journal of Artificial Intelligence and Soft Computing Research, 7*(1), 47–64. https://doi.org/10.1515/jaiscr-2017-0004

Alsadoon, E. (2020). The impact of an adaptive e-course on students’ achievements based on the students’ prior knowledge. *Education and Information Technologies, 25*, 3541–3551. https://doi.org/10.1007/s10639-020-10125-3

Alshammari, M., Anane, R., & Hendley, R. J. (2015). Design and usability evaluation of adaptive e-learning systems based on learner knowledge and learning style. In J. Abascal, S. Barbosa, M. Fetter, T. Gross, P. Palanque, & M. Winckler (Eds.), *Human-Computer Interaction – INTERACT 2015* (pp. 584–591). Springer: Cham. https://doi.org/10.1007/978-3-319-22668-2_45

Aslaksen, K., & Lorås, H. (2018). The modality-specific learning style hypothesis: A mini-review. *Frontiers in Psychology*, 1538. https://doi.org/10.3389/fpsyg.2018.01538

Barr, A., Beard, M., & Atkinson, R. C. (1976). The computer as a tutorial laboratory: The Stanford BIP project. *International Journal of Man-Machine Studies, 8*(5), 567–582. https://doi.org/10.1016/S0020-7373(76)80021-1

Bartolucci, A. A., & Hillegass, W. B. (2010). Overview, strengths, and limitations of systematic reviews and meta-analyses. In F. Chiappelli (Ed.), *Evidence-based practice: Toward optimizing clinical outcomes* (pp. 17–33). Berlin Heidelberg: Springer.

Beldagli, B., & Adiguzel, T. (2010). Illustrating an ideal adaptive e-learning: A conceptual framework. *Procedia-Social and Behavioral Sciences, 2*(2), 5755–5761. https://doi.org/10.1016/j.sbspro.2010.03.939

Bernacki, M. L., Greene, M. J., & Lobczowski, N. G. (2021). A Systematic Review of Research on Personalized Learning: Personalized by Whom, to What, How, and for What Purpose(s)? *Educational Psychology Review, 33*(4), 1675–1715. https://doi.org/10.1007/s10648-021-09615-8

Bingham, A. J., Pane, J. F., Steiner, E. D., & Hamilton, L. S. (2018). Ahead of the curve: Implementation challenge in personalized learning school models. *Educational Policy, 32*(3), 454–489. https://doi.org/10.1177/0895904816637688

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*(6), 4–16.

Botsios, S., Georgiou, D., & Safouris, N. (2008). Contributions to adaptive educational hypermedia systems via on-line learning style estimation. *Journal of Educational Technology & Society, 11*(2), 322–339.

Brusilovsky, P. (1996). Adaptive hypermedia: An attempt to analyze and generalize. In Proceedings of first international conference on multimedia, hypermedia and virtual reality, LNCS 1077 (pp. 288–304), Springer Verlag.

Brusilovsky, P. (2001). Adaptive hypermedia. *User Modeling and User-Adapted Interaction, 11*, 87–110. https://doi.org/10.1023/A:1011143116306

Brusilovsky, P., Karagiannidis, C., & Sampson, D. (2004). Layered evaluation of adaptive learning systems. *International Journal of Continuing Engineering Education and Lifelong Learning, 14*(4/5), 402–421.

Brusilovsky, P., & Maybury, M. T. (2002). From adaptive hypermedia to the adaptive web. *Communications of the ACM, 45*(5), 30–33.

Brusilovsky, P., & Millán, E. (2007). User models for adaptive hypermedia and adaptive educational systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The adaptive web: Methods and strategies of web personalization* (pp. 3–53). Springer: Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-72079-9_1

Brusilovsky, P., & Peylo, C. (2003). Adaptive and intelligent web-based educational systems. *International Journal of Artificial Intelligence in Education, 13*(2–4), 159–172.

Carbonell, J. (1970). AI in CAI: An artificial intelligence approach to computer aided instruction. *IEEE Transactions on Man-Machine Systems, 11*(4), 190–202. https://doi.org/ 10.1109/TMMS.1970.299942

Cha, H. J., Kim, Y. S., Park, S. H., Yoon, T. B., Jung, Y. M., & Lee, J. H. (2006). Learning styles diagnosis based on user interface behaviors for the customization of learning interfaces in an intelligent tutoring system. In M. Ikeda, K. D. Ashely, & T. W. Chan (Eds.), *Intelligent Tutoring Systems* (pp. 513–524). Springer, Berlin, Heidelberg.

Chang, H. H. (2015). Psychometrics behind computerized adaptive testing. *Psychometrika, 80*(1), 1–20. https://doi.org/10.1007/s11336-014-9401-5

Chen, C. M., Lee, H. M., & Chen, Y. H. (2005). Personalized e-learning system using item response theory. *Computers & Education, 44*(3), 237–255. https://doi.org/10.1016/j.compedu.2004.01.006

Cheng, L., Ritzhaupt, A. D., & Antonenko, P. (2019). Effects of the flipped classroom instructional strategy on students' learning outcomes: A meta-analysis. *Educational Technology Research and Development, 67*, 793–824. https://doi.org/10.1007/s11423-018-9633-7

Chou, C. Y., Lai, K. R., Chao, P. Y., Lan, C. H., & Chen, T. H. (2015). Negotiation based adaptive learning sequences: Combining adaptivity and adaptability. *Computers & Education, 88*, 215–226. https://doi.org/10.1016/j.compedu.2015.05.007

Chrysafiadi, K., & Virvou, M. (2013). Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications, 40*(11), 4715–4729. https://doi.org/10.1016/j.eswa.2013.02.007

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal and Social Psychology, 65*, 145153.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum. https://doi.org/10.4324/9780203771587

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and Instructional Methods: A Handbook for Research on Interactions*. Irvington.

