Optional ERIC Coversheet — Only for Use with U.S. Department of Education Grantee Submissions

This coversheet should be completed by grantees and added to the PDF of your submission if the information required in this form **is not included on the PDF to be submitted**.

GRANTEE SUBMISSION REQUIRED FIELDS

Title of article, paper, or other content

The Future of Intelligent Tutoring Systems for Writing

All author name(s) and affiliations on PDF. If more than 6 names, ERIC will complete the list from the submitted PDF.

Last Name, First Name	Academic/Organizational Affiliation	ORCID ID
Banawan, Michelle	Asian Institute of Management, Makati City, Pl	
Butterfuss, Reese	Minneapolis, MN, USA	
Taylor, Karen S.	Phoenix, AZ, USA	
Christhilf, Katerina	Tempe, AZ, USA	
Hsu, Claire	Paradise Valley, AZ, USA	
O'Loughlin, Connor	Phoenix, AZ, USA	

Publication/Completion Date—(if *In Press,* enter year accepted or completed) September 2023

Check type of content being submitted and complete one of the following in the box below:

- ☐ If article: Name of journal, volume, and issue number if available
- ☐ If paper: Name of conference, date of conference, and place of conference
- ☑ If book chapter: Title of book, page range, publisher name and location
- ☐ If book: Publisher name and location
- ☐ If dissertation: Name of institution, type of degree, and department granting degree

Digital Writing Technologies in Higher Education

DOI or URL to published work (if available) https://doi.org/10.1007/978-3-031-36033-6_23

Acknowledgement of Funding— Grantees should check with their grant officer for the preferred wording to acknowledge funding. If the grant officer does not have a preference, grantees can use this suggested wording (adjust wording if multiple grants are to be acknowledged). Fill in Department of Education funding office, grant number, and name of grant recipient institution or organization.

"This work was supported by U.S. Department of Education **[Office name]** National Center for Educat through **[Grant number]** R305A1801 to **Institution]** Arizona State U.The opinions expressed are those of the authors and do not represent views of the **[Office name]** National Center for Educa or the U.S. Department of Education.

The Future of Intelligent Tutoring Systems for Writing

Michelle Banawan, Reese Butterfuss, Karen S. Taylor, Katerina Christhilf, Claire Hsu, Connor O'Loughlin, Laura K. Allen, Rod D. Roscoe, and Danielle S. McNamara

Abstract Writing is essential for success in academics and everyday tasks, but the development of writing skills depends on consistent access to high-quality instruction, extended practice, and personalized feedback. To address these demands and meet students' needs, educators and researchers have turned to technology-based writing tools. Ideally, these tools integrate the core components of intelligent tutoring, including a domain model, student model, tutor model, and interface model to engage students with individualized feedback that is linked to adaptive writing instruction. However, the landscape of writing tools still has much room for improvement in terms of incorporating advanced artificial intelligence-enabled features to better approximate intelligent tutoring systems (ITSs). This chapter describes the key elements of ITS technologies and how they can be integrated to further develop ITS tools for writing. To this end, this chapter (1) summarizes evidence-based aspects of successful ITSs and how they might be integrated into computer-based tools for writing, (2) reviews how existing systems have leveraged intelligent tutoring approaches, and (3) articulates how future technology-based writing tools could implement advanced

M. Banawan
Asian Institute of Management, Makati City, Philippines

R. Butterfuss · L. K. Allen
Minneapolis, MN, USA
e-mail: lallen@umn.edu

K. S. Taylor · C. O'Loughlin
Phoenix, AZ, USA

K. Christhilf
Tempe, AZ, USA

C. Hsu
Paradise Valley, AZ, USA

R. D. Roscoe
Mesa, MN, USA
e-mail: Rod.Roscoe@asu.edu

D. S. McNamara (✉)
Arizona State University, Tempe, AZ, USA
e-mail: Danielle.McNamara@asu.edu

O. Kruse et al. (eds.), *Digital Writing Technologies in Higher Education*,
https://doi.org/10.1007/978-3-031-36033-6_23

intelligent tutoring features to better meet students' needs. The chapter concludes with the implications and future directions of intelligent tutoring for the teaching and learning of writing.

Keywords Intelligent tutoring systems for writing · Architecture of intelligent tutoring systems · Writing Pal

1 Overview

Strong writing skills are essential to academic performance across nearly all domains as well as for success in everyday life (Powell, 2009). However, writing is challenging because of the demands it places on cognitive skills and knowledge (Deane et al., 2008). Improving students' writing skills requires enormous amounts of high-quality instruction, deliberate practice, and individualized formative and summative feedback. Curricula developed to provide students with these resources can be challenging to implement in the classroom due to the time required for educators to read and provide individualized feedback on students' writing. Thus, educators have turned to intelligent writing tools as a means of supplementing classroom instruction and increasing students' opportunities to engage in deliberate writing practice. Most widely used for these purposes are automated essay scoring (AES) systems that provide valid and reliable scores and feedback-generating automated writing evaluation (AWE) systems (Cotos, 2018) (see Chapters S3C5, S3C6).

Many of the components of AES and AWE systems have also been integrated into prototypes of intelligent tutoring systems (ITSs), incorporating instructional content, game-based practice, and iterative practice with feedback into their architectures. Thus, the development of digital tools for writing has followed a trajectory from a focus on scoring to a focus on feedback and instruction, thereby becoming more ITS-like over time. However, the landscape of writing tools has much room for the design and development of cutting-edge systems for writing by implementing both traditional ITS elements and advances in artificial intelligence (AI), natural language processing (NLP) , and human–computer interaction (HCI). Compared to well-defined domains (e.g., algebra) for which ITSs have traditionally been developed, there are unique challenges in developing ITSs for writing. This chapter reviews several existing writing tools using ITS architecture as the analytic frame to identify challenges and forecast how intelligent tutoring for writing could be successfully implemented. In doing so, our goal is to capture the current state of the art in digital writing tools and stimulate future research regarding ITSs for writing.

