Computers & Graphics 102 (2022) 566–591

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Technical Section

# Cognitive load considerations for Augmented Reality in network security training[☆]

Bradley Herbert [a,*], Grant Wigley [a], Barrett Ens [b,a], Mark Billinghurst [c,a]

[a] University of South Australia, Mawson Lakes Boulevard, SA 5095 Mawson Lakes, Australia
[b] Monash University, Wellington Road, Clayton, Clayton Vic 3800, Australia
[c] University of Auckland, Auckland Bioengineering Institute University of Auckland, 70 Symonds Street, Auckland, Auckland 1010, New Zealand

ARTICLE INFO

*Article history:*
Received 24 February 2021
Received in revised form 17 August 2021
Accepted 3 September 2021
Available online 10 September 2021

*Keywords:*
Intelligent Tutoring Systems
Augmented Reality
Cognitive Load Theory
Instructional design
Training and education

ABSTRACT

This paper presents an Augmented Reality (AR)-based network cabling tutoring system that trains users how to interconnect cables within a Virtual Local Area Network (VLAN) on a physical switching rack. AR arrows combined with text-based instruction and a checklist provided assistance during practical learning. When learners made a mistake, an Intelligent Tutoring System (ITS) identified the source of the mistake and provided real-time feedback using text, AR and text-to-speech mechanisms. The design was motivated by the human-cognitive architecture and its five evolutionary principles (proposed by Sweller and Sweller (2006)). Users performed four consecutive network cabling training tasks with assistance from our ITS. We found that users made fewer errors when the AR cues, text-based instruction and checklist solutions were replaced with summarised information and then removed completely in the final task compared to those who used an identical system with fixed instruction. Cognitive Load Theory (CLT) explains our results by suggesting that the instructional mechanisms become redundant as knowledge increases. Implications of the study are discussed as well as how AR can help facilitate knowledge transfer in the network cabling domain.



## 1. Introduction

In response to growing threats against computer systems, organisations require for their networks to be secure [1]. A robust method of achieving this is to segment the network into logical groups of physical ports called Virtual Local Area Networks (VLANs) [1]. However, configuring them requires technical knowledge that can be difficult to learn [2,3]. Augmented Reality (AR) overlays virtual information on top of the real-world in real-time [4]. It could help improve learning in network cabling domains because ports are spatially mapped to AR cues in physical 3-Dimensional (3D) space [2].

Intelligent Tutoring Systems (ITSs) are computer-based training systems that facilitate learning of procedures by providing personalised feedback to learners [5]. The effectiveness of AR in education is likely to improve if combined with systems that replicate tutoring approaches (such as an ITS) [6–8].

In this paper, we propose a novel training system that combines AR techniques, real-world equipment, personalised tutoring support and finally, a fading mechanism (which gradually replaces examples with partial examples and then with no examples) to facilitate network security knowledge acquisition. Fading has been shown to improve knowledge retention in mathematical but not AR or networking training domains [9].

The design of our system is grounded in a cognitive load framework, which has robust empirical support [10] and strong theoretical support [11]. Cognitive Load Theory (CLT) suggests that novices waste mental resources trying to solve problems [12,13]. Novices tend to acquire knowledge effectively when problem-solving activities are eliminated or supplemented with examples of how to solve the problem [14]. Conversely, for learners with increasing expertise, the reverse is true. Instructional procedures should be supplemented with problem solving if not eliminated completely [15]. In this paper, expertise is defined as having the skills or knowledge to solve specific problems within a domain [15]. Problem solving is defined as the cognitive processes used to organise and structure information for achieving a goal, such as cabling a network [12]. Remediation is defined as real-time feedback provided by a computing system to guide the user through fixing errors made during a problem-solving learning task.

Among the studies reviewed by Akçayır and Akçayır [16], very few of them focus on CLT and the implications derived from it. Of the very few AR papers centred around the network cabling

☆ This article was recommended for publication by S. Zollmann.

