



www.ijtes.net

Systematic Review of Intelligent Tutoring Systems for Hard Skills Training in Virtual Reality Environments

Joakim Laine 

University of Helsinki, Helsinki, Finland

Timo Lindqvist 

Upknowledge Inc., Finland

Tiina Korhonen 

University of Helsinki, Helsinki, Finland

Kai Hakkarainen 

University of Helsinki, Helsinki, Finland

To cite this article:

Laine, J., Lindqvist, T., Korhonen, T., & Hakkarainen, K. (2022). Systematic review of intelligent tutoring systems for hard skills training in virtual reality environments. *International Journal of Technology in Education and Science (IJTES)*, 6(2), 178-203. <https://doi.org/10.46328/ijtes.348>

The International Journal of Technology in Education and Science (IJTES) is a peer-reviewed scholarly online journal. This article may be used for research, teaching, and private study purposes. Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material. All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations regarding the submitted work.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Systematic Review of Intelligent Tutoring Systems for Hard Skills Training in Virtual Reality Environments

Joakim Laine, Timo Lindqvist, Tiina Korhonen, Kai Hakkarainen

Article Info

Article History

Received:

14 October 2021

Accepted:

12 March 2022

Keywords

Intelligent tutoring systems

Virtual reality

Hard skills training

Systematic literature review

Abstract

Advances in immersive virtual reality (I-VR) technology have allowed for the development of I-VR learning environments (I-VRLEs) with increasing fidelity. When coupled with a sufficiently advanced computer tutor agent, such environments can facilitate asynchronous and self-regulated approaches to learning procedural skills in industrial settings. In this study, we performed a systematic review of published solutions involving the use of an intelligent tutoring system (ITS) to support hard skills training in an I-VRLE. For the seven solutions that qualified for the final analysis, we identified the learning context, the implemented system, as well as the perceptual, cognitive, and guidance features of the utilized tutoring agent. Generally, the I-VRLEs emulated realistic work environments or equipment. The solutions featured either embodied or embedded tutor agents. The agents' perception was primarily based on either learner actions or learner progress. The agents' guidance actions varied among the solutions, ranging from simple procedural hints to event interjections. Several agents were capable of answering certain specific questions. The cognition of the majority of agents represented variations on branched programming. A central limitation of all the solutions was that none of the reports detailed empirical studies conducted to compare the effectiveness of the developed training and tutoring solutions.

Introduction

Recent advances in virtual reality (VR) hardware and software technologies have made it possible to provide professional hard skills training within virtual environments. Furthermore, according to a meta-analysis conducted by Kulik and Fletcher (2016), intelligent tutoring systems (ITSs) have now surpassed the effectiveness of computer-assisted instruction and human tutoring (p. 67). We are interested in the combination of VR and ITS technologies in the context of developing immersive virtual reality learning environments (I-VRLEs) for hard professional skills training. We hope that studying and developing I-VRLEs will render the self-study of professional hard skills possible. The present work focuses on studying the different types of self-study-enabling hard skills training solutions that have been developed and assessed in the past. In this context, professional hard skills are the technical skills necessary to operate diverse pieces of equipment and conduct procedures related to professional work. The applied training methods are generally performance-based training

programs delivered by an expert in the relevant skill (Laker & Powell, 2011). Methods such as on-the-job training, master-apprentice teaching, and live-action simulations, whereby employees are trained in real working environments using a hands-on approach, represent popular means of hard skills training.

VR technologies utilize real-time motion tracking and three-dimensional (3D) graphics to create various imaginary yet realistic synthetic spaces. They are considered most appropriate for providing training in what would otherwise involve complex and high-risk situations (Zahabi & Razak, 2020). Prior studies have tended to overlook the fact that VR offers a more sustainable and cheaper way to provide training regarding certain resource-heavy work tasks, while virtualization allows product designs (e.g., machines, tools, and equipment) to be tested in order to detect potential faults or impracticalities (Steffen et al., 2019, p. 698). Regardless, VR technology can offer interactive immersion in realistic situations, including access to high-fidelity virtual objects and tools. In their theoretical analysis, Petukhov et al. (2017, p. 5) noted that the use of VR for professional training purposes can lower both the cost of training and the risk of damage to equipment and trainees. Moreover, they discovered that VR is already being used to perfect skilful actions and teach conceptual models of professional activity in various professional fields (e.g., industry, military, and education) (p. 3). In addition, VR has also been applied to help develop certain professionally important qualities, for example, developing the ability to maintain emotional stability during fire outbreak simulations (pp. 3-4).

Recent systematic literature reviews concerning ITSs have focused on their use in relation to psychomotor skills training (Neagu et al., 2020), learning path personalization (Nabizadeh et al., 2020), and research trends (Han et al., 2019; Soofi & Ahmed, 2019). ITSs have only rarely been integrated into VR due to the need for some type of virtual reality tutor extension to their interface model (Neagu et al., 2020). Yet, VR offers valuable opportunities for the more comprehensive tracking of learner behaviour, and the combination of these developing technologies has the potential to offer benefits in terms of professional training, meaning that their presence in companies' training programs will likely increase in the near future (Alcañiz et al., 2018, pp. 3-4). In the present systematic literature review, we searched well-known databases for papers on VR-based ITSs, reviewed the relevant papers, analysed the review data in light of our research questions, and reported our findings. This investigation was intended to support our efforts with regard to developing a machine-learning-trained artificial intelligence (AI) tutor to observe immersed learners and scaffold (i.e., offering appropriate, adaptive, and timely guidance) hard professional skills training within an I-VRLE. More specifically, the study focused on reviewing and cataloguing published literature in an attempt to answer the following questions:

1. What solutions featuring an ITS for hard skills training in an I-VRLE have been described in the literature?
2. What types of I-VRLEs have been implemented?
3. What kinds of tutoring systems have been implemented?

Immersive Virtual Reality Learning Environments

In this study, our intention was to focus on implementations that utilized highly immersive virtual realities (I-VRs) in their training solutions. Two types of immersion can be distinguished from one another, namely system

immersion and mental immersion. System immersion concerns the objective and measurable quality of a system, and highly immersive systems are capable of delivering an “inclusive, extensive, surrounding, and vivid illusion of reality” (Slater & Wilbur, 1997, pp. 3-4). Mental immersion, however, involves a subjective experience in which users achieve a sense of being in the VR space. Mental immersion involves constructing a spatialized mental model of the environment and then becoming convinced it is one’s primary frame of self-reference, that is, reasoning that one is located in the relevant space (Wirth et al., 2007, pp. 497-498). Some aspects of system immersion have stronger effects on mental immersion than others. According to Cummings and Bailenson (2016, pp. 296-297), tracking level, stereoscopy, and field of view are more influential aspects than sound quality and visuals. Moreover, according to a recent characterization by Kardong-Edgren et al. (2019), VR implementations involving head-mounted devices (HMDs) are likely to offer high levels of immersion. Other aspects they suggest should be considered include the system’s vividness, inclusiveness, extensiveness, and tracking acuity (Kardong-Edgren et al., 2019, p. 32).

I-VR technology seeks to replace users’ perception of the physical world with a high-fidelity computer-generated 3D and multi-sensory virtual environment. Knowledge is embedded in the environment’s activities, mechanisms, and objects. When the environment is constructed for the purpose of formal training, it can be referred to as an I-VRLE. In practice, I-VRLEs are used to achieve certain desired outcomes (i.e., learning goals). The constraints and affordances of a given I-VRLE can be fully modified, and they can be pointed out, hidden, or locked at different times. According to Steffen et al. (2019, pp. 689-690, 720-721), I-VR can be used to diminish negative aspects of the physical world, enhance positive aspects of the physical world, recreate aspects of the physical world, and create aspects that do not exist in the physical world. In terms of education and training, I-VR can be used with projective or interactive devices, on single-user or multi-user platforms, and to enhance the learning experience, engage the participants, motivate the participants, or improve participant achievement (Kurniawan et al., 2019, pp. 2-3). In this study, we were interested in I-VRLEs designed for hard skills training. When evaluating the existing I-VR and ITS applications, we paid particular attention to the intended learning objectives, the suggested training tasks, the utilized hardware, and the design of the virtual worlds.