ITSs are automated learning platforms that simulate tutor-tutee interaction while providing detailed feedback, assessments, and personalized learning, often through content adaptation that leverages the tutees' strengths and addresses their specific needs. ITS implementations emulate the known benefits of tutoring (Bloom, 1984) while simultaneously addressing limitations such as tutor subjectivity, fatigue, cost, and limited resources. ITSs employ a variety of pedagogical tools to support desired learning outcomes in a specific domain without intervention from human tutors or experts (Graesser et al., 2012; Ma et al., 2014).

2 Core Idea of the Technology: Architecture of Intelligent Tutoring Systems

There are four components of contemporary ITSs: *domain model*, *student model*, *tutor model*, and *interface model*. Earlier architectures encompassed the first three components (Derry et al., 1988). These core components provided the ITS with critical information on what to teach, who to teach, and how to teach (Nwana, 1990). The three-model architecture later expanded to include a fourth component, the user interface, and the four-model architecture has become the standard architecture for ITSs (Almasri et al., 2019).

2.1 Domain Model

The domain model represents the idealized expert knowledge domain, which may include the concepts, rules, skills, and strategies of the topics to be learned (Sottilare et al., 2016). It thus serves as the standard for evaluating students' performance and as the reference used to detect errors or deviations from expected knowledge and skills. This component is often organized into a curriculum that links all knowledge elements according to a pedagogical sequence. Domain models frequently implement a sequenced curriculum such that new material builds on prior knowledge and aspects of the curriculum that were previously administered.

2.2 Student Model

The student model focuses on students' cognitive and meta-cognitive states throughout the learning process. It represents what the students learn and how they learn, capturing the processes and strategies by which they learn. This component maps to the domain model, wherein students' knowledge is measured in terms of ideal expert knowledge (Sottilare et al., 2013). In other words, the student model captures the deviations from the expert knowledge base (represented by the curriculum) by highlighting gaps in students' knowledge. Therefore, it reflects the set of skills that students have mastered, thereby affording customized and individualized learning paths, feedback, and support. ITSs that recommend appropriate content or specific learning pathways based on students' progress, assessment results, or behaviors while using the system are usually informed by dynamic and adaptive student models.

2.3 Tutor Model

The tutor model, also known as the pedagogical model, teaching model, or expert model, relies on the interplay between the domain and student models to provide pedagogical strategies and actions that are most appropriate for a given student (e.g., providing a hint in response to an incorrect answer or assigning specific problems that target the skills that the student needs to improve upon). Additional tasks for this model include adjusting the speed of tutoring actions, checking the learning progress through questions, providing feedback, and offering additional information to assist with gaps in students' knowledge (Almasri et al., 2019). Knowledge tracing, or tracking students' progress while building a profile of strengths and weaknesses (Ahuja & Sille, 2013), is another important ability of the tutor model.

2.4 Interface Model

The user interface model, also referred to as the communication model, comprises the human–computer interaction features that are necessary to interpret and facilitate the learning process. The interface model provides the presentation of the learning material to the student and controls the communication and interaction between the student and the system. This component allows dialogue between students and the ITS to simulate tutor-tutee interaction. Intelligent interfaces focus on adaptive or adaptable interfaces to enhance user experience and learning (Sarrafzadeh et al., 2008). ITS interface models are implemented as pedagogical agents, menu-driven interfaces, text-driven interfaces, speech-driven interfaces, or via worked examples that demonstrate the steps necessary to complete a learning task. In addition, modern ITSs immerse students in a graphic environment enhanced by AI and virtual reality by employing animated and empathic pedagogical agents.

3 Functional Specifications: ITS Components in Action

ITSs implement the aforementioned components in different ways, but it is the integration of these components working together that greatly influences the effectiveness of intelligent tutoring. Considering that ITS for writing is still an underdeveloped area, in this section, we provide examples of representative ITSs from the domains of math and science. These exemplars depict the dynamic interplay between the four ITS components and how each component informs another, which is important to consider in the design of future ITSs for writing.

The Practical Algebra Tutor (PAT) is a system that mimics the steps a student would take to solve a problem and solves the problem at the same time as the student (Koedinger et al., 1997). The student model tracks students' steps in solving the problem and compares those steps against the domain model to check for discrepancies. In turn, the tutor model provides appropriate feedback at specific steps via the user interface. PAT has a domain model for each type of problem, as well as a representation of common student misconceptions. If students exhibit misconceptions in the domain model, the system leverages the tutor model to offer feedback that guides students back to the correct path. Students who used this step-based tutoring system performed significantly better compared to students following a traditional approach in a real-world problems assessment (Akyuz, 2020; Corbett et al., 1997).

The Andes Physics Tutoring System is another ITS that provides homework problem-solving support to students learning physics. Andes' tutor model consists of a coached problem-solving environment and provides immediate feedback through dialogue capabilities integrated into the interface model that provide students with increasingly specific hints for problem-solving. Importantly, the tutor models' feedback encourages the students to find the solution and not rely on the feedback system to provide the solution. The student model tracks students' responses and automatically notes when answers are inconsistent with the domain model. Because Andes allows students to perform tasks in no particular order, the student model cannot rely on accomplished tasks as the basis for the students' level of knowledge or mastery. Instead, the student model combines information on problem-specific knowledge that the students are working on or have completed along with information on the domain-general assessments that all problems have (Gertner & VanLehn, 2000). Encouraging results were found by many studies that evaluated the effectiveness of Andes in terms of increasing the learning gains of students as they are provided homework problem-solving support (VanLehn et al., 2005). The success of Andes' immediate feedback and hint progression strategies continue to spur similar implementations in more recent ITSs (Sale & Muldner, 2019).