* Corresponding author.
*E-mail addresses:* bradley.herbert@mymail.unisa.edu.au (B. Herbert), grant.wigley@unisa.edu.au (G. Wigley), barrett.ens@monash.edu (B. Ens), mark.billinghurst@unisa.edu.au (M. Billinghurst).

https://doi.org/10.1016/j.cag.2021.09.001
0097-8493/

domain (e.g. Nishino et al. [2], Haramaki and Nishino [17], Herbert et al. [7]), only Herbert et al. [7] used an ITS and no experiment was performed.

Similar papers in the AR/ITS domain (e.g. [18–22]) did not investigate cognitive load effects and only some of them performed an experiment. Those studies used novice participants and also did not use multiple learning tasks to test for knowledge transfer. This is an important consideration because novices require different instructional procedures compared to experts to learn effectively (see [23] for a review). Otherwise, redundant instruction is likely to result, which may interfere with learning [24].

### *1.1. AR-based instructional framework (C1)*

The first contribution made in this paper is a series of instructional principles and effects. We are not proposing extensions to existing cognitive frameworks. Rather, we are applying the existing principles and findings of related work to AR technologies. Our experiment aims to demonstrate these effects on learning to guide AR-based ITS design.

### *1.2. Empirical validation of our design (C2)*

The second contribution made in this paper is an evaluation of the robustness of improving an AR-based ITS by adding a fading mechanism. This demonstration was achieved in two stages. Stage 1 designed and implemented the system to show that the design was technically feasible. Stage 2 demonstrated that the implemented system achieved its intended role of reducing potential side effects that can result from using ITSs in network cabling or similar environments. For example, since problem-solving processes require additional cognitive resources, they can interfere with learning [12].

Problem solving and instructional procedures do not operate in isolation but seem to interact with each other [25]. The mechanisms combined in this paper have mostly been studied in isolation across separate disciplines (i.e. Artificial Intelligence in Education (AIED); AR and educational psychology). This paper brings them together to show how they operate when they interact.

## 2. Related work

This paper is grounded in the theoretical assumptions of Geary [26] and Sweller [27]. It uses these theories to underpin why instruction needs to explicitly guide the user using steps and be gradually replaced with problem-solving activities as expertise increases. This section describes the CLT framework, links this framework with AR and ITS technologies, discusses the limitations of the current state and finally, presents a series of AR instructional principles and effects linked to the human-cognitive architecture [28].

### *2.1. Cognitive Load Theory (CLT)*

CLT suggests that processing novel information during learning activities uses finite cognitive resources [12,13] (but only when processing novel information that humans have not evolved to learn automatically [29]). In particular, it is processed in working memory, which has a finite capacity [30] and duration [31]. CLT is underpinned by a series of principles that describe how humans acquire and process culturally-specific information or skills [11,26,32]. First, problem-solving skill is developed by building up large amounts of knowledge that can be adapted to a set of problems (information store principle).

Second, humans have evolved to automatically acquire information from others and structure it according to an existing schema in long-term memory (borrowing and reorganising principle). This follows from the information store principle by providing a mechanism to gradually construct the information store in long-term memory. Third, sometimes existing knowledge is not useful for problem solving, so learners compare problems and solutions to derive a novel solution (randomness as genesis principle). This is likely to impose a load on working memory and reduce learning [33].

Fourth, since human cognition has no mechanism for structuring novel information, it must be learned using problem-solving techniques which uses working-memory resources [12]. Therefore, working memory is limited in capacity [30] and duration [31] to limit erosion to the information store (narrow limits of change principle). Finally, working memory limitations only apply to novel information that is not structured in long-term memory and that there are no known limits on the amount of information in long-term memory that can be processed in working memory (environment, organising and linking principle).