Dede, C. (1986). A review and synthesis of recent research in intelligent computer-assisted instruction. *International Journal of Man-Machine Studies, 24*, 329–353. https://doi.org/10.1016/S0020-7373(86)80050-5

DeMink-Carthew, J., Olofson, M. W., LeGeros, L., Netcoh, S., & Hennessey, S. (2017). An analysis of approaches to goal setting in middle grades personalized learning environments. *RMLE Online*, *40*(10), 1–11. https://doi.org/10.1080/19404476.2017.1392689

Desmarais, M. C., & d Baker, R. S. (2012). A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction, 22*(1), 9–38. https://doi.org/10.1007/s11257-011-9106-8

Dewey, J. (1929). My pedagogic creed. *Journal of the National Education Association, 18*(9), 291–295.

Dhakshinamoorthy, A., & Dhakshinamoorthy, K. (2019). KLSAS—An adaptive dynamic learning environment based on knowledge level and learning style. *Computer Applications in Engineering Education, 27*(2), 319–331. https://doi.org/10.1002/cae.22076

Dorça, F. A., Lima, L. V., Fernandes, M. A., & Lopes, C. R. (2013). Automatic student modeling in adaptive educational systems through probabilistic learning style combinations: a qualitative comparison between two innovative stochastic approaches. *Journal of the Brazilian Computer Society, 19*(1), 43–58.

Dusseldorp, E., Li, X., & Meulman, J. (2016). Which combinations of behaviour change techniques are effective assessing interaction effects in meta-analysis. *European Health Psychologist, 18*, 563.

Duval, S. J., & Tweedie, R. L. (2000a). A non-parametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*, 89–98. https://doi.org/10.1080/01621459.2000.10473905

Duval, S. J., & Tweedie, R. L. (2000b). Trim and Fill: A simple funnel plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 276–284. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Dziuban, C., Howlin, C., Moskal, P., Johnson, C., Parker, L., & Campbell, M. (2018). Adaptive learning: A stabilizing influence across disciplines and universities. *Online Learning, 22*(3), 7–39. https://doi.org/10.24059/olj.v22i3.1465

Dziuban, C., Moskal, P., Johnson, C., & Evans, D. (2017). Adaptive learning: A tale of two contexts. *Current Issues in Emerging eLearning, 4*(1), 26–62.

Egger, M., Smith, G. D., & Phillips, A. N. (1997). Meta-analysis: *principles and procedures*. *BMJ, 315*(7121), 1533–1537. https://doi.org/10.1136/bmj.315.7121.1533

Erümit, A. K., & Çetin, İ. (2020). Design framework of adaptive intelligent tutoring systems. *Education and Information Technologies, 25*, 4477–4500. https://doi.org/10.1007/s10639-020-10182-8

Esichaikul, V., Lamnoi, S., & Bechter, C. (2011). Student modelling in adaptive e-learning systems. *Knowledge Management & E-Learning: An International Journal, 3*(3), 342–355. https://doi.org/10.34105/j.kmel.2011.03.025

Essa, A., & Laster, S. (2017). Bloom's 2 Sigma problem and data-driven approaches for improving student success. In R. Feldman (Eds.), *The first year of college: Research, theory, and practice on improving the student experience and increasing retention* (pp. 212-246). Cambridge: Cambridge University. https://doi.org/10.1017/9781316811764.009

Felder, R. M., & Silverman, L. (1988). Learning and teaching styles in engineering education. *Engineering Education, 78*(7), 674–681.

Felder, R. M., & Spurlin, J. (2005). Applications, reliability and validity of the index of learning styles. *International Journal of Engineering Education, 21*(1), 103–112.

Fleming, N. D. (2001). *Teaching and learning styles: VARK strategies*. Christchurch, New Zealand: Author

Fleming, N. D., & Mills, C. (1992). Not another inventory, rather a catalyst for reflection. *To Improve the Academy, 11*(1), 137–155. https://doi.org/10.1002/j.2334-4822.1992.tb00213.x

Fontaine, G., Cossette, S., Maheu-Cadotte, M. A., Mailhot, T., Deschênes, M. F., Mathieu-Dupuis, G., ... & Dubé, V. (2019). Efficacy of adaptive e-learning for health professionals and students: a systematic review and meta-analysis. *BMJ open*, *9*(8), e025252. http://dx.doi.org/10.1136/bmjopen-2018-025252

Ford, N. (1985). Learning styles and strategies of postgraduate students. *British Journal of Educational Technology, 16*(1), 65–79.

Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of augmented reality on students' learning gains. *Educational Research Review, 27*, 244–260. https://doi.org/10.1016/j.edurev.2019.04.00

Ge, X., Law, V., & Huang, K. (2012). Diagnosis, supporting, and fading: A scaffolding design framework for adaptive e-learning systems. In H. Wang (Eds.), *Interactivity in E-*

learning: Case Studies and Frameworks (pp. 116–142). Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-61350-441-3.ch006

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3–8.

Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research in Higher Education, 5*(1), 351–379.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied cognitive psychology, 9*(6), 495–522.

Graf, S., Kinshuk, Zhang, Q., Maguire, P., & Shtern, V. (2010). Facilitating learning through dynamic student modelling of learning styles. In P. Isaias, D. Ifenthaler, Kinshuk, D. Sampson, & J. Spector (Eds.), *Towards Learning and Instruction in Web 3.0* (pp. 3–16). Springer. https://doi.org/10.1007/978-1-4614-1539-8_1

Hartley, J. R., & Sleeman, D. H. (1973). Towards more intelligent teaching systems. *International Journal of Man-Machine Studies, 5*(2), 215–236. https://doi.org/10.1016/S0020-7373(73)80033-1

Hauptman, H., & Cohen, A. (2011). The synergetic effect of learning styles on the interaction between virtual environments and the enhancement of spatial thinking. *Computers & Education, 57*(3), 2106–2117. https://doi.org/10.1016/j.compedu.2011.05.008

Hawk, T. F., & Shah, A. J. (2007). Using learning style instruments to enhance student learning. *Journal of Innovative Education, 5*(1), 1–19. https://doi.org/10.1111/j.1540-4609.2007.00125.x

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics, 6*(2), 107–128. https://doi.org/10.3102/10769986006002107

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin, 92*, 490–499.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, NY: Academic Press.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65. https://doi.org/10.1002/jrsm.5

Honey, P., & Mumford, A. (1992). *The Manual of Learning Styles*. Maidenhead: Peter Honey.

Hong, H., and Kinshuk. (2004). Adaptation to student learning styles in Web based educational systems [Paper presentation]. *World Conference on Educational Multimedia, Hypermedia & Telecommunications*. Lugano, Switzerland

Hwang, G. J., Sung, H. Y., Chang, S. C., & Huang, X. C. (2020). A fuzzy expert system-based adaptive learning approach to improving students' learning performances by considering affective and cognitive factors. *Computers and Education: Artificial Intelligence, 1*, 100003. https://doi.org/10.1016/j.caeai.2020.100003

Hwang, G. J., Sung, H. Y., Hung, C. M., & Huang, I. (2013). A learning style perspective to investigate the necessity of developing adaptive learning systems. *Journal of Educational Technology & Society, 16*(2), 188–197.

Imhof, C., Bergamin, P., & McGarrity, S. (2020). Implementation of adaptive learning systems: Current state and potential. *Online teaching and learning in higher education*, 93–115. https://doi.org/10.1007/978-3-030-48190-2_6

Jagadeesan, S., & Subbiah, J. (2020). Real-time personalization and recommendation in Adaptive Learning Management System. *Journal of Ambient Intelligence and Humanized Computing, 11*(11), 4731–4741. https://doi.org/10.1007/s12652-020-01729-1

Jando, E., Hidayanto, A. N., Prabowo, H., & Warnars, H. L. H. S. (2017). Personalized e-learning model: A systematic literature review [Paper presentation]. *International Conference on Information Management and Technology (ICIMTech)*. Yogyakarta, Indonesia.

Jenkins, J. M., & Keefe, J. W. (2002). A special section on personalized instruction two schools: Two approaches to personalized learning. *Phi Delta Kappan, 83*, 449–456.

Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences, learning, and instruction*. Hillsdale, NJ: Erlbaum.

Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence, 2*, 1–12. https://doi.org/10.1016/j.caeai.2021.100017

Kasim, N. N. M., & Khalid, F. (2016). Choosing the right learning management system (LMS) for the higher education institution context: A systematic review. *International Journal of Emerging Technologies in Learning, 11*(6), 55–61.

Keefe, J. W. (1991). *Learning style: Cognitive and thinking skills*. Reston, VA: National Association of Secondary School Principals.

Kepes, S., Banks, G. C., McDaniel, M., & Whetzel, D. L. (2012). Publication bias in the organizational sciences. *Organizational Research Methods, 15*(4), 624–662. https://doi.org/10.1177/1094428112452760

Kerr, P. (2016). Adaptive learning. *ELT Journal, 70*(1), 88–93. https://doi.org/10.1093/elt/ccv055

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science, 36*, 757–798. https://doi.org/10.1111/j.1551-6709.2012.01245.x

Kulik, J., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research, 86*(1), 42–78. https://doi.org/10.3102/0034654315581420

Kumar, A., Singh, N., & Ahuja, N. J. (2017). Learning styles based adaptive intelligent tutoring systems: Document analysis of articles published between 2001 and 2016. *International Journal of Cognitive Research in Science, Engineering and Education, 5*(2), 83–97. https://doi.org/10.5937/IJCRSEE1702083K

Lee, D., Huh, Y., Lin, C.-Y., & Reigeluth, C. M. (2018). Technology functions for personalized learning in learner-centered schools. *Educational Technology Research and Development, 66*, 1269–1302. https://doi.org/10.1007/s11423-018-9615-9

Leutner, D. (2004). Instructional design principles for adaptivity in open learning environments. In N. M. Seel & S. Dijkstra (Eds.), *Curriculum, plans and processes of instructional design: international perspectives* (pp. 289–307). Mahwah, NJ: Lawrence Erlbaum.