As a final example (of many other potential exemplars), AutoTutor is a problem-oriented ITS that presents interactive content and uses conversational agents to help students learn. AutoTutor's domain model contains lessons and problems that cover the content of specific domains like computer literacy, critical scientific thinking, physics, and reading. The problems that the students work on are mapped to the knowledge components comprising the lessons. AutoTutor's tutor model leverages natural language and text-to-speech features in dialogue. Its interface model implements animated conversational agents that have facial expressions and can make various gestures (Cai et al., 2019; Graesser et al., 2007). Different versions of AutoTutor's student model also capture student affective states in real-time and modify the instruction that the tutor model provides to enhance student engagement (D'Mello et al., 2007). Students' affective states are derived from the dialogue patterns and physical markers that include facial expressions and posture students exhibit when interacting with the interface.

4 Main Products: A Landscape of Intelligent Writing Tools

Existing digital tools for writing have leveraged one or more components of intelligent tutoring that are present in the paradigmatic math and science ITSs presented in the previous section. In this section, we provide a selective overview of several digital tools for writing, including AWE systems (see S3C6) and highlight each tool's most noteworthy ITS-like component (see Table 1). By outlining the landscape of intelligent tools for writing and articulating how these tools integrate various components of intelligent tutoring, we clarify how an ITS for writing might be further improved.

Table 1 Intelligent features of digital writing tools

	Student model	Tutor model	Domain model	Interface model
Criterion		Formative and summative feedback on different writing traits and customized based on grade levels and prompts	Library of expository and argumentative prompts	Various learning artifacts
Research Writing Tutor		Formative feedback on rhetorical conventions of scientific writing	Annotated corpus of published discipline-specific scientific writing	Learning, demonstration, and feedback modules
Sword/ Peerceptiv		Open-ended feedback and weighted scores based on system-calculated accuracy of peer reviews	Double-blind reviews by students across disciplines	Task-driven user interface with elements reflecting different steps of the writing process and task
HARRY		Conversation-based feedback on narrative writing at word, sentence, and idea levels	Story themes and tasks organized based on writing stages	Scaffolding specific to writing stages
Writing Pal	Dynamic student model representation based on practice and summative performance	Formative and summative feedback on writing strategies Coached practice Gamified practice	Corpora of essay prompts Flexible sequencing of content / instruction Various pedagogical strategies	Freewriting, Planning, Introduction Building, Body Building, Conclusion Building, Paraphrasing, Cohesion Building, and Revising modules

4.1 Criterion® Online Writing Evaluation Service

Criterion® Online Writing Evaluation Service was developed by the Educational Testing Service (ETS) (see Chapter S3C6). Criterion is a representative AWE tool that exhibits built-in intelligence in providing feedback. Criterion's domain model is comprised of a content library of 180 essay topics and over 400 expository and argumentative assignments and prompts designed for students from fourth grade through college. Criterion uses NLP techniques to score and provide feedback on students' writing. Along with a holistic score, students also receive feedback on language errors (e.g., grammatical errors) and discourse elements (e.g., the absence of a thesis statement). Though teachers may provide the assignment and give feedback, the system is meant to be fairly independent by giving specific, timely feedback. The scoring and feedback are driven by the e-rater AES engine. Different scoring models are created for different grade levels and sometimes for specific prompts, and the resulting scores are displayed to students and teachers. The system's interface model serves as the platform for user interaction providing learning artifacts such as online portfolios with peer-to-peer feedback, teacher feedback, and two-way student–teacher communication.

ETS designed Criterion as a venue for frequent writing practice during self-study. Criterion's extensive types of feedback on the different writing traits (i.e., grammar, usage, mechanics, style, and organization) make it an exemplar of using real-time feedback as a pedagogy to achieve desired learning outcomes. Hence, Criterion's tutor model design contributes to its successful deployment.

4.2 Research Writing Tutor

Another representative intelligent AWE tool is Research Writing Tutor (RWT) (Cotos et al., 2020; see Chapter S3C6). RWT teaches students to write scientific discourse, specifically the Introduction, Methods, Results, and Discussion/Conclusion sections of a research article. RWT has a complex interface composed of three interactive modules for learning, demonstration, and feedback (Cotos, 2017). The learning module is designed for students to understand goals specific to research writing, and the demonstration module is comprised of a wide selection of pedagogically mediated research articles demonstrating the use of effective rhetorical strategies in various disciplines (currently an annotated corpus of 32 disciplines). Together, these modules can be considered the domain model of RWT representing the expert knowledge domain. This knowledge, derived from published research articles, is used to analyze students' drafts and generate automated discipline-specific feedback.

The implementation of an expansive representation of domain-specific content and applicable pedagogies are both resource-intensive and difficult in terms of domain modeling in ITS design. RWT has been successful in deploying one such approach that is aligned with the requirements of a curriculum for learning research writing.

RWT's interface is characterized by its alignment with scaffolding on the specific rhetorical strategies of the target genre that the system is designed for.

4.3 Scaffolded Writing and ReWriting in the Disciplines (SWoRD)/Peerceptiv

SWoRD supports peer review for high school and college students (Cho & Schunn, 2007). Having undergone rapid growth and significant improvements to its student and teacher interfaces, it was renamed Peerceptiv (Schunn, 2016) and now addresses common problems with peer review, such as a lack of effort on the part of the reviewer or a tendency to be overly positive (VanDeWeghe, 2004). When using Peerceptiv, teachers provide a list of topics, due dates, and the number of reviews they want each paper to receive. Students then choose which topic they want to write about, as well as which topics they would like to review. Students receive and write reviews for the initial draft, second draft, and final draft. Reviewers are asked to provide a rating on flow, logic, and insight; to give comments; and to provide a score on a seven-point scale for each essay. The peer reviews consist of both open-ended feedback and scores that reflect the average rating of the reviewers. Peerceptiv's domain model captures the double-blind review artifacts submitted by the students in their writing and rewriting tasks across disciplines. The student model represents the students' learning progress that is captured through the ratings and grades from submitted reviews of the drafts. Peerceptiv looks at systematic differences, consistency, and spread to determine the accuracy of each review. Peerceptiv then creates a weighted grade for each essay, with less accurate reviews receiving less weight. These review-based grades are presented to the students as feedback. The peer review mechanism affords students the knowledge of expected outcomes and competencies necessary to write effective research papers (Schunn et al., 2016).