These principles provide a framework for guiding the design of instruction and helps explain several effects observed in various experiments. In network cabling learning tasks, the following effects are particularly relevant:

1. **Worked-Example Effect:** Worked examples are partially-completed or fully-completed examples of how a problem should be solved (i.e. a checklist or a procedure). Working-memory resources could be wasted by connecting cables into random switch ports to discover the solution [14,34]. Cooper and Sweller [14] found that providing worked examples could reduce the redundant problem-solving mental search and improve learning. Therefore, novices benefit from being instructed using step-by-step instructions [35,36].
2. **Split-Attention Effect:** Connecting cables requires users to divide their attention between a switching rack and paper-based manuals. Similarly, in a network cabling tutoring system, the user interface (UI) needs to be designed so that information is integrated rather than presented across separate UIs [10]. Otherwise, a split-attention effect results and the benefits of the worked-example effect will be lost [37].
3. **Modality Effect:** In network cabling tasks, the potential for split-attention may be impossible to completely avoid. So, audio can be incorporated into the instruction to supplement the visual information. For example, short text phases could be spoken to reduce the amount of content that the user has to read. According to CLT, this effect can reduce cognitive load and increase learning [38].
4. **Expertise-reversal Effect:** Guiding novices through network cabling tasks helps them learn effectively but is not effective for experts. Instead, experts should use problem-solving to facilitate their learning [15,39]. AR may exhibit a similar effect, becoming less effective as knowledge is acquired. There seems to be too few studies comparing this effect and this serves as the motivation for our experiment in Section 4.
5. **Guidance Fading Effect:** Developing network cabling skills through practice should be more efficient when worked-examples are progressively faded out [39]. This is because redundant information imposes a high cognitive load, which interferes with working-memory processing.
6. **Human Movement Effect:** The human movement effect is an emerging effect in CLT [40]. It suggests that physically moving around a switching rack and connecting cables should reduce cognitive load. This effect occurs because human movement is a skill that we have presumably evolved to perform automatically [29].

### 2.2. Intelligent Tutoring Systems (ITSs)

ITSs are a collection of software components that aim to provide personalised learning support by providing remediation or affirmative feedback to student solutions during problem solving [41,42]. They have been shown to improve grades in mathematical domains [43] and knowledge acquisition in AR-assisted motherboard assembly tasks [18]. They differ from traditional e-learning systems because unlike the latter, they integrate cognitive mechanisms [15].

#### 2.2.1. Intelligent tutoring limitations

ITS research aims to understand how ITSs can improve educational outcomes. Therefore, there is a growing need to investigate the potential limitations of ITSs in network education and how these limitations might be mitigated.

Despite the potential of ITSs, they tend to be ineffective for conceptual knowledge acquisition [44,45]. As such, it has been proposed that they be combined with worked-example instruction [45,46].

Combining ITSs with worked-example instruction may reduce their effectiveness when training experts or novices that have gained skills. For example, Menozzi et al. [20] found that the gap between ITS and non-ITS narrowed overtime as the participants' knowledge increased. This suggests that there could be some instructional effect that interferes with their learning [14]. In a CLT framework, the expertise-reversal effect explains these findings [23]. It suggests that instructional techniques that are effective for novices become extraneous and redundant for experts.

One way to overcome this problem is to use fading instruction, which transitions from complete examples, to partially-completed examples and finally, to providing minimal/no examples [47]. For example, when practising a task, the complete solution is provided for the first task. A similar task with a partial example or a more summarised completed example is then presented. The purpose of fading is to facilitate practice using imagination, which helps the user integrate the knowledge into their long-term memory [48]. Various studies have shown that fading out instruction improves performance and transfer [24,49,50].

### 2.3. Augmented Reality (AR)

AR overlays virtual information in real-world 3D space in real-time using a combination of tracking, display and interaction technologies [4,51]. It has promising applications in education [8,52]. For example, in the network cabling domain, virtual information or arrows could be overlaid on physical switch ports through AR [7].

Some works integrated AR into network cabling activities (such as VLAN management) [2,7,53]. Their rationale was that AR improved the mental mapping between the ports and the abstract information. This mental mapping has been well established in non-AR research. For example, in mental model construction and in schema development theories (e.g. Craik [54],Al-Diban [55]). However, cognitive load effects have not been explored in AR cabling domains, despite them potentially interfering with knowledge integration (e.g. [56]).