Li, F., He, Y., & Xue, Q. (2021). Progress, challenges and countermeasures of adaptive learning. *Educational Technology & Society, 24*(3), 238–255.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Liu, M., Kang, J., Zou, W. T., Lee, H., Pan, Z. L., & Corliss, S. (2017). Using data to understand how to better design adaptive learning. *Technology, Knowledge and Learning, 22*(3), 271–298. https://doi.org/10.1007/s10758-017-9326-z

Lowendahl, J. M., Thayer, T. L. B., Morgan, G., Yanckello, R., Resnick, M., & Revang, M. (2016). Top 10 strategic technologies impacting higher education in 2016. *Research Note G, 294732*, 15.

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology, 106*(4), 901–918. http://dx.doi.org/10.1037/a0037123

Ma, Y., Chu, H., & Mazumdar, M. (2016). Meta-analysis of proportions of rare events–a comparison of exact likelihood methods with robust variance estimation. *Communications in Statistics-Simulation and Computation, 45*(8), 3036–3052. https://doi.org/10.1080/03610918.2014.911901

Malekzadeh, M., Mustafa, M. B., & Lahsasna, A. (2015). A review of emotion regulation in intelligent tutoring systems. *Educational Technology & Society, 18*(4), 435–445.

Mampadi, F., Chen, S., Ghinea, G., & Chen, M. (2011). Design of adaptive hypermedia learning systems: A cognitive style approach. *Computers & Education, 56*(4), 1003–1011. https://doi.org/10.1016/j.compedu.2010.11.018

Martins, A. C., Faria, L., Vaz de Carvalho, C., & Carrapatoso, E. (2008). User modeling in adaptive hypermedia educational systems. *Educational Technology & Society, 11*(1), 194–207.

Martin, F., Chen, Y., Moore, R. L., & Westine, C. D. (2020). Systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018. *Educational Technology Research and Development, 68*, 1903–1929. https://doi.org/10.1007/s11423-020-09793-2

Mavroudi, A., Giannakos, M., & Krogstie, J. (2018). Supporting adaptive learning pathways through the use of learning analytics: Developments, challenges and future opportunities. *Interactive Learning Environments, 26*(2), 206–220. https://doi.org/10.1080/10494820.2017.1292531

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282.

Messick, S. (1984). The nature of cognitive styles: problems and promise. *Educational Psychologist, 19*(2). 59–74. https://doi.org/10.1080/00461528409529283

Mödritscher, F., Garcia-Barrios, V. M., & Gütl, C. (2004). The past, the present and the future of adaptive e-learning. An approach within the scope of the research project AdeLE. In M. Auer & U. Auer (Eds.), *Proceedings of the International Conference on Interactive Computer Aided Learning (ICL 2004)*. Villach, Austria: Carinthia Tech Institute.

Mohamed, H., Bensebaa, T., & Trigano, P. (2012). Developing adaptive intelligent tutoring system based on item response theory and metrics. *International Journal of Advanced Science and Technology*, *43*, 1–14.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine, 151*(4), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Mohd, C. K. N. C. K., & Shahbodin, F. S. (2015). Personalized learning environment (PLE) experience in the twenty-first century: Review of the literature. In A. Abraham, A. Muda,

& Y. H. Choo (Eds.), *Pattern Analysis, Intelligent Security and the Internet of Things* (pp. 179–192). Springer, Cham. https://doi.org/10.1007/978-3-319-17398-6_17

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364–386. https://doi.org/10.1177/1094428106291059

Mousavinasab, E., Zarifsanaiey, N., Kalhori, S. R. N., Rakhshan, M., Keikha, L., & Saeedi, M. G. (2021). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments, 29*(1), 142–163. https://doi.org/10.1080/10494820.2018.1558257

Murray, T. (1999). Authoring intelligent tutoring systems: An analysis of the state of the art. *International Journal of Artificial Intelligence in Education (IJAIED)*, *10*, 98–129.

Muñoz, J. L. R., Ojeda, F. M., Jurado, D. L. A., Peña, P. F. P., Carranza, C. P. M., Berríos, H. Q., ... & Vasquez-Pauca, M. J. (2022). Systematic Review of Adaptive Learning Technology for Learning in Higher Education. *Eurasian Journal of Educational Research, 98*(98), 221–233. https://doi.org/10.14689/ejer.2022.98.014

Mustafa, Y. E. A., & Sharif, S. M. (2011). An approach to adaptive e-learning hypermedia system based on learning styles (AEHS-LS): Implementation and evaluation. *International Journal of Library and Information Science, 3*(1), 15–28.