Peerceptiv's interface is among its strengths as a platform for learning writing. An interface feature worth highlighting is the students' timeline view, which clearly shows the status and progress of each writing assignment. Peerceptiv's forms reflect the appropriate affordances necessary for collaborative learning and optimizing the benefits of feedback from relevant peer reviews. For example, the Reviewing form allows students to scroll through the document while giving open-ended feedback and view the appropriate rating rubric to be used for a specific task. Peerceptiv continues to contribute to the overall classroom review process as more current work use Peerceptiv's review artifacts and corpora for further analysis related to review relevance and metareview criteria (Lam, 2021; Zhang et al., 2020).

4.4 Harry

HARRY is a web-based tutor designed to support narrative writing for students in elementary grades to engage in higher-level thinking (Holdich & Chung, 2003). HARRY's tutor model provides individualized comments and feedback to guide students as they write stories. HARRY's domain model has four story themes (pirates, space, woodland adventure, and enchanted journey) and writing tasks organized into three stages (story composition, editing, and finalizing). The students' progress in the different stages of narrative writing is instantiated in the student model. HARRY's interface model presents its narrative writing scaffolds specific to each stage (e.g., writing prompts and stylistic suggestions) as the students go through the different stages of the writing task.

HARRY's strength is its tutor model, which guides students through the writing process via conversation-based prompting. This addresses the "what next" approach of beginners as their writing evolves across revisions. With conversational dialogues, the tutor does not just deliver information or instructions but guides the students as they engage in meaning-making processes. This dialogue-based pedagogical strategy is a notable implementation of the ITS' tutor model that is anchored in educational theories with strong evidence of positive outcomes (Lefstein & Snell, 2013; Liu et al., 2019). In addition, HARRY provides help via prompts for word, sentence, and idea levels that encourage students to review and revise their work. HARRY-assisted narratives of elementary school-aged children were characterized to have varied vocabulary use, more sophisticated sentence construction, and appropriate use of punctuation than control narratives that were written without using this tool (Beam & Williams, 2015; Holdich et al., 2004).

4.5 Writing Pal

Writing Pal (Roscoe & McNamara, 2013) is the only ITS for writing developed to date. It is an online tutoring platform designed for struggling writers. It provides instructional video modules for each stage of the writing process, game-based practice, and essay-writing practice with formative and summative feedback (similar to AWE tools). Compared to most writing systems, Writing Pal has more features that depict underlying domain, student, tutor, and interface models typical of representative ITSs. Specifically, writing Pal's domain model is represented across its eight modules (Freewriting, Planning, Introduction Building, Body Building, Conclusion Building, Paraphrasing, Cohesion Building, and Revising; see Fig. 1) spanning the three main phases of writing: prewriting, drafting, and revising. Each module starts with an introductory video lesson, followed by lessons on specific strategies. For example, the Planning module includes lessons on "Positions, Arguments, and Evidence" and "Outlines and Flowcharts". The interface model includes three virtual characters (i.e., a teacher and his two students) that present instructional content. At

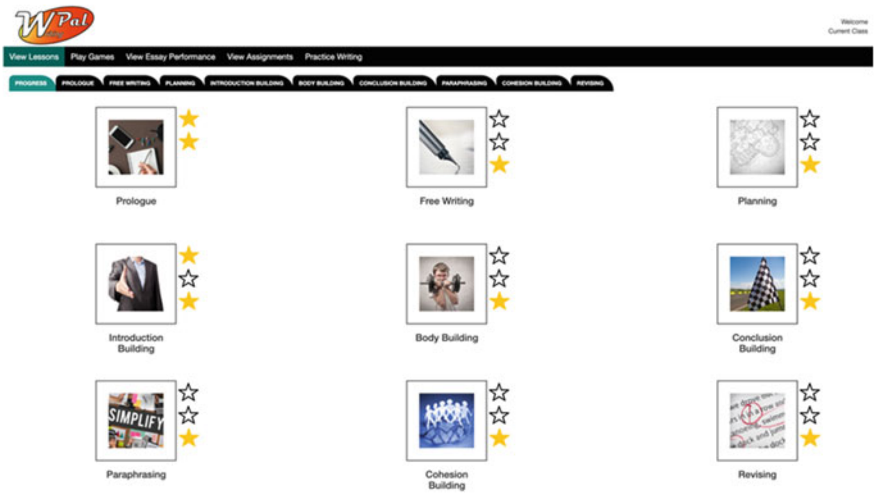

Fig. 1 Writing Pal Modules

the end of each lesson, students are tested on their knowledge via a short quiz, which is in turn incorporated into Writing Pal's evolving dynamic student model.

Writing Pal's tutor model includes multiple opportunities to practice writing strategies in the context of game-based practice and coached practice (Roscoe et al., 2014b). The games (see Fig. 2) allow students to better understand the individual strategies as well as practice using them to promote automaticity of strategy use. Specifically, identification games require students to recognize example strategies and text features (via multiple choice), such as irrelevant information in a body paragraph. Generative games require constructed responses, such as writing a topic sentence and providing evidence in response to a thesis. Practice games are inherently adaptive because advancing, leveling, and earning points during gameplay are based on performance within the game. In essence, gamification within the Writing Pal is also a form of intelligent tutoring, albeit veiled in the guise of short, dynamic games.