Intelligent Tutoring Systems

The purpose of all ITSs is to automate instruction and personalize the learning experience. They are computer systems that typically consist of multiple modules, including a domain module, learner module, tutor module, interface module, and pedagogical module (Amokrane et al., 2008, p. 186). Their goal is to observe and assist learners. It has been suggested that ITSs have the capability to replace traditional educational methods with more adaptive methods delivered through or augmented with digital technologies (Soofi & Ahmed, 2019, p. 106). Such methods are intended to take into consideration the individual learner’s needs and capabilities and then to deliver personalized learning through adapting the learning materials accordingly (Virvou et al., 2003, p. 4872). In recent years, the most popular foundations for the development of tutoring systems have been web-based and computer-application-based modes (Soofi & Ahmed, 2019, p. 104). Moreover, ITSs have generally been studied from the computer sciences perspective and at a university level (Han et al., 2019, pp. 156-157;

Mousavinasab et al., 2021, p. 16; Soofi & Ahmed, 2019, p. 100). Generally speaking, it has been concluded that ITSs for academic subjects exert moderately strong effects on learning outcomes (Kulik & Fletcher, 2016, p. 67). However, in the present study, we sought to uncover more novel implementations of ITSs specifically designed for I-VRLEs and hard skills training. Thus, we focused on these tutoring systems’ perceptual capabilities, underlying cognitive processing, and guidance actions.

Methods

Database Queries

We adopted the systematic review process outlined by Neagu et al. (2020, pp. 2-4) and Soofi and Ahmed (2019, p. 100), which included several phases and review steps, as shown in **Hata! Başvuru kaynağı bulunamadı.** During the preparation phase, we discussed the purpose of this review paper and established the research questions, which we then used to develop the inclusion and exclusion criteria. The relevant online databases were identified based on prior systematic reviews concerning ITS technology. We performed comprehensive searches in Scopus, IEEE Xplore, Web of Science, ScienceDirect, and ACM Digital Library and systematically narrowed the query down to one that could be repeated in all of the databases.

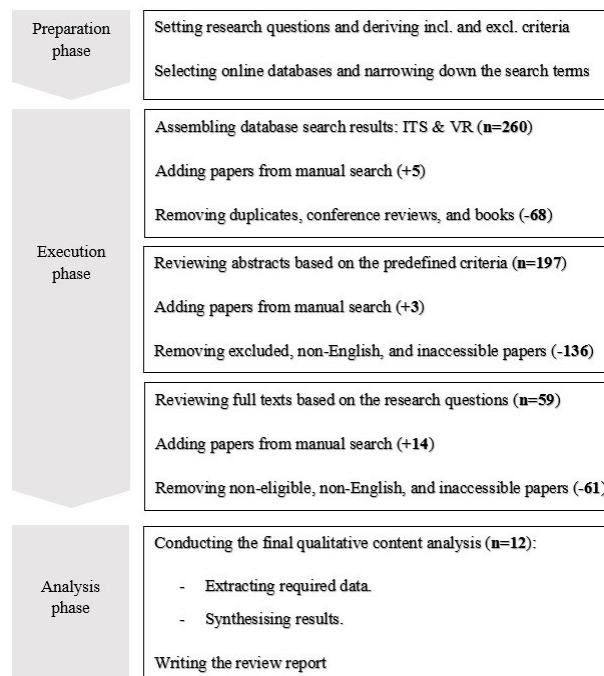


Figure 1. Phases and Steps of the Systematic Review Process

We then moved on to the execution phase, which involved performing the query in a case-insensitive form using the title, abstract, and keyword fields of each database, with the exception of ACM Digital Library, where only the abstract field was available for the search. The query regarding “intelligent tutoring system” resulted in 8340 papers being identified (see Table 1). As we were not interested in papers on ITSs outside of the VR context, we included the term “virtual reality” in the search, which significantly reduced the number of results to 260. Scopus contributed the largest number of results. We then added five papers that we had previously identified as

being potentially relevant when conducting the comprehensive queries. In addition, we removed all of the duplicate papers, conference reviews, and books from the results before reviewing the abstracts of the target articles.

Table 1. Number of Query Results in Each Database

Database	Target fields	Results for “intelligent tutoring system”	Results for “intelligent tutoring system” AND “virtual reality”
Scopus	TITLE-ABS-KEY	5203	189
IEEE Xplore	“All metadata”	595	45
Web of Science	Topic (title, abstract, key words)	1749	19
ScienceDirect	Title, abstract, or author-specified keywords	404	6
ACM Digital Library	Abstract	139	1
ALL	-	8340	260

Abstract Review

Next, we read through the abstracts of all the remaining papers (n=197) and assessed them based on the previously established inclusion and exclusion criteria (see Table 2). Papers that met at least two inclusion criteria and did not meet any exclusion criteria were accepted for a full-text review. In the case of 62 papers, merely reading the abstract was not enough to make a decision regarding their eligibility. We could not determine whether they had implemented an immersive form of VR technology, what the exact training topic was, or whether a tutoring system was implemented. To address these issues, we performed an additional content-based inspection of the relevant papers. Three articles were added to the review following the inspection, having been discovered when we searched for the full-text versions of the papers. After the inspection, a total of 59 papers were approved for a full-text review, while 136 papers were removed. The removed papers did not meet the inclusion criteria, could not be retrieved, or were not available in English.

Table 2. Inclusion and Exclusion Criteria

Criteria	Action
Concerns hard skills training	Include
Describes a tutoring system	Include
The tutoring system guides learners towards achieving learning goals	Include
Describes the VR context	Include
A framework proposition or a case study	Include
The training context is clearly non-immersive VR	Exclude
Concerns training in soft skills or academic skills, such as communication or mathematics	Exclude
The focus is different, such as authoring tool development, child development, etc.	Exclude
A review or a meta-analysis.	Exclude

During the additional inspection, we learned that some authors had published multiple papers concerning the same solution. These papers were closely linked together throughout the remainder of the systematic review. Moreover, we combined various authors into supposed research groups based on their affiliations with one another. This also helped us to locate papers on the same solution during the full-text review step.

Full-text Review

We performed a full-text review of the approved papers (n=59) after the abstract review. We re-visited our research questions and identified within them aspects that the papers should describe (see Table 3). The reviewed papers' eligibility for the final analysis was determined based on whether they discussed and described the proposed hard skills training solution integrating VR and ITS.

Table 3. Reviewed Aspects of the I-VRLEs and ITSs

RQ1 Overview	RQ2 I-VRLE implementation	RQ3 Tutoring solution
Authors	Learning objectives	Perception
Domain	Training tasks	Guidance actions
Overall concept	Virtual reality hardware Virtual reality world	Cognition

The first author of the present work read each paper during the full-text review and marked all instances where they described the aspects of interest. More papers were linked together based on the studied training solutions. Furthermore, 14 new papers that detailed some aspects of the same training solutions were added to the review during this step. Thus, a total of 73 papers were fully reviewed. Prior to determining the final sample, two researchers re-reviewed the suggestions for the final analysis and discussed the papers' eligibility. Overall, 61 papers that could not be assessed (inaccessible or non-English) or did not provide an adequate description of using an ITS in a VR context were removed. The final accepted papers (n=12) provided the required descriptions of the utilized I-VRLEs and tutoring solutions.

Final Analysis

During the final generic qualitative content analysis step (Mihás, 2019), two researchers re-read and marked the accepted papers (n=12) individually. All of the marked phrases and paragraphs relevant to the tutoring systems and I-VRLEs were entered into a Microsoft Excel sheet and organized according to the research questions. Furthermore, we collected images of each solution and used them to interpret aspects of the constructed VR worlds and tutoring systems. After extracting and organizing the relevant material, we analysed its content across each aspect category (see Table 3). More specifically, for each aspect category, we read through the material and established a general sense of it. Then, we began coding the descriptions with labels one aspect and one training solution at a time. Whenever a new label emerged, we rechecked the previous papers' descriptions for any instances of that label. The generated labels were inserted into landscape tables and more detailed tables that are broken down in the results section. A research question sometimes demanded descriptive answers that

did not require such an in-depth content analysis. In such cases, we focused on achieving the correct interpretation of the extracted data. The reliability of the interpretations was assessed by means of discussions between the two researchers. Finally, all of the researchers assessed the results tables and made any necessary revisions.

Results

Solutions Described in the Literature

The 12 papers included in our final analysis described seven separate solutions utilizing an ITS for hard skills training within an I-VRLE (see Table 4). Among them, Solutions 1–3 were described in single papers, while Solutions 4–7 were discussed in more than one reviewed paper. Solution 2 (“Anatomy”) was included as an exception. It passed the review process and implemented the latest I-VR; however, it merely showed promise as a platform for hard skills training (it included no drills or practice). Solution 6 (“Osprey”) presented two training solution variants applied in two different training scenarios that had different learning objectives (Buck et al., 2018).