Nakic, J., Granic, A., & Glavinic, V. (2015). Anatomy of student models in adaptive learning systems: A systematic literature review of individual differences from 2001 to 2013. *Journal of Educational Computing Research, 51*(4), 459–489. https://doi.org/10.2190/EC.51.4.e

National Association of State Boards of Education. (2001). *Any time, any place, any path, any pace: Taking the lead on e-learning policy*. Alexandria, VA: Author. Retrieved from http://www.nasbe.org/e-Learning.html

Nesbit, J. C., Adesope, O. O., Liu, Q., & Ma, W. (2014). How effective are intelligent tutoring systems in computer science education? *2014 IEEE International Conference on Advanced Learning Technologies*. https://doi.org/ 10.1109/ICALT.2014.38

Nguyen, L., & Do, P. (2008). Learner model in adaptive learning. *World Academy of Science, Engineering and Technology, 45*(70), 395–400.

Njenga, J. K., & Fourie, L. C. H. (2010). The myths about e-learning in higher education. *British Journal of Educational Technology, 41*(2), 199–212. https://doi.org/10.1111/j.1467-8535.2008.00910.x

Normadhi, N. B. A., Shuib, L., Nasir, H. N. M., Bimba, A., Idris, N., & Balakrishnan, V. (2019). Identification of personal traits in adaptive learning environment: Systematic literature review. *Computers & Education, 130*, 168–190. https://doi.org/10.1016/j.compedu.2018.11.005

Nwana, H. S. (1990). Intelligent tutoring systems: An overview. *Artificial Intelligence Review*, *4*(4), 251–277.

Nye, B. D. (2015). Intelligent tutoring systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education, 25*, 177–203. https://doi.org/10.1007/s40593-014-0028-6

Ohlsson, S. (1996). Learning from performance errors. *Psychological review, 103*(2), 241–262. https://doi.org/10.1037/0033-295X.103.2.241

Orwin, R. F. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8*, 157–159.

Ounaies, H. Z., Jamoussi, Y., & Ghezala, H. H. B. (2012). Multi-perspective measurement framework for adaptive learning system. *International Journal of Software Engineering and Its Applications, 6*(2), 23–34.

Özyurt, Ö., & Özyurt, H. (2015). Learning style based individualized adaptive e-learning environments: Content analysis of the articles published from 2005 to 2014. *Computers in Human Behavior, 52*, 349–358. https://doi.org/10.1016/j.chb.2015.06.020

Papanagnou, D., Serrano, A., Barkley, K., Chandra, S., Governatori, N., Piela, N., … Shin, R. (2016). Does tailoring instructional style to a medical student's self-perceived learning style improve performance when teaching intravenous catheter placement? A randomized controlled study. *BMC Medical Education, 16*, 205. https://doi.org/10.1186/s12909-016-0720-3

Paredes, P., & Rodriguez, P. (2004). A mixed approach to modelling learning styles in adaptive educational hypermedia. *Advanced Technology for Learning*, *1*(4), 210–215.

Pardos, Z. A., & Heffernan, N. T. (2010, June). Modeling individualization in a bayesian networks implementation of knowledge tracing. In *International conference on user modeling, adaptation, and personalization* (pp. 255–266). Springer, Berlin, Heidelberg.

Park, J. H., & Choi, H. J. (2009). Factors influencing adult learners' decision to drop out or persist in online learning. *Journal of Educational Technology & Society, 12*(4), 207–217.

Pask, G. (1976). Styles and strategies of learning. *British Journal of Educational Psychology, 46*(2), 128–148. https://doi.org/10.1111/j.2044-8279.1976.tb02305.x

Phobun, P., & Vicheanpanya, J. (2010). Adaptive intelligent tutoring systems for e-learning systems. *Procedia – Social and Behavioral Sciences, 2*(2), 4064–4069. https://doi.org/10.1016/j.sbspro.2010.03.641

Pliakos, K., Joo, S.-H., Park, J. Y., Cornillie, F., Vens, C., & Noortgate, W. (2019). Integrating machine learning into item response theory for addressing the cold start problem in adaptive learning systems. *Computers & Education, 137*, 91–103. https://doi.org/10.1016/j.compedu.2019.04.009

Popescu, E. (2008). *Dynamic adaptive hypermedia systems for e-learning* [Unpublished doctoral dissertation]. Université de Craiova.

Porta, M. S. (2014). *A dictionary of epidemiology*. Oxford University Press.

Premlatha, K. R., & Geetha, T. V. (2015). Learning content design and learner adaptation for adaptive e-learning environment: A survey. *Artificial Intelligence Review, 44*, 443–465. https://doi.org/10.1007/s10462-015-9432-z

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science, 23*(3), 425–438. https://doi.org/10.1007/s11121-021-01246-3

Qazdar, A., Cherkaoui, C., Er-Raha, B., & Mammass, D. (2015). AeLF: mixing adaptive learning system with learning management system. *International Journal of Computer Applications, 119*(15), 1–8.

R Core Team. (2022). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Rasheed, F., & Wahid, A. (2021). Learning style detection in E-learning systems using machine learning techniques. *Expert Systems with Applications, 174*, 114774. https://doi.org/10.1016/j.eswa.2021.114774

Redding, S. (2016). Competencies and personalized learning. In M. Murphy, S. Redding, & J. Twyman (Eds.), *Handbook on personalized learning for states, districts, and schools* (pp. 3–18). Philadelphia, PA.