Writing Pal's tutor model incorporates many opportunities for practice. At the end of each module, students can write an essay in response to a prompt (i.e., whole-task practice). The essay gives students practice executing and combining the strategies they learned throughout the instructional modules and games. Each essay is automatically evaluated and scored using NLP techniques (McNamara et al., 2013, 2015). Students are presented with a score from "Poor" to "Great" along with specific suggestions for improving the essay. For instance, a short essay might receive recommendations for using freewriting to substantiate their ideas. Students are encouraged to use the feedback to revise and resubmit their essays for the second round of feedback. Although the scoring and feedback features are similar to the functionality of AES and AWE tools, Writing Pal is unique because of its dynamic tutor model—that is, formative feedback points specifically to writing strategies introduced within the

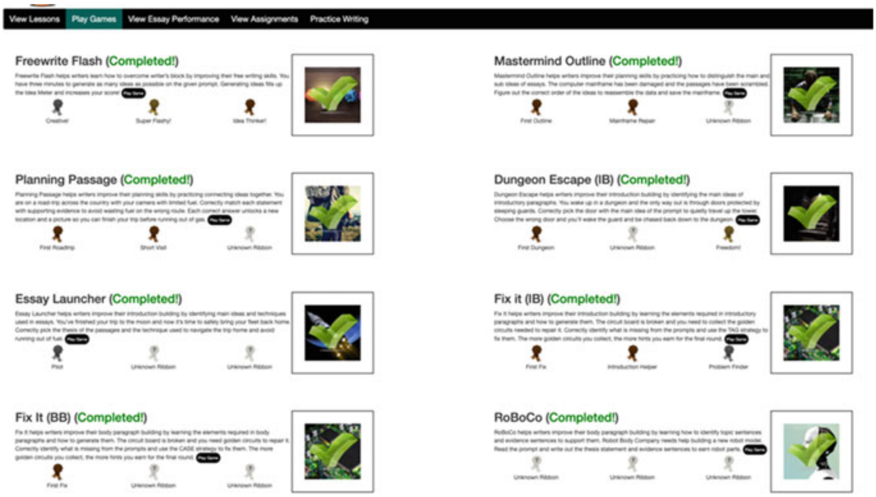

Fig. 2 Writing Pal Games

lessons and games, rather than solely to aspects of the essay that need to be fixed. Thus, there is an explicit link between the tutor model, as manifested through the lessons and games, and the feedback provided to the students based on their writing performance.

Students can either complete the modules in a fixed sequence, or flexibly choose which modules they complete, how long they interact with each module, and which games they play. The modular format of Writing Pal allows teachers to flexibly tailor instruction, including which modules to cover, their sequencing, which games to include, and the extent to which students engage in writing practice with automated feedback. Although Writing Pal's domain model comprises corpora of essay prompts that the teachers can readily use in their classes, Writing Pal also allows instructors the flexibility to incorporate their own essay prompts.

In sum, there are multiple ITS components and principles incorporated within Writing Pal, as well as functionality to customize its intelligent features. Foremost, what makes Writing Pal "intelligent" is the NLP algorithms embedded in the grading of the essays to provide formative feedback to students, which is intrinsically tied to the tutor model. The tutor model implements a wide variety of pedagogical strategies to enhance student writing, such as modular or adaptive instruction, formative and summative feedback, and gamified practice.

5 Research

ITS research has focused on investigating educational outcomes and which parameters, features, and scaffolding make ITSs effective tools for learning. Various reviews reported ITSs to be more effective than small-group instruction and some to be equivalent to one-to-one tutoring (Kulik & Fletcher, 2016; Steenbergen-Hu & Cooper, 2014; VanLehn, 2011). Across the many years of development, deployments, and subsequent commercialization, ITSs evolved to become important tools for both educators and students, as in the case of PAT (Kelkar, 2022). Significant learning gains were observed in AutoTutor implementations compared to the students reading the learning materials on their own for the same amount of time, and equivalent learning gains were observed as compared to human tutoring with experienced tutors (Graesser, 2016).

Specific to writing, ITS research continues to explore whether the integration of both cognitive and meta-cognitive processes in writing within ITSs may hold strong potential for effective scaffolding in explicit strategy instruction, increased practice opportunities, and individualized formative and summative feedback. When ITSs are designed such that their educational and theoretical anchors are clear and well-implemented in their components (i.e., domain, student, tutor, and interface models), writing instruction becomes more effective and results in the achievement of positive learning outcomes (Godwin-Jones, 2018; Roscoe et al., 2014a, 2014b). For example, classroom implementations of Writing Pal and teacher focus groups indicate that some instructors require the flexibility to cover various writing topics and modules at the classroom level, rather than allowing students to cover the material at individualized pace (Roscoe & McNamara, 2013). This is a natural tension between intelligent tutoring and the inherent nature of classroom instruction. Flexible sequencing of instructional modules is somewhat antithetical to adaptive sequencing that follows a more traditional intelligent tutoring design. Thus, parameterization in the Writing Pal interface model enables this function, allowing this tool to monitor student progress and performance and to suggest subsequent modules, lessons, or practice games. Research continues to suggest that Writing Pal's adaptive strategy instruction shows successful uptake of feedback from the tutor model during training and improves the quality of students' essays overall, as well as the more specific dimensions of essay quality (Butterfuss et al., 2022). However, ITSs, digital writing tools included, do not always lead to positive learning outcomes, especially in the absence of teacher regulation and intervention, as it found in one of Criterion's implementations (Heffernan & Otoshi, 2015).