Table 4. Identified Solutions

Solution	Domain	Concept description	Reviewed papers
1 “Driving”	Driving	An evaluative driving simulation with adaptive selection of training activities	Ropelato et al. (2018)
2 “Anatomy”	Biology	Free-form exploration of a frog’s anatomic structures and concepts	Ahn et al. (2018)
3 “Printer”	Additive manufacturing	Perform machine setup procedures according to instructions	Mogessie et al. (2020)
4 “Blood”	Industrial blood analysis	Perform procedures in the guidance and with the assistance of an embodied tutor agent	Taoum et al. (2016) Querrec et al. (2018)
5 “Steve”	Naval bridge operations	Perform procedures in the guidance and with the assistance of an embodied tutor agent	Rickel & Johnson (1998) Rickel & Johnson (1999)
6 “Osprey”	(a) Aircraft maintenance (b) Flight crew coordination	(a) An evaluative maintenance simulation within a virtual reality environment (b) Perform technical and teamwork tasks with a virtual teammate	Buck et al. (2018) Buck et al. (2017)
7 “Machine”	Machine operation (subject-independent)	Acquire abilities by executing activities within an inhabited virtual reality environment (i.e., I-VRLE featured other learners, auxiliary characters, and virtual tutors)	Sánchez & Imbert (2007a) Imbert et al. (2007) Sánchez & Imbert (2007b)

Notes. Each solution was randomly numbered and assigned a nickname for easy reference. For Solution 6 (“Osprey”), the papers described two separate training scenarios (denoted a and b when a distinction was required)

We wanted to determine whether the solutions were intended for a specific application domain, or if they could be applied more generally. Except for Solution 7 (“Machine”), all of the reviewed solutions were described within the context of a specific task domain. While the “Machine” solution was prototyped in the machine operation domain, it was proposed as a subject-independent multi-agent tutoring solution. A similar desire for generalization was also mentioned in relation to Solution 3 (“Printer”); however, the latter’s generalization was intended to support the learning of similar procedures using other additive manufacturing equipment models.

Following our evaluation of the overall concepts, we noted the active role of the learner in all of the solutions. For instance, Solution 2 (“Anatomy”) was based on learner-initiated explorative knowledge acquisition, while Solution 6b (“Osprey” b) was based on learner participation in the role-playing of a real-world situation accompanied by a virtual non-player character (NPC) teammate. In all of the other solutions, the learner was engaged in performing simulated real-world tasks. In Solutions 6a (“Osprey” a) and 3 (“Printer”), the VR training was part of a blended learning solution comprising a declarative learning part performed on a personal computer (PC) followed by an evaluative practice part performed in the I-VRLE.

The papers describing Solutions 5 (“Steve”) and 7 (“Machine”) were published prior to 2008, while the papers concerning the other solutions were published after 2015. The earlier publications described a training solution or tutoring system design, although they did not include any validation or evaluative study of the proposed solution. In fact, only Solutions 1 (“Driving”) and 2 (“Anatomy”) involved user studies. In “Driving,” Ropelato et al. (2018) studied the VR setup’s effects on simulator sickness and presence, while in “Anatomy,” Ahn et al. (2018) studied the ease of use of the interactive elements as well as the helpfulness of their system’s open educational resources (OER) search mechanism. For the other solutions, no validation studies were performed, although the authors expressed the intention to perform such studies in the future: “test and prove the generalized framework” with end users (Mogessie et al., 2020, p. 358; “Printer”), test “the impact of our proposition on the performance of the learner... influence of the presence of the virtual agent when learning a procedure” (Taoum et al., 2016, pp. 347-348; “Blood”), and test the effectiveness of the blended approach and its elements (Buck et al., 2017, p. 6; “Osprey” a).

I-VRLE Implementations

Learning Objectives and Training Tasks

Each solution offered a description of the relevant training scenario. We collected and reviewed each solution’s learning objectives and training tasks (see Table 5). The learning objectives focused on the development of unique skills related to the use of various machines and pieces of equipment (in all of the solutions except “Anatomy”), of procedural models related to various setup, maintenance, and operating tasks (in all of the solutions except “Anatomy” and “Driving”), and of adherence to specified safety procedures (in “Printer” and “Osprey” a). “Anatomy” sought to teach the definitions of the anatomical parts of a frog and tested the OER mechanism. The two-dimensional (2D) PC learning environments utilized in “Printer” and “Osprey” involved declarative learning objectives related to acquiring knowledge regarding the topic of interest. The I-VRLEs were used to practice tasks (“Printer”) or

evaluate abilities (“Osprey” a) at the end of the blended learning courses. Finally, the prototype I-VRLE associated with “Machine” aimed to educate the learner regarding the attributes of clothing and how they impact laundry tasks.

Table 5. Training Solutions’ Learning Objectives and Training Tasks

Solution	Learning objective(s)	Training tasks
“Driving”	Abilities of a good driver	Drive around following directions and react to both traffic and events
“Anatomy”	Anatomical structures and parts of a frog	Select virtual frog parts, receive instructions and descriptions, and search for more information using the OER search
“Printer”	Safety procedures, equipment components, and operating steps associated with an EOS M290 3D printer	Follow instructions and carry out tasks related to the printer’s operating and safety procedures
“Blood”	Operational actions using an automaton and preparation of reagents	Follow a predefined sequence of pedagogical assistance and conduct a blood analysis setup process
“Steve”	Operating procedures associated with naval equipment and machinery	Interact with a pedagogical agent demonstrating naval machine operational procedures
“Osprey”	(a) Correct aircraft maintenance tasks, use of equipment, fault diagnosis, and repair procedures (b) Flight crew coordination abilities and flight deck use	(a) Navigate around an aircraft, perform practical maintenance tasks, and follow safety protocols (b) Work with a “Virtual Pilot”, operate the flight deck, and react to guidance and flight crew coordination situations
“Machine”	Machine operating skills and correct procedures	Choose and carry out tasks related to a washing machine’s setup process

For all of the solutions, the training tasks complied with the relevant learning objectives. In the case of “Driving,” the learner followed directions from a satellite navigation display and practiced their driving skills in a driving simulation. In “Anatomy,” the learner picked up frog parts and received information about them. Additionally, the learner could request more information about the parts using the simulation’s OER search. In “Printer,” the learner set up a 3D printer by following the tutorial’s instructions. In “Blood,” the learner followed instructions, operated an automaton, and prepared reagents for a blood analysis process. In “Steve,” the learner practiced naval equipment operating procedures under the tutelage of an embodied virtual tutor. In “Osprey” a, the learner’s aircraft maintenance skills were put to the test as they navigated around an aircraft and performed maintenance and diagnostic tasks. In “Osprey” b, the learner operated a flight desk and carried out flight crew management tasks in the company of a “Virtual Pilot” (an NPC). In “Machine,” the learner trained to perform laundry service activities that increased in complexity. Overall, the training tasks more or less expected the learners to pay attention to pre-defined instructions, had them move or drive around a VR world, and prompted interaction with the surrounding virtual objects, tools, and mechanisms.

VR Hardware and Worlds

Each solution offered some type of information regarding the applicable VR hardware and the VR world. “Driving,” “Anatomy,” “Printer,” and “Osprey” applied the latest generation of immersive VR technology. All of them used HTC Vive devices and base station trackers. Moreover, the “Driving” solution’s hardware included a steering wheel, driver’s seat, gearbox, and pedals, whereas the other solutions used the default motion controllers. The “Steve” solution’s hardware was the only so-called legacy VR, that is, HMDs with mics, data gloves, position sensors on the head and hands, and 3D mice. “Blood” and “Machine” did not specify which VR devices were used, although vague mention was made of the possibility of using immersive VR peripherals. The “Machine” solution’s environment could also be accessed using a desktop PC or a multi-user, room-sized VR system known as CAVE™ (Cruz-Neira et al., 1993).

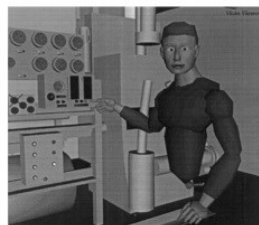
All of the solutions’ VR worlds were 3D and portrayed solely from the first-person perspective (see Figure 2). The “Anatomy” solution’s world was the only abstract virtual space, while the “Driving” solution’s driving simulation was set in a mock-up of a lifelike city and inhabited by AI controlled cars. The rest of the VR worlds were mock-ups of real-life working environments featuring realistic virtual objects and tools.



“Driving” from Ropelato et al. (2018)



“Printer” from Mogessic et al. (2020)



“Steve” from Rickel & Johnson (1999)



MAEVIF application from de Antonio et al. (2005)



“Anatomy” from Ahn et al. (2018)



“Blood” from Taoum et al. (2016)



“Osprey” from Buck et al. (2018)

Figure 2. Virtual Reality Worlds of the Reviewed Solutions

In addition, the “Osprey” b scenario’s VR world was inhabited by the “Virtual Pilot”. Moreover, the “Blood,” “Steve,” and “Machine” solutions’ VR worlds included an embodied humanoid virtual tutor avatar, while the “Anatomy” solution’s tutor avatar resembled a floating robot. According to the description offered by Sánchez and Imbert (2007a), the “Machine” solution’s VR world was presumably a room-like space with a virtual washing machine and other laundry-service related objects. Furthermore, the VR world’s display fidelity was likely quite low given the image we retrieved from the report by de Antonio et al. (2005, p. 50) of a VR world produced using the same software.