Riding, R., & Rayner, S. (1998). *Cognitive Styles and Learning Strategies: Understanding Style Differences in Learning and Behaviour*. David Fulton, London.

Riding, R. J., & Buckle, C. F. (1990). Learning styles and training performance. *Sheffield: Training Agency*.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. Psychological *Bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Rosita Cecilia, M., Vittorini, P., & di Orio, F. (2016). An adaptive learning system for developing and improving reading comprehension skills. *Journal of Education Research, 10*(4). 195–236

Rothstein, H. R. (2008). Publication bias as a threat to the validity of meta-analytic results. *Journal of Experimental Criminology, 4*(1), 61–81. https://doi.org/10.1007/s11292-007-9046-9

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464

Shemshack, A., Kinshuk, & Spector, J. M. (2021). A comprehensive analysis of personalized learning components. *Journal of Computers in Education, 8*, 485–503. https://doi.org/10.1007/s40692-021-00188-7

Shemshack, A., & Spector, J. M. (2020). A systematic literature review of personalized learning terms. *Smart Learning Environment, 7*(33), 1–20. https://doi.org/10.1186/s40561-020-00140-9

Shute, V. J. (1993). A comparison of learning environments: All that glitters. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as Cognitive Tools* (pp. 4774). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Shute, V. J., & Towle, B. (2003). Adaptive e-learning. *Educational Psychologist, 38*(2), 105–114. https://doi.org/10.1207/S15326985EP3802_5

Shute, V. J. & Zapata-Rivera, D. (2008). Using an evidence-based approach to assess mental models. In D. Ifenthaler, P. Pirnay-Dummer, & J. M. Spector (Eds.), *Understanding models for learning and instruction: Essays in honor of Norbert M. Seel* (pp. 23–41). Springer.

Siddique, A., Durrani, Q. S., & Naqvi, H. A. (2019). Developing adaptive e-learning environment using cognitive and noncognitive parameters. *Journal of Educational Computing Research, 57*(4), 811–845. https://doi.org/10.1177/0735633118769433

Sugawara, R., Okuhara, S., & Sato, Y. (2020). Study about the aptitude-treatment interaction between learning using the e-learning system and learning type of learner. International *Journal of Information and Education Technology, 10*(7), 488–493. https://doi.org/10.18178/ijiet.2020.10.7.1412

Sunitha, R., Shyamala, R., & Aghila, G. (2011). Towards learner model sharing among heterogeneous e-learning environments. *International Journal on Computer Science and Engineering, 3*(5), 2034–2040.

Surjono, H. D. (2011). The design of adaptive e-learning system based on students' learning styles. *International Journal of Computer Science and Information Technologies, 2*(5), 2350–2353.

Stash, N., & De Bra, P. (2004). Incorporating cognitive styles in AHA! (The adaptive hypermedia architecture) [Paper presentation]. The World Wide Web Conference 2004, New York, USA.

Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K-12 students' mathematical learning. *Journal of Educational Psychology, 105*(4), 970–987. https://doi.org/10.1037/a0032447

Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology, 106*(2), 331–347. https://doi.org/10.1037/a0034752

Stemler, S., & Tsai, J. (2008). Best Practices in Interrater Reliability Three Common Approaches. In J. Osborne (Ed.), *Best practices in quantitative methods* (pp. 29–50). Thousand Oaks, CA: SAGE Publications, Inc.

Sterne, J. A., & Harbord, R. M. (2004). Funnel plots in meta-analysis. *The stata journal, 4*(2), 127–141. https://doi.org/10.1177/1536867X0400400204

Tacoma, S., Drijvers, P., & Jeuring, J. (2021). Combined inner and outer loop feedback in an intelligent tutoring system for statistics in higher education. *Journal of Computer Assisted Learning, 37*(2), 319–332. https://doi.org/10.1111/jcal.12491

Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research, 81*(1), 4–28. https://doi.org/10.3102/0034654310393361

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analysis data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology, 2*, 85–112. https://doi.org/10.1007/s40865-016-0026-5

The Endnote Team. (2013). *Endnote (Version Endnote X9) (64 bit)*. Philadelphia: Clarivate.

Thomson, D., & Mitrovic, A. (2009). Towards a negotiable student model for constraint-based ITSs [Paper presentation]. 17th International Conference on Computers in Education, Hong Kong, China.

Triantafillou, E., Pomportsis, A., Demetriadis, S., & Georgiadou, E. (2004). The value of adaptivity based on cognitive style: an empirical study. *British Journal of Educational Technology, 35*(1), 95–106. https://doi.org/10.1111/j.1467-8535.2004.00371.x

Truong, H. M. (2016). Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities. *Computers in Human Behavior, 55*, 1185–1193. https://doi.org/10.1016/j.chb.2015.02.014

Tseng, J. C., Chu, H. C., Hwang, G. J., & Tsai, C. C. (2008). Development of an adaptive learning system with two sources of personalization information. *Computers & Education, 51*(2), 776–786. https://doi.org/10.1016/j.compedu.2007.08.002

Uruchrutu, E., Mackinnon, L., & Rist, R. (2005). User cognitive style and interface design for personal, adaptive learning: What to model? [Paper presentation]. The International Conference on User Modeling UM 2005, Edinburgh, Scotland.