Notably, much of the work on ITS (as well as on AWE) has focused on the development and implementation of machine learning algorithms and scaling AI to provide students with more accurate feedback necessary to monitor and assess their work. These algorithms typically leverage information from and about texts, but keystroke data have also emerged as valuable because they reveal temporal characteristics and offer insights into students' writing processes. For example, in Writing

Pal deployments, behavioral data derived from keystroke dynamics serve as important indicators of the processes that unfold in the production of written output (Allen et al., 2016; Conijn, 2020; Likens et al., 2017).

6 Conclusions and Implications for Writing Theory and Practice

The foundational ITS implementations described in this chapter, PAT, Andes Physics Tutoring System, Auto Tutor, and Writing Pal - demonstrate how the dynamic interplay of the domain, student, tutor, and interface models scale AI or intelligence to afford effective scaffolding in support of personalized learning. Specific to the writing domain, existing digital tools similarly demonstrate intelligence and emulate ITS components that result in positive learning outcomes. If the design of intelligent writing tools adheres to the underlying architecture of paradigmatic ITSs, writing instruction can become more personalized relative to the evolving context of the students. This entails designing comprehensive and adaptive ITS components that dynamically inform each other.

The scope of possible knowledge domains that might be integrated within writing ITSs is incredibly vast, and designing a complete domain model is nearly impossible. Domain models need to encompass knowledge of the language, general world knowledge, as well as knowledge of the writing task. Also, domain models should embed expertise that is sufficiently general yet representative of specialized and targeted topics, writing strategies (e.g., paraphrasing, bridging, question-asking), and writing tasks (e.g., summarization, source-based writing, argumentative writing). Student models are equally (if not more) challenging, as they should capture the dynamic and diverse students' contexts, prior knowledge, baseline skills, and individual progress. For example, student models should represent the distinct contexts of L1 and L2 student populations. Capturing the heterogeneity of the students' learning requirements should allow a respective tutor model to provide scaffolding and support pertinent to the specific needs of the students via an equally dynamic and personalized instantiation of the interface model. Furthermore, writing ITSs may benefit from a greater focus on enhancing the user experience through the implementation of more engaging and immersive student interfaces. Future writing systems have the potential to improve system interaction when navigating the system, recovering from errors, and receiving feedback by implementing dialogue-based interfaces as in Andes, empathic chatbots as in Auto Tutor, animated pedagogical agents as in Writing Pal, and augmented reality-enhanced user interfaces, among others. The user interface should be flexibly designed to be conducive to a specific learning goal given students' learning context and writing task at hand. For example, enhanced user experience and learning outcomes can be achieved by ensuring the correspondence between the expected written output and the size of the text boxes (as in Peerceptiv), using mini-games for supplemental practice opportunities for complex writing tasks (as in Writing Pal), and implementing text-to-speech functionality for longer texts.

Future ITSs for writing will continually face the challenges of (1) personalized instruction adapted to evolving student attributes, (2) provision of appropriate and relevant instruction contingent on the domain and student, (3) provision of formative and summative feedback, (4) appropriate design of user interface elements to facilitate learning, and (5) tensions between classroom instruction and adaptive instruction, to name a few. Nonetheless, incorporating intelligent tutoring principles within digital writing technologies has strong potential to improve performance for the learning and teaching of writing. In their present form, digital writing tools have yet to fully optimize the canonical and cutting-edge features of modern ITSs when it comes to AI-enabled domain, pedagogical, tutoring, and intelligent interface designs. Despite their known benefits, there is still untapped potential and much room for improvement to serve as an impetus for subsequent work in this area.

7 Tools

No	Tools	Descriptors	References/links
1	Andes Physics Tutoring System	Non-writing ITS, physics, homework problem-solving support	Gertner and VanLehn (2000) and VanLehn et al. (2005)
2	AutoTutor	Non-writing ITS, computer literacy, physics, conversational ITS, NLP-enabled dialogue system	Graesser et al. (2001)
3	Criterion Online Writing Evaluation Service	NLP-based assessment and formative error-correction feedback	Burstein et al. (2004), Burstein et al. (2013), and Ramineni and Deane (2017) https://www.ets.org/criterion.html
4	HARRY	Web-based tutor, narrative writing, dialog-based prompts, conversational dialogues	Holdich and Chung (2003)
5	Practical Algebra Tutor	Non-writing ITS, algebra, step-based tutor, cognitive task analysis	Koedinger et al. (1997)
6	Research Writing Tutor	Discipline-specific rhetorical feedback on scientific writing, genre-based learning	Cotos (2017)
7	Scaffolded Writing and ReWriting in the Disciplines (Sword)/ Peerceptiv	Peer review platform, feedback and scores based on reviewer ratings	Cho and Schunn (2007) https://peerceptiv.com/

(continued)

(continued)

No	Tools	Descriptors	References/links
8	Writing Pal	Web-based Tutor, platform for struggling readers, NLP algorithms, adaptive instruction;	Roscoe and McNamara (2013) http://www.adaptiveliteracy.com/

Acknowledgements The authors disclosed receipt of the following financial support for the research, author-ship, and/or publication of this chapter. This work was supported in part by IES Grants R305A180261 and R305A180144, as well as the Office of Naval Research (Grant: N00014-17-1-2300). Opinions, conclusions, or recommendations do not necessarily reflect the view of the Department of Education, IES, or the Office of Naval Research.

References

Ahuja, N. J., & Sille, R. (2013). A critical review of development of intelligent tutoring systems: Retrospect, present, and prospect. *International Journal of Computer Science Issues, 10*, 39–48.

Akyuz, Y. (2020). Effects of intelligent tutoring systems (ITS) on personalized learning (PL). *Creative Education, 11*(6), 953–978.

Allen, L. K., Jacovina, M. E., Dascalu, M., Roscoe, R., Kent, K., Likens, A., & McNamara, D. S. (2016). {ENTER}ing the time series {SPACE}: Uncovering the writing process through keystroke analyses. *9th International Conference on Educational Data Mining (EDM 2016)* (pp. 22–29). International Educational Data Mining Society.