Tutoring Systems

As shown in Table 6, each solution featured a unique tutoring system. The purposes of these tutoring systems varied from one solution to the next. Some resembled tutorials (“Anatomy” and “Printer”), some focused on evaluative practical tests (“Driving” and “Osprey” a), and some attempted to interact with the learner (“Blood,” “Steve,” “Osprey” b, and “Machine”). Two distinct forms of tutors stood out, namely embedded and embodied tutors. The “Driving” and “Printer” solutions’ intelligent tutoring was primarily embedded within the respective I-VRLEs, that is, the tutor was not represented by a virtual character and guided the learner by augmenting the I-VRLE with necessary information. In “Driving,” the tutor guided the learner via the virtual car’s satellite navigation system display, while in “Printer,” instructions and hints were readable from a virtual tablet and relevant parts were highlighted in the I-VRLE. In “Anatomy,” “Blood,” “Steve,” and “Machine,” intelligent tutoring was mostly delivered to the learner by an embodied agent, that is, the tutor was represented in the I-VRLE by a distinctive virtual character. The “Blood” and “Steve” solutions’ tutor avatars were made to look like humans, whereas the “Anatomy” solution’s avatar resembled a futuristic robot. The “Machine” solution’s virtual tutor most likely had an embodied human-like frame, albeit with low fidelity (see Figure 2).

In the case of “Osprey,” the intelligent tutoring appeared to have been influenced by both embedded and embodied forms. More specifically, in “Osprey” a, the tutoring system evaluated the learner’s performance in the background, although based on the excerpted image (see Figure 2), its messages were relayed to the learner either via a virtual notifications board featuring a picture of a human-like face or through augmented information in the I-VRLE. In addition, in “Osprey” b, the tutor exhibited various embedded information augmentation capabilities, but at the same time, the authors implied that the “Virtual Pilot” had an embodied form in terms of using gestures and expressing behavioural intentions, the incorrectness of which could possibly be recognized by the learner.

The solutions also differed with regard to how directly the learner could interact with the tutoring solution itself. In “Printer,” “Blood,” “Steve,” “Osprey” a, and “Machine,” the learner was able to interact directly with the tutor system, whereas in “Anatomy” and “Osprey” b, the learner could only indirectly prompt the tutor’s actions by interacting with objects and mechanisms in the I-VRLE. Yet, in the case of the latter solution, the learner could also engage in a direct reciprocal relationship with a synthetic teammate, which could be considered a representation of the tutoring system. In “Driving,” the learner had no direct channel of communication with the

underlying tutoring system. The learner’s evolving driving skills were constantly evaluated, although the authors did not specify how reckless driving would impact either the simulation or the tutoring.

Table 6. Overview of the Proposed Intelligent Tutoring Systems

Solution	Overview of the tutor	The tutor’s form	Interaction with the tutor
“Driving”	Tutor is an algorithm that determines the optimal path during the simulation and operates in the background	A satellite navigation system display	Learner follows the satellite navigation display’s directions and distances
“Anatomy”	Tutor system consists of prescribed text-to-speech (TTS) instructions regarding virtual parts and enabled open-ended research (OER) searching	A robotic virtual avatar and a virtual video display	Learner interacts with virtual parts and listens to the tutor’s instructions and explanations; learner may request more information using the OER search
“Printer”	The application itself is a tutorial wherein the tutor’s role is to determine step correctness and monitor completion	A virtual tablet with hints and instructions in text form and information augmentation available	Learner uses the tablet to follow their progress, read instructions, and request hints
“Blood”	The tutor communicates with the learner via verbal and non-verbal behaviours and initiates pedagogical actions in response to the learner’s progress and behaviour	An embodied humanoid virtual avatar	Learner can choose to interact with the learning content in the I-VRLE or directly with the embodied tutor
“Steve”	The tutor monitors the state of the world and of the learner, responds to them, and makes changes to the I-VRLE	An embodied humanoid virtual avatar	Learner may ask questions of the tutor and observe as it demonstrates tasks
“Osprey”	(a) The “Virtual Instructor” assesses learner performance in real time and adapts the learning scenario (b) The “Virtual Reality Instructor” is a version of the above that guides the learner and compares their performance to pre-defined performance measures	(a) Information augmentation and a floating board (b) Information augmentation and a synthetic teammate	(a) Learner may request help, which impacts their score (b) Learner interacts with the I-VRLE and the tutor adapts the scenario in real time
“Machine”	The virtual tutor is the representational part of a multi-agent system that builds activity plans, monitors and controls the simulation, and monitors and communicates with the learner	An embodied virtual avatar	Learner may ask questions of the tutor and observe its clues, warnings, and descriptions of the syllabus items (course, phase, activities, main goals, and procedures)

Perception

As can be seen from Table 7, all of the solutions monitored the learner's actions ("Anatomy" and "Steve"), the learner's progress through the training ("Driving"), or both (all of the other solutions). This included but was not limited to actions such as selecting virtual objects ("Anatomy," "Blood," and "Machine"), interacting with the mechanisms of a virtual flight deck ("Osprey" b), operating virtual touch-screen monitors ("Printer"), viewing objects and places ("Anatomy"), and carrying objects from one place to another ("Machine"). The progress monitoring involved assessing the learner's performance during the training by monitoring one or more of the following: the correctness of the attempted actions, errors made, assistance consultation frequency, or progress towards the predefined learning objectives. Additionally, the authors of "Osprey" characterized the context of their scenarios as open-ended problems that could be solved in numerous ways. They utilized the behaviour tree technique to interpret the likely correctness of the learner's actions, whereas the other solutions monitored the learner's progress in relation to a pre-defined optimal or expert path.

Table 7. Perception Capabilities of the Intelligent Tutoring Systems

Perceptive capability	Solutions with the perceptive capability						
Progress monitoring	1	3	4	6a	6b	7	
Learner action monitoring		2	3	4	5	6a	6b 7
Head tracking	1	2	4			6b	
Locomotion tracking	1			5			7
Questions			4	5			7
Field of view monitoring				5			7
I-VRLE state monitoring				5		6b	
Non-verbal signals			4			6b	
Requests for help		3			6a		
Requests for more information		2					

Notes. In Solutions 3 ("Printer") and 6 ("Osprey" a), only the perceptual capabilities of the VR-based ITSs were considered.

Four solutions utilized head tracking data. For instance, in the "Driving" solution, the learner's head rotation while driving influenced the scoring related to making turns, while in "Anatomy," head tracking data were used to determine whether the learner was looking at the correct object. Furthermore, head tracking data were used to evaluate the learner's performance in relation to pre-defined performance measures ("Osprey" b) or to monitor and interpret the learner's non-verbal communication ("Blood"). Moreover, in "Blood," the learner's facial expressions and voice intonation were monitored for the same purpose.

In total, six solutions were responsive to the learner's requests. For example, in "Printer" and "Osprey" a, the learner could request help. In the case of the former, the learner could decide to turn hints on or off using the virtual tablet, while in the latter, requesting help impacted the final evaluative score. In "Blood," "Steve," and "Machine," the learner could pose questions regarding the simulations' goals, virtual objects, or actions. In

“Anatomy,” the learner could initiate an expanded query concerning a selected frog part in order to gain more information (e.g., from YouTube, Wikipedia, etc.) using an OER search based on a self-made hierarchical ontology of frog anatomy.

The tutoring systems included in three of the solutions tracked the learner’s locomotion. The “Driving” solution’s tutor environment used locomotion data to monitor the virtual car’s route, position, and location within the simulation. The “Steve” solution’s tutor agent used such data to determine where the learner was situated in the I-VRLE. The “Machine” solution’s tutor agent was capable of tracking the learner’s itineraries and comparing them with the optimal path plan. The tutoring systems in the “Steve” and “Osprey” b solutions monitored the state of the I-VRLE, that is, changes in the state of the simulation and its objects. Finally, in the case of the older solutions (“Steve” and “Machine”), the tutoring systems were capable of monitoring the learner’s field of view.

Guidance Actions

As shown by Table 8, the most common forms of guidance actions provided to the learner by the tutoring systems were hints and clues. The other commonly included actions were action demonstrating, guiding attention, answering questions, and providing feedback. The unique guidance actions included describing objects’ properties, providing directions, interjecting events into the I-VRLE, modifying the task difficulty, and offering warnings.

Table 8. Guidance Actions of the Intelligent Tutoring Systems

Guidance actions	Solutions with the guidance action				
Hints and clues	3	4		6a	7
Demonstrations	2*	4	5		7**
Attention guidance		4	5	6b	
Answer questions		4	5		7
Feedback		3	4	6b	
Descriptions	2				
Directions	1				
Event interjection				6b	
Modify task difficulty	1				
Warnings					7

Notes. For Solutions 3 (“Printer”) and 6 (“Osprey” a), only the guidance actions of the VR part’s ITS have been included. *Solution 2 (“Anatomy”) included a virtual monitor that displayed video material discovered via the OER search. **Solution 7’s (“Machine”) virtual tutor avatar demonstrated actions as a detailed hint.