U.S. Department of Education. (2012). *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief*. Washington, DC: U.S. Department of Education, Office of Education Technology.

U.S. Department of Education Office of Educational Technology. (2010). *Transforming American Education: Learning Powered by Technology.* Alexandria, VA: U.S. Department of Education.

Vandewaetere, M., & Clarebout, G. (2014). Advanced technologies for personalized learning, instruction, and performance. In J. Spector, M. Merrill, J. Elen, & M. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 425–437). Springer, New York, NY.

Vandewaetere, M., Desmet, P., & Clarebout, G. (2011). The contribution of learner characteristics in the development of computer-based adaptive learning environments. *Computers in Human Behavior, 27*(1), 118–130. https://doi.org/10.1016/j.chb.2010.07.038

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education, 16*, 227–265.

Verdú, E., Regueras, L. M., Verdú, M., J., De Castro, J. P., & Perez, M. A. (2008). Is adaptive learning effective? A review of the research [Paper presentation]. The 7th WSEAS

International Conference on Applied Computer & Applied Computers Science. Hangzhou, China.

Vevea J. L., Coburn K., Sutton A. (2019). Publication bias. In Cooper H., Hedges L.V., Valentine J.C. (Eds.), *The Handbook of Research Synthesis and Meta-Analysis*, 3rd ed (pp. 383–429). New York: Russell Sage.

Walkington, C., & Bernacki, M. L. (2020). Appraising research on personalized learning: Definitions, theoretical alignment, advancements, and future directions. *Journal of Research on Technology in Education, 52*(3), 235–252. https://doi.org/10.1080/15391523.2020.1747757

Wang, Y.-Q., Liu, H., & You, X. (2013). Learning diagnosis—from concepts to system development. Paper presented at the Annual Meeting of Assessment and Evaluation, the Chinese Society of Education, Dalian, China, May.

Weiss, D. J. (1982). Improving measurement quality and efficiency with adaptive testing. Applied psychological measurement, 6(4), 473–492. https://doi.org/10.1177/014662168200600408

What Works Clearinghouse. (2013). *Procedures and standards handbook* (Version 3.0). Retrieved from http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the econometric society*, *50*(1), 1–25.

Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. (1977). Field-dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research, 47*(1), 1–64.

Xie, H., Chu, H. C., Hwang, G. J., & Wang, C. C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education, 140*, 103599. https://doi.org/10.1016/j.compedu.2019.103599

Xu, Z., Wijekumar, K., Ramirez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K-12 students' reading comprehension: A meta-analysis. *British*

Journal of Educational Technology, 50(6), 3119–3137. https://doi.org/10.1111/bjet.12758

Yang, T.-C., Hwang, G.-J., & Yang, S. J. H. (2013). Development of an adaptive learning system with multiple perspectives based on students' learning styles and cognitive styles. *Journal of Educational Technology & Society, 16*(4), 185–200.

Zhang, L., Basham, J. D., & Yang, S. (2020). Understanding the implementation of personalized learning: A research synthesis. *Educational Research Review, 31*, 100339. https://doi.org/10.1016/j.edurev.2020.100339

Zhang, L., Pan, M., Yu, S., Chen, L., & Zhang, J. (2021). Evaluation of a student-centered online one-to-one tutoring system. Interactive Learning Environments, 1–19. https://doi.org/10.1080/10494820.2021.1958234

Zhang, S., & Chang, H. H. (2016). From smart testing to smart learning: How testing technology can assist the new generation of education. *International Journal of Smart Technology and Learning, 1*(1), 67–92.

Appendix A

The Titles and Publication Years of the Existing Literature Reviews on ALSs

Title	Publication Year
Adaptive e-learning	2003
User modeling and user profiling in adaptive e-learning systems	2005
Adaptive educational hypermedia systems in technology enhanced learning: A literature review	2010
The contribution of learner characteristics in the development of computer-based adaptive learning environments	2011
Adaptive systems: A content analysis on technical side for e-learning environments	2015
Learning style based individualized adaptive e-learning environments: Content analysis of articles published from 2005 to 2014	2015
Learning content design and learner adaptation for adaptive e-learning environment: A survey	2015
Application of data mining in adaptive and intelligent tutoring systems: A review	2015
Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities	2016
Adaptivity in educational systems for language learning: A review	2017
Learning styles based adaptive intelligent tutoring systems: Document analysis of articles published between 2001 and 2016	2017
Review of data mining techniques and parameters for recommendation of effective adaptive e-learning system	2017
Learner modelling: Systematic review of the literature from the last 5 years	2019
Identification of personal traits in adaptive learning environment: Systematic literature review	2019
A literature review of the adaptive algorithms adopted in adaptive learning systems	2019
Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017	2019
A systematic review of process modelling methods and its application for personalised adaptive learning systems	2019
Design framework of adaptive intelligent tutoring systems	2020
A systematic review: Machine learning based recommendation system for e-learning	2020
Learning path personalization and recommendation methods: A survey of the state-of-the-art	2020
Enhancing students' ability in learning process of programming language using adaptive learning systems: A literature review	2020