### 2.4. Limitations of AR education experiments

Constructivist educational philosophies emphasise individualised learning experience, active and discovery learning often through problem solving [3,57–59].

AR research in training, maintenance, assembly and education focuses heavily on discovery, exploratory and other constructivist approaches. For example, Westerfield et al. [18] focused on learning styles and engagement as factors for AR's usefulness. Ibáñez and Delgado-Kloos [60] found that most AR education studies between 2010 and 2017 focused primarily on exploration learning.

While useful, more studies on CLT, evolutionary psychology and instructional design would also be beneficial. For example, such studies could help inform AIED researchers of the suitable mechanisms that need to be incorporated into an ITS. For example (1) instructional differences between novices and experts, (2) the implications of instruction for transferring knowledge and finally, (3) how different cognitive load effects interact with one another in an AR training task. Our experiment is motivated by these limitations and gaps.

#### 2.4.1. Expertise considerations

The first limitation stems from the underlying assumption that instructional procedures effective for novices should likewise be effective for experts. CLT suggests that this assumption may require reconsideration [39]. For example, some studies (i.e. Henderson and Feiner [61],Vargas González et al. [62],Morillo et al. [63]) featured participants with existing domain knowledge and did not demonstrate AR/ITS effectiveness over Virtual Reality (VR) or desktop-based conditions. Morillo et al. [63] found that AR was not significantly better than a video condition for knowledge acquisition. Lin et al. [64] found that AR only marginally helped low-achieving learners while providing no measurable benefit for high-achievers. Radkowski et al. [65] found that the effectiveness of AR annotations varied according to task difficulty.

Many works that we reviewed used novices in their AR education experiments (e.g. Westerfield et al. [18],Henderson and Feiner [66],Hou Lei et al. [67],Tsuruzoe et al. [68]). Other experiments (e.g. Henderson and Feiner [69],Funk et al. [70],Thees et al. [71]) did not seem to report or control for prior learner knowledge.

Some AR studies have controlled for expertise–novice differences. For example, in non-ITS AR-based studies, participants were provided with background information prior to using the system (e.g. Wiedenmaier et al. [72] and Thees et al. [71]). Despite these controls, they neither used an ITS nor a worked-example in any condition.

So, it unknown to what extent the ITS or AR can transfer conceptual knowledge without prior instruction. While we did find some studies using an ITS in an AR training domain, these were too few to demonstrate robustness. For instance, only Westerfield et al. [18],Menozzi et al. [20] and Herbert et al. [22] performed an experiment to test for knowledge retention and/or usability. Westerfield et al. [18] used novices for their motherboard assembly experimental task.

Interestingly, Gavish et al. [73] found that experienced expert users trained using AR made fewer uncorrected errors compared to video-based or VR training systems. This serves as a motivation for why expertise instructional considerations may require adaptation when implementing them in an AR training system. It also suggests that, despite expertise considerations, AR is still superior to video alternatives for problem-solving tasks.

#### 2.4.2. Transfer considerations

The second limitation is the lack of investigation on testing for knowledge transfer, which is when a task is repeated multiple times and users build on their previous knowledge to understand the slightly different scenario in the task. This is important when the goal is to build up a knowledge store as per the information store principle. Evidence is emerging that those with a greater information store automate their skills and perform it unconsciously without attention slips [14,74].

The AR-based ITS studies that we reviewed (e.g. Herbert et al. [7],Westerfield et al. [18],LaViola et al. [19],Menozzi et al. [20], Herbert et al. [21,22],Vargas González et al. [62]) did not test for the effects of transfer on learning [14]. For example, Westerfield et al. [18] performed one simple task to facilitate training. Similarly, Thees et al. [71] used a single task to test if skills or knowledge had improved.

#### 2.4.3. Interaction effects

The third limitation is the lack of integration of certain types of skills and how these interact to influence cognitive load. Some skills and knowledge are described as biologically-primary because we have evolved to perform them unconsciously [32]. The extent to which these skills interact with worked-example instruction and remediation in an AR-based network cabling task is unclear. These implications serve as another motivation for why the different effects need to be combined and investigated in an AR experiment.