Almasri, A., Ahmed, A., Al-Masri, N., Sultan, Y. A., Mahmoud, A. Y., Zaqout, I., Akkila, A. N., & Abu-Naser, S. S. (2019). Intelligent tutoring systems survey for the period 2000–2018. *International Journal of Academic Engineering Research, 3*, 21–37.

Beam, S., & Williams, C. (2015). Technology-mediated writing instruction in the early literacy program: Perils, procedures, and possibilities. *Computers in the Schools, 32*(3–4), 260–277.

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*, 4–16.

Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The Criterion online writing system. *AI Magazine, 25*, 27–36.

Burstein, J., Tetreault, J., & Madnani, N. (2013). The e-rater automated essay scoring system. In M. D. Shermis, & J. Burstein (Eds.), *Handbook of automated essay evaluation: Current applications and new directions* (pp. 55–67). Routledge.

Butterfuss, R., McCarthy, K. S., Roscoe, R. D., Allen, L. K., & McNamara, D. S. (2022). Strategy uptake in W-Pal: Adaptive feedback and instruction. *Journal of Educational Computing Research, 60*, 696–721.

Cai, Z., Hu, X., & Graesser, A. C. (2019). Authoring conversational intelligent tutoring systems. In R. Sottilare, & J. Schwarz (Eds.), *Adaptive instructional systems. HCII 2019a. Lecture notes in computer science* (Vol. 11597). Springer.

Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education, 48*, 409–426.

Conijn, R. (2020). *The keys to writing: A writing analytics approach to studying writing processes using keystroke logging* (Doctoral dissertation). Tilburg University/University of Antwerp.

Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Intelligent tutoring systems. In M. G. Helander, T. K. Landauer, & P. V. Prabhu (Eds.), *Handbook of human-computer interaction* (2nd Rev. Ed., pp. 849–874). Elsevier Science & Technology.

Cotos, E. (2018). Automated writing evaluation. In J. I. Liontas (Ed.), *The TESOL encyclopedia of English language teaching*. Wiley.
Cotos, E. (2017). Computer-assisted research writing in the disciplines. In S. A. Crossley & D. S. McNamara (Eds.), *Adaptive educational technologies for literacy instruction* (pp. 225–242). Routledge.
Cotos, E., Huffman, S., & Link, S. (2020). Understanding graduate writers' interaction with and impact of the Research Writing Tutor during revision. *Journal of Writing Research, 12*(1), 187–232.
D'Mello, S., Picard, R. W., & Graesser, A. (2007). Toward an affect-sensitive AutoTutor. *IEEE Intelligent Systems, 22*(4), 53–61.
Deane, P., Odendahl, N., Quinlan, T., Fowles, M., Welsh, C., & Bivens-Tatum, J. (2008). *Cognitive models of writing: Writing proficiency as a complex integrated skill* (Research Report No. RR-08–55). Educational Testing Service.
Derry, S. J., Hawkes, L. W., & Ziegler, U. (1988). A plan-based opportunistic architecture for intelligent tutoring. *Proceedings of Intelligent Tutoring Systems (ITS-88),* 116–123.
Gertner, A. S., & VanLehn, K. (2000, June). Andes: A coached problem solving environment for physics. In *International conference on intelligent tutoring systems* (pp. 133–142). Springer.
Godwin-Jones, R. (2018). Second language writing online: An update. *Language Learning and Technology, 19*, 1–15.
Graesser, A. C. (2016). Conversations with AutoTutor help students learn. *International Journal of Artificial Intelligence in Education, 26*(1), 124–132.
Graesser, A. C., Conley, M. W., & Olney, A. (2012). Intelligent tutoring systems. In K. R. Harris, S. Graham, T. Urdan, A. G. Bus, S. Major, & H. L. Swanson (Eds.), *APA educational psychology handbook: Application to learning and teaching* (Vol. 3, pp. 451–473). American Psychological Association.
Graesser, A. C., Jackson, G. T., & McDaniel, B. (2007). AutoTutor holds conversations with learners that are responsive to their cognitive and emotional states. *Educational Technology, 47*, 19–23.
Graesser, A. C., Person, N., Harter, D., & Tutoring Research Group. (2001). Teaching tactics and dialog in AutoTutor. *International Journal of Artificial Intelligence in Education, 12*(3), 257–279.
Heffernan, N., & Otoshi, J. (2015). Comparing the pedagogical benefits of both Criterion and teacher feedback on Japanese EFL students' writing. *JALT CALL Journal, 11*(1), 63–76.
Holdich, C. E., & Chung, P. W. H. (2003). A 'computer tutor' to assist children develop their narrative writing skills: Conferencing with HARRY. *International Journal of Human-Computer Sciences, 59*, 631–669.
Holdich, C. E., Chung, P. W., & Holdich, R. G. (2004). Improving children's written grammar and style: Revising and editing with HARRY. *Computers & Education, 42*(1), 1–23.
Kelkar, S. (2022). Between artificial intelligence and learning science: The evolution and commercialization of intelligent tutoring systems. *IEEE Annals of the History of Computing*.
Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*(1), 30–43.
Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research, 86*(1), 42–78.
Lam, S. T. E. (2021). A web-based feedback platform for peer and teacher feedback on writing: An Activity Theory perspective. *Computers and Composition, 62*, 102666.
Lefstein, A., & Snell, J. (2013). *Better than best practice: Developing teaching and learning through dialogue*. Routledge.
Likens, A. D., Allen, L. K., & McNamara, D. S. (2017). Keystroke dynamics predict essay quality. In *Proceedings of the 39th annual meeting of the cognitive science meeting (CogSci 2017)* (pp. 2573–2578). London, UK.
Liu, T., Yuizono, T., Lu, Y., & Wang, Z. (2019). Application of human-machine dialogue in foreign language teaching at universities. In *IOP Conference Series: Materials Science and Engineering* (Vol. 573, No. 1, p. 012047). IOP Publishing.