In “Printer” and “Osprey” a, the tutoring systems offered hints on an on-demand basis. More specifically, the “Printer” solution’s tutoring system displayed an instructional message concerning the next correct action and highlighted the appropriate objects, mechanisms, and directions. Similarly, the “Osprey” a scenario’s tutoring

system helped the learner to locate the necessary objects when they requested help. The specificity of the information varied based on the frequencies of the requests and errors made in the latter scenario. Otherwise, the solution focused more on evaluating than guiding the learner. The “Blood” solution’s tutoring system highlighted objects and provided explanations of those objects and the related training goals in response to negative expressions on the part of the learner. In “Machine,” the tutor agent was able to offer hints with varying levels of detail when the learner appeared to be confused, looked or moved away from the target, or had not attempted an action for a while. The most detailed level of hints involved demonstrations of the next correct action. Likewise, in “Steve,” the embodied tutor avatar could perform demonstrations of the expected actions. In “Blood,” the embodied tutor avatar was stationed between the automaton (an automated blood sample analyser machine) and the desk for preparing the reagents (substances in vials and test tubes). However, the authors claimed that the demonstrations could be viewed as animations in the I-VRLE. Finally, “Anatomy” included a virtual video display in the I-VRLE through which demonstrations and other video material discovered using the OER search could be viewed.

The tutor agents in “Blood,” “Steve,” and “Machine” were designed to answer the learner’s questions about the virtual objects, training goals, purposes of actions, or next expected actions. More specifically, in “Blood,” the embodied tutor avatar guided the learner’s attention via verbal signals and by pointing at objects. Similarly, the “Steve” solution’s embodied tutor avatar used gaze and gestures to direct the learner’s attention towards relevant objects. By contrast, in “Osprey” b, the learner received tactile feedback, such as the vibration of the controller, intended to guide their attention towards relevant instruments.

Three tutoring systems provided feedback to the learner within the I-VRLE. For instance, in “Printer,” the feedback was intended to let the learner know if they were following the correct procedural steps. The “Blood” solution’s embodied tutor avatar could respond to the learner’s mistakes by providing negative feedback through verbal signals and facial expressions. The “Osprey” b scenario’s tutoring system provided feedback through verbal or textual comments or by highlighting areas of the I-VRLE.

The unique guidance actions available in the “Driving” solution’s driving simulation included providing the learner with directions based on their location in relation to the next event in the virtual city and adjusting the difficulty of such events as the learner progressed through the simulation. The “Anatomy” solution’s robotic virtual avatar provided informative descriptions of virtual objects after the learner had selected them in the I-VRLE. The “Osprey” b scenario’s tutoring system would interject events into the I-VRLE when the learner did not appear to be paying attention. For example, it would make the “Virtual Pilot” perform a task incorrectly to provoke a reaction on the part of the learner. Finally, the “Machine” solution’s virtual tutor avatar would warn the learner if they failed to follow the optimal route to the required destination.

Cognition

The cognitive features of the reviewed solutions, as presented in Table 9, spanned three of the four major components considered typical of an ITS (see, e.g., Pavlik et al., 2013), that is, the domain model, the student

model, and the pedagogical model. The domain model comprises a set of skills, knowledge, and strategies concerning the topic being tutored. The student model consists of the cognitive, affective, motivational, and other psychological states inferred based on the learner’s performance during the training. The pedagogical model takes the domain and student models as inputs and selects strategies, steps, and actions regarding what the tutor should do next (Pavlik et al., 2013). The fourth major component, namely the tutor-student model or user interface, was implemented through the perception and guidance actions aspects.

Table 9. Cognitive Features of the Reviewed Solutions

Feature	“Driving”	“Anatomy”	“Printer”	“Blood”	“Steve”	“Osprey”	“Machine”
<i>Framework</i>	-	-	CTAT	MASCARET	Soar	-	MAEVIF
<i>Domain model</i>	Activities graph	Branched programming	Behaviour graph	UML ontology	Procedural network	Behaviour graph	Syllabus*
<i>Student model</i>	ZPD	-	KC mastery probability	Memory model	Hierarchical plan state		Student modelling agent*
<i>Pedagogical model</i>	ZPDES algorithm	Fixed user-triggered actions	Example tracing BKT	UML Pedagogical scenarios	Simplified partial order planner	Example tracing	Tutoring agent*
<i>NLU/ Dialogue</i>	-	-	-	AIML/FIPA messages	Keywords Text fragments		Question types*

Notes. A dash (-) indicates that the feature was not present. An empty value indicates that information concerning the feature was not available. *Architectural concept. NLU = natural language understanding; ZPD = zone of proximal development; ZPDES = zone of proximal development and empirical success; CTAT = cognitive tutor authoring tools; KC = knowledge component; BKT = Bayesian knowledge tracing; MASCARET = Multi-Agent System for Collaborative, Adaptive & Realistic Environments for Training; UML = unified modelling language; AIML = artificial intelligence markup language; AIML = artificial intelligence markup language; FIPA = the foundation for intelligent physical agents; MAEVIF = model for the application of intelligent virtual environments to education.

Framework and Domain Model

Among the seven solutions, four were based on an existing cognitive or ITS framework. The applied framework determined each solution’s architecture and features, including the domain model and how it was authored. “Printer” utilized an external ITS service based on the cognitive tutor authoring tools (CTAT) toolset (Aleven et al., 2015), which was consulted to determine the correctness of the learner’s actions. By using the CTAT, an author can build a domain model by recording all of the foreseen solution paths that the learner could take as a behaviour graph. Furthermore, the author can annotate the graph to make it apply to variations in the path, in addition to being able to add incorrect solution paths. “Blood” was implemented using MASCARET (Buche et al., 2003), a framework for modelling the ontology of a domain using an extension of the unified modelling

language (UML), which allows for the use of widely available UML modelling tools and enables the division of modelling tasks between domain experts and pedagogical experts (Chevaillier et al., 2012). The “Steve” solution differed from the other solutions due to not building its domain model on an ITS framework. Instead, it utilized the generic Soar cognitive architecture (Laird et al., 1987), which was reflected in the tutoring solution’s name (Soar training expert for virtual environments = Steve). While building Soar-based intelligent agents normally requires the costly authoring of production rules, “Steve” allowed the author to build the domain model using a procedural network consisting of hierarchical task plans, which were automatically converted into Soar production rules (Rickel & Johnson, 1999). In the case of “Machine,” the authors restricted their description of the solution to its abstract architecture. The tutoring solution was based on MAEVIF (de Antonio et al., 2005), an architectural framework centred on the principle of implementing the functional parts of an ITS as software components known as “agents.” The authors substituted the basic tutoring module in MAEVIF with their own tutoring module and described the domain model in their solution as a “syllabus,” a hierarchical structure specifying the courses, phases, actions, activities, and objectives of the solution.

Among the solutions that were not based on a framework, “Driving” authored the domain model as an activity graph reflecting the possible transitions from one exercise to the next. In the case of “Osprey,” the authors offered no description of their domain model apart from stating that it was a behaviour graph. “Anatomy” did not model the domain using an overlay model (Pavlik et al., 2013), instead providing fixed branched programming actions triggered by user interactions.

Pedagogical and Student Models

The pedagogical model of “Anatomy” relied on simple branched programming, executing fixed triggers such as checking where the learner was looking before proceeding to the next action. The solution had no student model, which precluded individual adaptation. “Printer” based its pedagogical model on example tracing (Aleven et al., 2009), whereby the system matched the learner’s solution path to multiple parallel candidate paths authored in the domain model. To guide the outer loop of the pedagogical model (i.e., problem selection), the system maintained a student model that calculated the probability that the student had mastered each knowledge component (KC) defined by the author. In this context, KCs represent the smallest units into which the knowledge to be learned can be decomposed (Aleven & Koedinger, 2013). The probabilities were updated by means of Bayesian knowledge tracing (BKT) as the learner traversed links in the domain model (Corbett & Anderson, 1994). Again, the authors of “Osprey” did not describe their pedagogical model in detail, although their description resembled an example-tracing mechanism.

In its pedagogical model, “Driving” considered the problem of selecting the best sequence of driving exercises in order to achieve the best learning outcomes. The authors described their model as an application of the ZPDES algorithm (Clement et al., 2014), which relies on a reinforcement learning (RL) (Sutton & Barto, 2018) algorithm to find the learner-specific zone of proximal development (ZPD) by selecting exercises that are not too easy while also maximizing the overall success rates of the exercises. This was the only reviewed tutoring solution for which the authors described the application of a machine-learning algorithm. In “Blood,” the ITS

loops executed pedagogical scenarios that made use of a library of actions intended to guide or correct the learner, which the authors referred to as the “pedagogue.” These constructs were essentially rules that could consider the learner and the world state, as authored in UML by a pedagogical expert (Buche et al., 2010). The authors described a student model imitating human memory, wherein model concepts (entities, actions) representing what the learner is instructed to do were placed in the “working memory.” The working memory represented what the learner was expected to work on. If the learner performed an action that was inconsistent with the contents of the working memory, the system performed guidance actions.