### 2.5. AR instructional principles

We adapted the five evolutionary cognitive principles of Sweller and Sweller [11] to AR to devise three AR instructional principles. First, AR facilitates learning by integrating information. Second, AR allows intuitive interpretation of information using biologically-primary skills, and finally, improves knowledge acquisition when combined with an ITS [18].

#### 2.5.1. AR integration principle

The AR Integration principle predicts that learning using AR will result in more effective outcomes because AR creates a mental mapping between the object and the virtual information. This helps in two ways. First, by creating a link between temporary working memory information and long-term memory information by chunking associations into single elements to reduce working memory usage [67]. Second, since the information is tied to the real-world, it reduces the split-attention effect. For example, a study in university physics education showed that integrating information using AR reduces cognitive load as predicted by the split-attention effect [71].

#### 2.5.2. AR interpretation principle

The AR Interpretation principle predicts that AR will enhance understanding of spatial concepts because humans have evolved to interpret 3D information [32]. Hou Lei et al. [67] found that users of an AR training system for Lego assembly required less information compared to the Non-AR group, which used an instruction manual.

#### 2.5.3. AR transfer principle

The AR Transfer principle predicts that AR's effectiveness will be enhanced if it is combined with instruction, such as an ITS [18]. In an AR physics study, AR was demonstrated to have reduced cognitive load [71]. Despite this, no differences in conceptual understanding were found. We suspect that this might have been because there was no remediation or worked-example instruction to facilitate the transfer of knowledge [14]. Therefore, the integration of AR with worked-example instruction should improve AR's effectiveness.

### 2.6. AR instructional effects

The works discussed in this paper (e.g. [64,72,73]) suggest that the AR instructional principles previously described may not universally hold true. We derive a series of effects to explain the inconsistent results.

#### 2.6.1. AR bias effect

The AR bias effect occurs when AR mechanisms implicitly and unconsciously biases the perception of the task or the system used. In a review by Schomaker and Meeter [75], novel stimuli (such as AR) can prime the user to respond more efficiently to subsequent stimuli, which can bias the emotional states. Users may report that the AR cues helped them learn or felt enjoyable. Often, there is no measurable change in learning outcomes relative to non-AR controls because no actual learning effect was created by the AR cues (e.g. Morillo et al. [63]).

#### 2.6.2. AR unconscious-stimulation effect

The AR Unconscious-stimulation effect occurs when users of an AR training system exceed performance expectations on easy tasks. It occurs because AR, due to being used in a novel way, results in a stronger response because humans have evolved to respond more attentively to novel stimuli [75]. AR produces a novel response if it is used for learning of novel information [75].

#### 2.6.3. AR acceleration effect

The AR Acceleration effect occurs when users of an AR-based ITS learn significantly faster than those trained with AR only. For example, users trained using a non-AR ITS, Sherlock, demonstrated roughly 4 years of professional experience with just 25–30 hours of training [76]. The effect is facilitated by cognitive mechanisms designed to operate unconsciously. The implication of this effect is that novices trained with an AR-based ITS may quickly acquire the knowledge and become experts.

#### 2.6.4. AR problem-solving effect

The AR problem-solving effect occurs when users successfully solve a problem in an AR context regardless of the quality of instructional procedures. The effect may result in AR-based instructional principles or knowledge acquisition having no measurable effect. For example, Wiedenmaier et al. [72] found that AR provided no measurable gain over paper manuals for *"intuitive and repetitive tasks."*. So, as a task becomes easier due to gaining experience, the instructional procedures are redundant so should be removed.

#### 2.6.5. AR conditioning effect

The AR conditioning effect predicts that users of an AR training system perform significantly slower or produce more mistakes than non-AR controls due to inattention. Working memory regulates attention and control [77]. Overloading it results in attention lapses and misfires due to incorrect reactions to stimuli [75,78,79].