A systematic literature review on adaptive content recommenders in personalized learning environments from 2015 to 2020	2021
AI-enabled adaptive learning systems: A systematic mapping of the literature	2021
A review of adaptive and intelligent online learning systems	2022
Systematic review of adaptive learning technology for learning in higher education	2022
Personalized and adaptive context-aware mobile learning: Review, challenges and future directions	2022

Appendix B

An Overview of Studies on ALSs Selected in the Present Meta-Analysis

No.	Author	Title
1	Adnan et al. (2019)	Cloud-supported machine learning system for context-aware adaptive m-learning
2	Alsadoon (2020)	The impact of an adaptive e-course on students' achievements based on the students' prior knowledge
3	Alshalabi et al. (2018)	An automated adaptive mobile learning system using optimal shortest path algorithms
4	Alshammari et al. (2015)	An e-learning investigation into learning style adaptivity
5	Alshammari et al. (2016)	Usability and effectiveness evaluation of adaptivity in e-learning systems
6	Cabada et al. (2018)	An affective and Web 3.0-based learning environment for a programming language
7	Chang et al. (2016)	Yet another adaptive learning management system based on Felder and Silverman's learning style and Mashup
8	Chen (2014)	An adaptive scaffolding e-learning system for middle school students' physics learning
9	Chu et al. (2021)	Development of an adaptive game-based diagnostic and remedial learning system based on the concept-effect model for improving learning achievements in mathematics
10	Dhakshinamoorthy et al. (2018)	KLSAS—An adaptive dynamic learning environment based on knowledge level and learning style
11	EI-Ghool et al. (2016)	Designing an adaptive learning environment to improve writing skills and usability for EFL students at the faculty of education
12	Hariyanto (2020)	An adaptive e-learning system based on student's learning styles and knowledge level
13	Hsieh et al. (2012)	A fuzzy logic-based personalized learning system for supporting adaptive English learning

14	Huang et al. (2012)	A user-centric adaptive learning system for e-learning 2.0
15	Hwang et al. (2013)	A learning style perspective to investigate the necessity of developing adaptive learning systems
16	Jeong (2016)	UX based adaptive e-learning hypermedia system (U-AEHS): An integrative user model approach
17	Jong et al. (2012)	Applying learning portfolios and thinking styles to adaptive remedial learning
18	Joseph et al. (2019)	Adaptive e-learning system for slow learners based on Felder-Silverman learning style model
19	Kavcic (2004)	Fuzzy user modeling for adaptation in educational hypermedia
20	Lamia et al. (2017)	A model for an adaptive hypermedia learning system based on data mining technique
21	Lin et al. (2014)	The influence of using affective tutoring system in accounting remedial instruction on learning performance and usability
22	Lo et al. (2004)	Effects of confidence scores and remedial instruction on prepositions learning in adaptive hypermedia
23	Mampadi et al. (2011)	Design of adaptive hypermedia learning systems: A cognitive style approach
24	Mazaheri et al. (2014)	An adaptive e-learning system based on Gardner's eight intelligence theory
25	Mohamed et al. (2012)	Developing adaptive intelligent tutoring system based on item response theory and metrics
26	Naik et al. (2015)	Adaptive and Gamified Learning Environment (AGLE)
27	Niknam (2017)	LPR: An adaptive learning path recommendation system using ACO and meaningful learning theory
28	Pandey et al. (2014)	CBR based approach for adaptive learning in e-learning system
29	Siadaty et al. (2007)	PALS2: Pedagogically adaptive learning system based on learning styles

30	Siddique et al. (2019)	Developing adaptive e-learning environment using cognitive and noncognitive parameters
31	Srisawasdi et al. (2015)	Personal learning activity approach for developing adaptive web-based learning systems
32	Su (2014)	A self-regulated learning system to support adaptive scaffolding in hypermedia-based learning environments
33	Su (2017)	Designing and developing a novel hybrid adaptive learning path recommendation system (ALPRS) for gamification mathematics geometry course
34	Suraweera et al. (2004)	An intelligent tutoring system for entity relationship modelling
35	Tseng, Chu, et al. (2008)	Development of an adaptive learning system with two sources of personalization information
36	Tseng, Su, et al. (2008)	An object-oriented course framework for developing adaptive learning systems
37	Wan and Yu (2020)	A recommendation system based on an adaptive learning cognitive map model and its effects
38	Wang (2014)	Developing and evaluating an adaptive business English self-learning system for EFL vocabulary learning
39	Wang (2016)	Promoting contextual vocabulary learning through an adaptive computer-assisted EFL reading system
40	Wang et al. (2011a)	Data mining for adaptive learning in a TESL-based e-learning system
41	Wang et al. (2011b)	Adaptive learning for ESL based on computation
42	Wang et al. (2011c)	Application of context-aware and personalized recommendation to implement an adaptive ubiquitous learning system
43	Wang et al. (2020)	When adaptive learning is effective learning: comparison of an adaptive learning system to teacher-led instruction
44	Wu et al. (2017)	An adaptive e-learning system for enhancing learning performance: Based on dynamic scaffolding theory

45	Yu and Xu (2020)	The practice strategy for adaptive Chinese grammar learning system
46	Zafar et al. (2014)	Evaluation study of eLGuide: A framework for adaptive e-learning