In “Machine,” the authors described an abstract “tutoring agent” module that executed ITS loops over the “syllabus” in coordination with a “student modelling agent,” which was a module that maintained the student model. The implementation of this architecture in the pilot system was not described. Among the reviewed solutions, “Steve” was associated with the most complex and most flexible pedagogical model. During every decision cycle, “Steve” consulted the underlying Soar truth management system, which provided information about which goals of the procedural task model (represented in Soar as production rules) were currently satisfied and used that information to mark parts of the task model as relevant. The information regarding the relevant tasks constrained the planning algorithm and greatly reduced the search space for planning the next action. In addition to guiding the learner in the next step, the current plan could guide the avatar in performing an action sequence involving demonstrating the task. The supporting student model consisted of the current state of the production rules representing the task plan as well as an episodic memory that recorded the state of the world when each step was executed (Rickel & Johnson, 1999). The history of the world state allowed the system to respond to unexpected changes in the environment, for example, to re-execute parts of the plan that were undone, skip over parts that had had their goals serendipitously achieved, or present cue phrases to indicate the relation between the current and previous steps. In addition, the causal links in the procedural task model allowed “Steve” to answer the learner’s queries about “why” something should be done, and combined with the additional “debrief” rules, to answer another or repeated query regarding a demonstration it has already performed.

Natural Language Understanding and Dialogue

Among the reviewed tutoring solutions, “Blood,” “Steve,” and “Machine” all mentioned the cognitive capability to engage in a question-answer dialogue with the learner. Again, “Machine” only prescribed an architectural taxonomy of basic to expert questions that the learner could answer, which meant that it did not describe a mechanism for handling the required cognitive processing of the dialogue. The natural language understanding (NLU) functionality in “Blood” considered an I-VRLE architecture wherein the ITS functionality and the VR environment were implemented in separate components. The authors described an NLU interface between these components based on the formal FIPA-SL language (Poslad, 2007). To recognize the semantic content of learner utterances, the system armed the VR environment with authored artificial intelligence markup language (AIML) templates that mapped incoming speech patterns to their semantic meanings. The VR environment matched incoming speech to the templates and, if a match was found, communicated the semantic meaning of the speech utterance to the ITS using FIPA-SL. The semantic content of a recognized utterance fired rules in the

UML pedagogical scenarios. If such a rule triggered a speech response, the ITS communicated the semantic content of the response as FIPA-SL to the VR environment, which translated the semantic content into the appropriate computer-generated speech. Through this arrangement, the ITS rules could work with clean semantic symbols, while the mapping of those symbols to natural language was left to the VR environment. “Steve” described a similar mechanism involving authored text templates for its “speak” actions. By way of speech input, it only recognized a few keywords (e.g., “Why?”), although it maintained a dialogue context and included actions for listening for an active speaker. In addition, by maintaining a discourse focus stack (see Grosz & Sidner, 1986), “Steve” could resume a dialogue after an interruption (Rickel & Johnson, 1999).

Discussion

The present literature review has revealed that ITS and I-VR technologies have only rarely been applied together in the context of hard skills learning over time and in different domains. The learning objectives within I-VRLEs have generally focused on helping learners to attain procedural knowledge. Learning objectives concerning declarative knowledge were exclusively realized in “Anatomy” and in blended learning lessons (“Printer” and “Osprey” a) before the learner used the I-VR. However, various information augmenting methods were utilized by most of the tutoring systems to display descriptions, explanations, and demonstrations as guiding actions within the I-VRLEs. As such, learning objectives related to declarative knowledge also seem attainable within the studied I-VRLEs.

In this review, roughly half of the tutoring systems were embedded in the environment, while half employed a virtual avatar as the tutor. Presenting the tutor in an embodied form has previously received support due to the persona effect (Lester et al., 1997), that is, that fact that the presence of a life-like avatar can have a strong positive impact on the learning experience. This effect was first discovered among middle school children and later noticed among adults who followed virtual presentations (Van Mulken et al., 1998). To the best of our knowledge, the persona effect has not previously been studied in the context of learning within I-VRLEs.

The majority of the reviewed training solutions resembled tutorials with high regulation and an expected order of actions. One might even ask, where is the pedagogy and how impactful are these types of tutorials? The student models associated with the reviewed tutoring systems concentrated on the learner’s capability to perform the correct actions and remain on the optimal path. However, I-VR technology can enable more detailed tracking capabilities and offer new possibilities for monitoring the learner. The aim when creating a tutoring system should be to develop a system equally capable of providing both cognitive and affective guidance (Lepper et al., 1993). In addition, Roscoe and Chi (2007, p. 552) discovered that guiding learners was more impactful when elements of knowledge-building (e.g., engaging learners in producing knowledge through reflecting on their actions) were incorporated into the tutoring when compared with tutoring based on knowledge-telling elements (e.g., the tutor telling the learner what they are supposed to know or do).

The “Blood” solution’s tutoring system was the only one in which some level of affect monitoring was applied. The tutoring system was designed to infer the learner’s negative expressions based on non-verbal

communication and offer guidance accordingly. Cognitive learning theories have long been favoured over affective theories when it comes to developing tutoring solutions (Woolf et al., 2009, p. 131). Consequently, most of the tutoring systems were concerned with whether or not the learner was aware of the next correct action. Furthermore, we noticed that the training solutions exhibited three different ways of dealing with learner idleness. The first way was to try and prevent it by allowing the learner to request hints at any time. The second way was to interject events into the I-VRLE in an effort to force the learner to react. Finally, the third way was simply to enforce hints when enough time had passed since the learner last progressed in the training.

The different cognitive architectures underlying the intelligent behaviour of the reviewed solutions represented different implementation traditions. Several of the solutions extended existing ITS frameworks, retaining a student model that was originally developed to support web-based training. Such frameworks may severely constrain the tutoring capabilities of an I-VRLE, as their domain modelling approach would become excessively costly when the learner action space expanded from simple web-based exercises to long, stateful procedural sequences performed in a simulated environment. By contrast, the generic cognitive approaches exemplified by “Steve” can implement many intelligent behaviours that are useful in terms of supporting procedural learning in such complex settings. However, none of the reviewed solutions attempted to utilize more recent generic cognitive architectures, such as ACT-R or Clarion (for a review, see Bach, 2009). Even “Anatomy,” which was based on the CTAT toolset, chose example tracing over the ACT-R-based cognitive tutoring functionality in the toolset. One possible explanation for this is that the cost of authoring a detailed domain model for a generic cognitive architecture remains too high when using current tools.

Although a natural language speech dialog is especially important in relation to an I-VRLE with limited input affordances, the natural language understanding and generation capabilities of the reviewed systems were clearly very limited, with none of the systems attempting an auto-tutor-like tutoring dialogue (Graesser et al., 1999). “Driving” employed a control loop capable of online learning based on the user’s behaviour. The authors also mentioned the possibility of pre-training their model using data collected from user sessions. Interestingly, none of the other solutions employed either machine learning or connectionist approaches, instead relying on old fashioned AI (McDermott & O’Reilly, 2015). The future incorporation of generic cognitive architectures, affective learning models using machine-learning-based perception, and state-of-the-art NLU may require revisiting the discussion concerning what should be considered the minimum level for an ITS in this context. Indeed, earlier definitions (e.g., Alevan et al., 2015, p. 235; VanLehn, 2006) may not adequately describe the capabilities expected from a tutor intended to support hard skills learning in an I-VRLE.

Finally, we acknowledge that only a small number of solutions met our inclusion criteria. In part, we consider this to be a fundamental issue associated with the literature review method, as authors may not include canonical terms within their keywords and abstracts. The majority of prior studies concerning ITSs have focused on the education (formal schooling) domain, which we deliberately excluded due to our focus on the professional learning context. In addition, the latest work conducted on the new and inexpensive class of VR hardware that has been available since 2016 may not yet feature in the literature, or despite the work exhibiting many ITS-like behaviours, may not be considered to incorporate a formal ITS element.

Conclusion

We have reviewed papers that described training solutions in which ITS technology was implemented in order to guide hard skills acquisition in an I-VRLE. While there is a long tradition of ITS research, its implementation in I-VRLEs remains a novel research branch. However, separate from this tradition, it is possible that some researchers have developed other adaptive tutoring solutions for I-VRLEs. The scope of our study only included papers featuring ITSs.