If cognitive load is too low, unconscious conditioning may occur because there is no effort invested into regulating attention [80,81]. For example, Wiedenmaier et al. [72] found that for easy tasks, AR provided no benefit. Therefore, difficulty should be increased by removing or reducing AR-based cues.

#### 2.6.6. AR novice effect

The AR conditioning effect describes how AR itself becomes redundant with increasing expertise. Conversely, the AR novice effect predicts that the effectiveness of an ITS decreases in an AR training environment as users gain skills or knowledge. The implication is that an AR training system without an ITS should be used to train experts. The effect occurs because the AR conditioning and problem-solving effects interact with the ITS which increases redundant use of mental resources which depletes limited working memory [23].

### 2.7. Summary

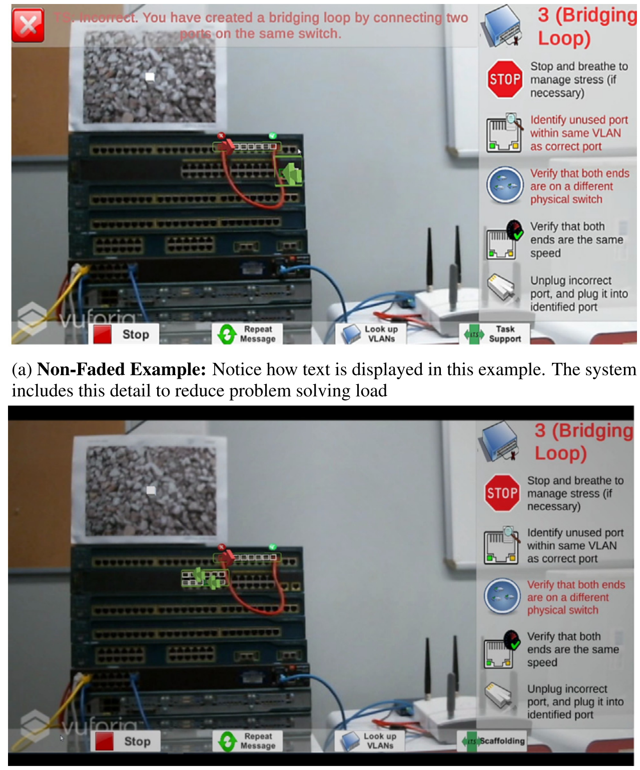
In summary, CLT proposes that humans have evolved to be able to perform problem-solving by structuring information in long-term memory and that when this information is not structured, a heavy load must be processed in working memory [11]. Since working memory is limited in resources and the amount of time it can hold knowledge, this can interfere with learning [12,30,31]. Therefore, Kirschner et al. [82] suggests structuring information and explicitly presenting it to novices during problem-solving learning activities such as network cabling. However, as expertise increases, an effect results where structured instruction interferes with learning and should be faded out [50].

These mechanisms could be combined together in an AR training system to benefit from the advantages that structured instruction, problem-solving and AR can provide. For example, an ITS in an AR-based network cabling training system could provide guided problem-solving. Despite the potential advantages, potentially confounding factors or limitations with this kind of system are unknown. We did not find any user study which investigated these effects in a combined fashion. For example, most AR training tasks either lacked an ITS, did not use worked-example instruction or did not fade out instruction. They were motivated by constructivism, which leans on the assumption that problem solving should precede instruction. Evidence is emerging that this is not ideal for novices [36].

Using CLT, we have designed an AR-based network cabling ITS that switches from full instruction, to partial instruction, to finally, no instruction. This system is described in Section 3. Finally, the related work, our instructional framework and the gaps in existing AR/AIED areas serve as our motivations for evaluating our designed AR cabling system in Section 4.

## 3. System description

Our system, AR EtherGuide, is an AR handheld tablet application that guides users through interconnecting cabling between switches on a physical network rack in accordance with VLAN concepts. When any of the concepts are violated, the system provides feedback and identifies which concept was violated. The system operates in two modes: (i) Task support mode, which provides the same instructional detail across all four experimental tasks and (ii) Scaffolding mode, which fades out instruction after each task is successfully completed until eventually, no examples and annotations are shown (Fig. 1).