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*, 901–918.

McNamara, D. S., Crossley, S. A., & Roscoe, R. D. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods, 45*, 499–215.

McNamara, D. S., Crossley, S. A., Roscoe, R. D., Allen, L. K., & Dai, J. (2015). A hierarchical classification approach to automated essay scoring. *Assessing Writing, 23*, 35–59.

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

Powell, P. R. (2009). Retention and writing instruction: Implications for access and pedagogy. *College Composition and Communication, 60*, 664–682.

Ramineni, C., & Deane, P. (2017). The Criterion® Online Writing Evaluation Service. In S. A. Crossley & D. S. McNamara (Eds.), *Adaptive educational technologies for literacy instruction* (pp. 178–198). Routledge.

Roscoe, R. D., Allen, L. K., Weston, J. L., Crossley, S. A., & McNamara, D. S. (2014a). The Writing Pal intelligent tutoring system: Usability testing and development. *Computers and Composition, 34*, 39–59.

Roscoe, R. D., Brandon, R. D., Snow, E. L., & McNamara, D. S. (2014b). Game-based writing strategy practice with the Writing Pal. In K. E. Pytash, & R. E. Ferdig (Eds.), *Exploring technology for writing and writing instruction* (pp. 1–20). IGI Global.

Roscoe, R. D., & McNamara, D. S. (2013). Writing Pal: Feasibility of an intelligent writing strategy tutor in the high school classroom. *Journal of Educational Psychology, 105*, 1010–1025.

Sale, K., & Muldner, K. (2019). Learning with an algebra computer tutor: What type of hint is best?. In *CogSci* (pp. 2708–2714).

Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C., & Bigdeli, A. (2008). "How do you know that I don't understand?" A look at the future of intelligent tutoring systems. *Computers in Human Behavior, 24*(4), 1342–1363.

Schunn, C. D. (2016). Writing to learn and learning to write through SWoRD. In S. A. Crossley & D. S. McNamara (Eds.), *Adaptive educational technologies for literacy instruction* (pp. 243–260). Routledge.

Schunn, C., Godley, A., & DeMartino, S. (2016). The reliability and validity of peer review of writing in high school AP English classes. *Journal of Adolescent & Adult Literacy, 60*(1), 13–23.

Sottilare, R. A., Graesser, A., Hu, X., & Holden, H. (Eds.). (2013). *Design recommendations for intelligent tutoring systems: Learner modeling* (Vol. 1). US Army Research Laboratory.

Sottilare, R. A., Graesser, A. C., Hu, X., Olney, A., Nye, B., & Sinatra, A. M. (Eds.). (2016). *Design recommendations for intelligent tutoring systems: Domain modeling* (Vol. 4). US Army Research Laboratory.

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.

VanDeWeghe, R. (2004). "Awesome, Dude!" Responding helpfully to peer writing. *English Journal, 94*, 95–99.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*, 197–221.

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., & Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education, 15*, 147–204.

Zhang, F., Schunn, C., Li, W., & Long, M. (2020). Changes in the reliability and validity of peer assessment across the college years. *Assessment & Evaluation in Higher Education, 45*(8), 1073–1087.

Dr. Michelle Banawan is a postdoctoral research scholar at Arizona State University. Her research interests include machine learning, educational data mining/learning analytics and artificial intelligence in education.

Dr. Reese Butterfuss is a postdoctoral research scholar at Arizona State University. His research focuses on the cognitive processes that underlie learning from texts, as well as improving students' literacy skills using technology-based literacy instruction.

Dr. Karen S. Taylor is an Education Researcher at American Institutes for Research (AIR). Her work focuses on identifying and applying evidence-based strategies to support reading and writing development and instruction.

Katerina Christhilf is a graduate student in Psychology at Arizona State University under Dr. Danielle McNamara. She obtained a B.S. in Cognitive Sciences, summa cum laude, from the University of California, Irvine. Her interests include reading comprehension, educational interventions, cognitive training, intelligent tutoring systems, and educational data mining. Her ongoing research projects include dynamic analyses of readers' constructed responses, game-based stealth assessment of literacy skills, and predicting course performance as a function of prior knowledge and reading skill.

Claire Hsu is undergraduate Research Aides for Arizona State University's SoLET Lab majoring in Psychology. The focus of their research is on examining the process of learning and misconceptions influencing knowledge revision.

Connor O'Loughlin is undergraduate Research Aides for Arizona State University's SoLET Lab majoring in Psychology. The focus of their research is on examining the process of learning and misconceptions influencing knowledge revision.

Dr. Laura K. Allen is an Assistant Professor of Educational Psychology at University of Minnesota Twin Cities. The primary aim of her research is to examine how individuals learn and communicate with text and to apply those insights to educational practice through the development of interventions and educational technologies.

Dr. Rod D. Roscoe is an Associate Professor of Human Systems Engineering in the Ira A. Fulton Schools of Engineering at Arizona State University. His research combines insights from learning science, computer science, and design science to improve the implementation and effectiveness of equitable educational technologies.

Dr. Danielle S. McNamara Professor and Director of the Science of Learning and Educational Technology (SoLET) Lab, Arizona State University. Dr. McNamara develops educational technologies (iSTART, iSTART-ME, Coh-Metrix, Writing-Pal) and conducts research to better understand cognitive processes of comprehension, learning, text coherence, and individual differences. With decades of experience as a senior researcher, Dr. McNamara has solidified herself as a one of the world's premier experts in cognitive psychology, publishing hundreds of scholarly works as well as delivering several keynote addresses at conferences.