This review indicated that the described hard skills training using VR was exclusively completed either in realistic mock-ups of the learner's real-life environment or using full mock-ups of the work equipment to be trained. A recent meta-analysis (Kaplan et al., 2021, p. 10) argued that digital learning environments are generally made to resemble real-life working contexts in an effort to minimize the negative effects on performance transfer. The meta-analysis revealed that, to date, training involving extended reality technologies, such as I-VR, has matched up with training in traditional settings. Yet, none of the papers in this review included actual empirical studies comparing the effectiveness of their I-VRLE/ITS or examining the impact on learning results such as performance transfer.

Overall, it appears that I-VR technology will be increasingly used to provide training through emulations of realistic situations. However, we should not ignore the fact that in its current phase of development, I-VR technology may increase inequality in relation to education. Indeed, not all learners can afford it, while some learners may not be able to use it. Yet, I-VR technology could help with rehabilitation, and it could also be harnessed to render certain practical training and manufacturing processes more sustainable. Ultimately, I-VR is a tool that needs to be combined with various technologies, and we can still influence its development and use.

Recommendations

Based on the findings of this review, we suggest that future studies involving I-VRLEs should: 1) clearly state which I-VR devices were applied, 2) offer a detailed description of the proposed virtual learning environments, 3) specify which learning theories influenced the design and use of the I-VR system, and 4) empirically study the learning results of differently equipped I-VRLEs and compare their effectiveness to other digital tools. Moreover, if an ITS of any kind is developed and applied, researchers should: 5) provide detailed descriptions of its perceptual, cognitive, and guidance capabilities, and 6) study and compare the impacts of the ITS and its features (such as embodied and embedded tutor forms) on learning processes and results. This would enable researchers to perform detailed reviews and meta-analyses as well as to offer design suggestions based on empirical research.

Acknowledgements

This research was supported by Business Finland (Co-innovation project AI in Learning). The authors report there are no competing interests to declare.

References

- Ahn, J. W., Tejwani, R., Sundararajan, S., Sipolins, A., O'Hara, S., Paul, A., Kokku, R., Kjallstron, J., Dang, N. H., & Huang, Y. (2018). Intelligent virtual reality tutoring system supporting open educational resource access. In R. Nkambou, R. Azevedo, & J. Vassileva (Eds.), *International conference on intelligent tutoring systems* (Vol. 14, pp. 280-286). Springer. https://doi.org/10.1007/978-3-319-91464-0_28
- Alcañiz, M., Parra, E., & Chicchi Giglioli, I. A. (2018). Virtual reality as an emerging methodology for leadership assessment and training. *Frontiers in Psychology, 9*, 1658. <https://doi.org/10.3389/fpsyg.2018.01658>
- Aleven, V., & Koedinger, K. R. (2013). Knowledge component (KC) approaches to learner modeling. In R. A. Sottolare, A. Graesser, X. Hu, & H. Holden (Eds.), *Design recommendations for intelligent tutoring systems* (Vol. 1, pp. 165-182). U.S. Army Research Laboratory. <https://skillsmarket.net/wp-content/uploads/resources/DesignRecommendations.pdf#page=188>
- Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2009). A new paradigm for intelligent tutoring systems: Example-tracing tutors. *International Journal of Artificial Intelligence in Education, 19*(2), 105-154. https://doi.org/10.1007/11774303_7
- Aleven, V., Sewall, J., Popescu, O., van Velsen, M., Demi, S., & Leber, B. (2015). Reflecting on twelve years of ITS authoring tools research with CTAT. In R. A. Sottolare, A. Graesser, X. Hu, & K. Brawner (Eds.), *Design recommendations for adaptive intelligent tutoring systems* (Vol. 3, pp. 263-283). U.S. Army Research Laboratory. https://www.researchgate.net/profile/Robert-Sottolare/publication/278620644_Design_Recommendations_for_Intelligent_Tutoring_Systems_-_Volume_3_Authoring_Tools_and_Expert_Modeling_Techniques/links/558176f908ae607ddc33c9f6/Design-Recommendations-for-Intelligent-Tutoring-Systems-Volume-3-Authoring-Tools-and-Expert-Modeling-Techniques.pdf#page=285
- Amokrane, K., Lourdeaux, D., Barthès, J. P., & Burkhardt, J. M. (2008). An intelligent tutoring system for training and learning in a virtual environment for high-risk sites. In *2008 20th IEEE international conference on tools with artificial intelligence* (Vol. 2, pp. 185-193). IEEE. <https://doi.org/10.1109/ictai.2008.151>
- Bach, J. (2009). *Principles of synthetic intelligence PSI: An architecture of motivated cognition* (Vol. 4). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195370676.001.0001>
- Buche, C., Bossard, C., Querrec, R., & Chevaillier, P. (2010). PEGASE: A generic and adaptable intelligent system for virtual reality learning environments. *International Journal of Virtual Reality, 9*(2), 73-85. <https://doi.org/10.20870/ijvr.2010.9.2.2772>
- Buche, C., Querrec, R., Maffre, E., Chevaillier, P., & De Loor, P. (2003). MASCARET: Multiagent system for virtual environment for training. In *VRIC 03* (pp. 159-164). <https://hal.archives-ouvertes.fr/hal-00611187/>
- Buck, B., Genova, M., Dargue, B., & Biddle, E. (2018). Adaptive learning capability: User-centered learning at the next level. In R. A. Sottolare (Ed.), *Proceedings of the ITS 2018 industry track co-located with the 14th international conference on intelligent tutoring systems (ITS 2018)* (pp. 3-10). U.S. Army Research Laboratory. <http://ceur-ws.org/Vol-2121/>

- Buck, B., Genova, M., Sottolare, R. A., & Goldberg, B. (2017). *A blended approach to adaptive learning*. The Boeing Company and U.S. Army Research Laboratory. https://www.gifttutoring.org/attachments/download/2687/05_GIFTSym6_Inst%20Mgt_paper_27.pdf
- Chevallier, P., Trinh, T. H., Barange, M., De Loor, P., Devillers, F., Soler, J., & Querrec, R. (2012). Semantic modeling of virtual environments using MASCARET. In *2012 5th workshop on software engineering and architectures for realtime interactive systems (SEARIS)* (pp. 1-8). IEEE. <https://doi.org/10.1109/searis.2012.6231174>
- Clement, B., Roy, D., Oudeyer, P. Y., & Lopes, M. (2014, October). Developmental learning for intelligent tutoring systems. In *4th International Conference on Development and Learning and on Epigenetic Robotics* (pp. 426-433). IEEE. <https://doi.org/10.1109/devlrm.2014.6983019>
- Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, 4(4), 253-278. <https://doi.org/10.1007/BF01099821>
- Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, 19(2), 272-309. <https://doi.org/10.1080/15213269.2015.1015740>
- Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE. In *Proceedings of the 20th annual conference on computer graphics and interactive techniques* (pp. 135-142). Association for Computing Machinery. <https://doi.org/10.1145/166117.166134>
- de Antonio, A., Ramírez, J., Imbert, R., Méndez, G., & Aguilar, R. A. (2005). A software architecture for intelligent virtual environments applied to education. *Revista Facultad de Ingeniería-Universidad de Tarapacá*, 13(1), 47-55. <https://doi.org/10.4067/s0718-13372005000100006>
- Graesser, A. C., Wiemer, K., Wiemer-Hastings, P. M., & Kreuz, R. J. (1999). AutoTutor: A simulation of a human tutor. *Cognitive Systems Research*, 1, 35-51.
- Grosz, B., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3), 175-204. <https://dash.harvard.edu/handle/1/2579648>
- Han, J., Zhao, W., Jiang, Q., Oubibi, M., & Hu, X. (2019). Intelligent tutoring system trends 2006-2018: A literature review. In *2019 eighth international conference on educational innovation through technology (EITT)* (pp. 153-159). IEEE. <https://doi.org/10.1109/eitt.2019.00037>
- Imbert, R., Sánchez, L., de Antonio, A., Méndez, G., & Ramírez, J. (2007). A multiagent extension for virtual reality based intelligent tutoring systems. In *Seventh IEEE international conference on advanced learning technologies (ICALT 2007)* (pp. 82-84). IEEE. <https://doi.org/10.1109/icalt.2007.23>
- Kaplan, A. D., Cruit, J., Endsley, M., Beers, S. M., Sawyer, B. D., & Hancock, P. A. (2021). The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis. *Human Factors*, 63(4), 706-726. <https://doi.org/10.1177/0018720820904229>
- Kardong-Edgren, S. S., Farra, S. L., Alinier, G., & Young, H. M. (2019). A call to unify definitions of virtual reality. *Clinical Simulation in Nursing*, 31, 28-34. <https://doi.org/10.1016/j.ecns.2019.02.006>
- Kulik, J. A., & 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>