(a) **Non-Faded Example:** Notice how text is displayed in this example. The system includes this detail to reduce problem solving load

(b) **Faded Example:** Notice how no text is displayed in this example. The system has faded out the text to reduce redundancy

**Fig. 1.** AR EtherGuide in non-faded and faded modes.

### 3.1. Artificial Intelligence (AI) techniques

Our system uses an ITS to provide problem-solving support, error detection and error correction. Our ITS decouples problem-solving support from remediation. Ashman et al. [36] showed that novices benefit from examples that precede problem solving for difficult learning tasks.

#### 3.1.1. Domain modelling

The ITS uses an Artificial Intelligence (AI) technique called constraint-based modelling (CBM) to represent ideal or satisfactory knowledge as an ontology in a series of constraints [83–85]. Errors are implicit in the solution and are not explicitly defined. This varies from other types of domain modelling where errors are explicitly defined as rules [84]. This has the advantage of not requiring a large library of buggy rules.

Use of CBM is motivated by the fact that network cabling systems have not been extensively studied in AIED research. So, there is a lack of knowledge on the explicit rules needed to be modelled. For example, in ITSs that use explicit rules, a rule could be added to catch certain types of misconceptions. For example, in networking, learners often confuse the switch number with the port identifier (ID). This could be modelled as a bug within a misconception library. In CBM, however, misconceptions are not explicitly defined. A meta-analysis found that while providing feedback on misconceptions was perceived to be more effective, no significant differences were found [86].

Finally, CBM is combined with a variant of model-tracing that checks the constraints each time a cable is either plugged in or unplugged. If any of the constraints in the solution are violated, then the system determines that an error has been made. The system does not have a list of ports hard coded into the system, which are correct. Rather, it infers if both ends of a port is valid by checking to see if any of the constraints are in violation using the domain model in Appendix B. For example, if the two ports belong to the same switch, then a constraint violation (bridging loop) has occurred.

#### 3.1.2. Pedagogical modelling

Our system performs pedagogical modelling, which aims to replicate teaching strategies. Once either a correct or erroneous action has been detected, the system then needs to know what it should do. In our case, this typically involves displaying a feedback message and showing some AR arrows, ticks or crosses.

The pedagogical model comprises of a series of properties that are mapped to variables in Extensible Markup Language (XML) format (Appendix C). When the fading mode is enabled, some of the pedagogical properties are disabled (i.e. displaying of annotations is set to false, which disables it). As a result, annotations are no longer shown. The pedagogical model is consistent between users, and only varies in the case of the fading condition.


**Fig. 2.** The User Interface button pane.

#### 3.1.3. User modelling

Our system performs user modelling, which involves personalising the instruction based on data contained within the user model. Since our study focuses on certain instructional interventions, the user modelling capabilities are limited.

The first type of user modelling is keeping a record of previous violations in the domain model. So, if the user repeats an error that they have already committed, the system provides instruction that they have already been instructed on that error. This helps reduce unnecessary redundancy. Similarly, the previous cabling configurations already tried are also stored in the user model. So, if the user interconnects the same two ports, then the system provides feedback that the ports have already been tried.

The second type of user modelling is a list of feedback options, which the system iterates through. Each message is paired with an "if statement" that consists of a property and a Boolean value. If the property is found in the user model, the Boolean evaluates to true. Otherwise, it evaluates to false. For example, it may display shorter messages for subsequent attempts of the same cabling error rather than merely repeating the previous message again.

### 3.2. User interface design

The UI was designed in accordance with CLT and the AR instructional principles (proposed in Section 2). Our UI was developed over time and evaluated through several pilot studies (see [7] for an earlier iteration of the system) as part of a larger project. The AR EtherGuide UI consists of three components: (1) the control pane; (2) feedback pane and finally (3) the check list and VLAN pane.