- Kurniawan, C., Rosmansyah, Y., & Dabarsyah, B. (2019). A systematic literature review on virtual reality for learning. In *2019 IEEE 5th international conference on wireless and telematics (ICWT)* (pp. 1-4). IEEE. <https://doi.org/10.1109/icwt47785.2019.8978263>
- Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence*, 33(1), 1-64. <https://doi.org/10.21236/ada188742>
- Laker, D. R., & Powell, J. L. (2011). The differences between hard and soft skills and their relative impact on training transfer. *Human Resource Development Quarterly*, 22(1), 111-122.
- Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. Lajoie & S. Derry (Eds.), *Computers as cognitive tools* (pp. 75-105). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203052594-8>
- Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997). The persona effect: Affective impact of animated pedagogical agents. In *Proceedings of the ACM SIGCHI conference on human factors in computing systems* (pp. 359-366). <https://doi.org/10.1145/258549.258797>
- McDermott, J., & O'Reilly, U. M. (2015). Genetic programming. In J. Kacprzyk & W. Pedrycz (Eds.), *Springer handbook of computational intelligence* (pp. 845-869). Springer. https://doi.org/10.1007/978-3-662-43505-2_43
- Mihas, P. (2019). Qualitative data analysis. In G. W. Noblit (Ed.), *Oxford research encyclopedia of education*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190264093.013.1195>
- Mogessie, M., Wolf, S. D., Barbosa, M., Jones, N., & McLaren, B. M. (2020). Work-in-progress: A generalizable virtual reality training and intelligent tutor for additive manufacturing. In *2020 6th international conference of the immersive learning research network (iLRN)* (pp. 355-358). IEEE. <https://doi.org/10.23919/ilrn47897.2020.9155119>
- Mousavinasab, E., Zarifsanaiey, N. R., Niakan Kalhori, S., Rakhshan, M., Keikha, L., & Ghazi Saeedi, M. (2021). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, 29(1), 142-163.
- Nabizadeh, A. H., Leal, J. P., Rafsanjani, H. N., & Shah, R. R. (2020). Learning path personalization and recommendation methods: A survey of the state-of-the-art. *Expert Systems with Applications*, 159, 113596. <https://doi.org/10.1080/10494820.2018.1558257>
- Neagu, L. M., Rigaud, E., Travadel, S., Dascalu, M., & Rughinis, R. V. (2020). Intelligent tutoring systems for psychomotor training: A systematic literature review. In V. Kumar & C. Troussas (Eds.), *International conference on intelligent tutoring systems* (pp. 335-341). Springer. https://doi.org/10.1007/978-3-030-49663-0_40
- Pavlik Jr, P. I., Brawner, K., Olney, A., & Mitrovic, A. (2013). A review of student models used in intelligent tutoring systems. In R. A. Sottolare, A. Graesser, X. Hu, & H. Holden (Eds.), *Design recommendations for intelligent tutoring systems: Learner modeling* (Vol. 1, pp. 39-67). U.S. Army Research Laboratory. https://blogs.memphis.edu/aolney/files/2019/10/pavlik_studentmodels_2013.pdf
- Petukhov, I., Steshina, L., & Glazyrin, A. (2017). Application of virtual reality technologies in training of man-machine system operators. In *2017 international conference on information science and communications technologies (ICISCT)* (pp. 1-7). IEEE. <https://doi.org/10.1109/icisct.2017.8188569>
- Poslad, S. (2007). Specifying protocols for multi-agent systems interaction. *ACM Transactions on Autonomous*

- and *Adaptive Systems (TAAS)*, 2(4), 1-24. <https://doi.org/10.1145/1293731.1293735>
- Querrec, R., Taoum, J., Nakhhal, B., & Bevacqua, E. (2018). Model for verbal interaction between an embodied tutor and a learner in virtual environments. In *Proceedings of the 18th international conference on intelligent virtual agents* (pp. 197-202). Association for Computing Machinery. <https://doi.org/10.1145/3267851.3267895>
- Rickel, J., & Johnson, W. L. (1998). STEVE (video session): A pedagogical agent for virtual reality. In *Proceedings of the second international conference on autonomous agents* (pp. 332-333). Association for Computing Machinery. <https://doi.org/10.1145/280765.280851>
- Rickel, J., & Johnson, W. L. (1999). Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. *Applied Artificial Intelligence*, 3(4-5), 343-382. <https://doi.org/10.1080/088395199117315>
- Ropelato, S., Zünd, F., Magnenat, S., Menozzi, M., & Sumner, R. (2018). Adaptive tutoring on a virtual reality driving simulator. *International SERIES on information systems and management in creative emedia (CreMedia)*, 2017(2), 12-17. <https://core.ac.uk/download/pdf/228470743.pdf>
- Roscoe, R. D., & Chi, M. T. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, 77(4), 534-574. <https://doi.org/10.3102/0034654307309920>
- Sánchez, L., & Imbert, R. (2007a). An agent-based adaptable and configurable tutoring module for intelligent virtual environments for training. In *International conference on technologies for e-learning and digital entertainment* (pp. 499-510). Springer. https://doi.org/10.1007/978-3-540-73011-8_49
- Sánchez, L., & Imbert, R. (2007b). Effective student assistance in virtual reality based intelligent tutoring systems. In *Seventh IEEE international conference on advanced learning technologies (ICALT 2007)* (pp. 424-426). IEEE. <https://doi.org/10.1109/icalt.2007.136>
- Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6(6), 603-616. <https://doi.org/10.1162/pres.1997.6.6.603>
- Soofi, A. A., & Ahmed, M. U., (2019). A systematic review of domains, techniques, delivery modes and validation methods for intelligent tutoring systems. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 10(3), 99-107. <https://doi.org/10.14569/ijacsa.2019.0100312>
- Steffen, J. H., Gaskin, J. E., Meservy, T. O., Jenkins, J. L., & Wolman, I. (2019). Framework of affordances for virtual reality and augmented reality. *Journal of Management Information Systems*, 36(3), 683-729. <https://doi.org/10.1080/07421222.2019.1628877>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press. https://www.academia.edu/download/54674740/Reinforcement_Learning.pdf
- Taoum, J., Nakhhal, B., Bevacqua, E., & Querrec, R. (2016). A design proposition for interactive virtual tutors in an informed environment. In D. Traum, W. Swartout, P. Khooshabeh, S. Kopp, S. Scherer, & A. Leuski (Eds.), *International conference on intelligent virtual agents* (pp. 341-350). Springer. https://doi.org/10.1007/978-3-319-47665-0_30
- VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education*, 16(3), 227-265.

- Van Mulken, S., Andre, E., & Müller, J. (1998). The persona effect: How substantial is it? In H. Johnson, L. Nigay, & C. Roast (Eds.), *People and computers XIII* (pp. 53-66). Springer. https://doi.org/10.1007/978-1-4471-3605-7_4
- Virvou, M., Manos, K., & Katsionis, G. (2003). An evaluation agent that simulates students' behaviour in intelligent tutoring systems. In *SMC'03 conference proceedings. 2003 IEEE international conference on systems, man and cybernetics. Conference theme - System security and assurance (Cat. No. 03CH37483)* (Vol. 5, pp. 4872-4877). IEEE. <https://doi.org/10.1109/icsmc.2003.1245754>
- Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., Saari, T., Laarni, J., Ravaja, N., Gouveia, F. R., Biocca, F., Sacau, A., Jancke, L., Baumgartner, T., & Jancke, P. (2007). A process model of the formation of spatial presence experiences. *Media Psychology*, 9(3), 493-525. <https://doi.org/10.1080/15213260701283079>
- Woolf, B., Bursleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: Recognising and responding to student affect. *International Journal of Learning Technology*, 4(3-4), 129-164. <https://doi.org/10.1504/ijlt.2009.028804>
- Zahabi, M., & Abdul Razak, A. M. (2020). Adaptive virtual reality-based training: A systematic literature review and framework. *Virtual Reality*, 24(4). <https://doi.org/10.1007/s10055-020-00434-w>

Author Information

Joakim Laine

 <https://orcid.org/0000-0002-3407-6551>

Faculty of Educational Sciences

University of Helsinki


Siltavuorenpenger 5A, PL 9

00014 Helsinki

Finland

Contact e-mail: joakim.x.laine@helsinki.fi

Timo Lindqvist

 <https://orcid.org/0000-0002-9577-6092>


Upknowledge Inc.

Hiomotie 19

00380 Helsinki

Finland

Tiina Korhonen

 <https://orcid.org/0000-0003-2875-4915>

Faculty of Educational Sciences


University of Helsinki

Siltavuorenpenger 5A, PL 9

00014 Helsinki

Finland

Kai Hakkarainen

 <https://orcid.org/0000-0003-3507-7537>

Faculty of Educational Sciences

University of Helsinki

Siltavuorenpenger 5A, PL 9

00014 Helsinki

Finland