#### 3.2.1. Control pane

The control pane contains touchable buttons that the user infrequently pushes (Fig. 2). They are primarily used for initially configuring the UI for the experiment.

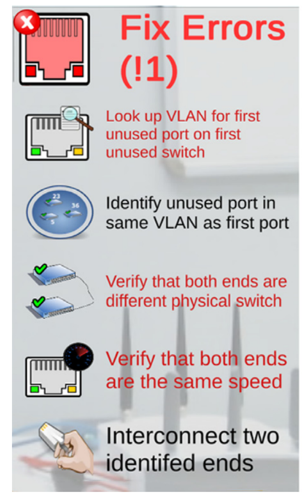
- **Stop:** Manually stops the current task and proceeds to the next cabling task.
- **Repeat Message:** Text messages on the tablet screen stay visible for 10 s. Pressing on this button forces the last played message to be repeated.
- **Look up VLANs:** Toggles the VLAN table on or off (Fig. 4(b)).
- **Task Support:** Toggles between Task Support and Scaffolding modes. It cannot be changed once the task has been started.

#### 3.2.2. Feedback pane

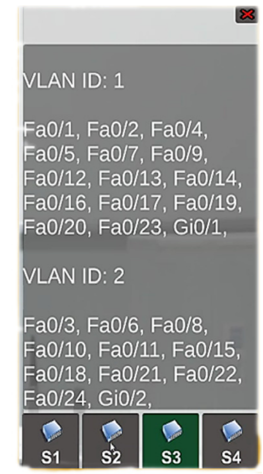
The Feedback pane shows instructional text to the user (Fig. 3). When the action by the user is determined to be correct, virtual tick marks are displayed over the two ports (Fig. 1(a)). The user receives feedback that their solution is correct. The text also explains why the solution is correct. This helps clarify tentative but correct actions in novices [87].

Synthesised audio of feedback messages were generated on-demand using the open source Mary text-to-speech system [88]. Mary was installed on a backend Intel i7 Nvidia GTX 950 gaming personal computer (PC). Audio was streamed to a Galaxy Samsung Tab S5e Wi-Fi 64GB tablet over a web connection. The tablet then played the audio, helping the user avoid splitting their attention between trying to read the feedback and the wiring rack [38,89].


**Fig. 3. The Feedback Pane:** shows the text-based feedback to the user. This is typically feedback about the learner's progress.




(a) **Check List Pane:** Items the user should check when diagnosing a fault




(b) **VLAN Table Pane:** The digital VLAN table shows the ports and which VLAN ID they are mapped to

**Fig. 4. Checklist/VLAN Table pane:** Toggles between VLAN and check list pane.

#### 3.2.3. Check list and virtual local area network panes

The Check List pane allowed the user to quickly check that they have covered all the items before connecting two ports (Fig. 4(a)).

The check list provided the items that users needed to mentally learn. Since these items were mentally checked before removing or connecting a cable into the rack, they are probably best displayed simultaneously as a group. Users must rehearse the entire check list because missing out even a single item would result in an invalid connection. So, separating them creates a temporal (time-based) separation that results in a split-attention effect [10].

The VLAN pane showed all of the ports within the entire rack and their corresponding mappings to VLAN IDs (Fig. 4(b)). Ports were automatically and randomly assigned to randomly-created VLANs when the task was started. Users could toggle the pane on/off by tapping on the *Look up VLANs* button on the control pane. When the learner made an error, the VLAN pane was hidden automatically and the check list pane was displayed in its place.

### 3.3. Instructional support

Both variants of the system provided three types of instructional support: (a) AR-based worked-example instruction; (b) check list-based worked-examples and (c) remediation (which is provided using the AI techniques described previously). The experimental condition included an additional instructional support type called fading, which is performed after each task is successfully completed. It does not use the user model to adapt to the user. The ITS primarily provides the remediation instruction. The worked-example instructions are provided by the AR front-end and do not depend on an ITS. The combination of checklist and AR provide the examples needed to reduce working-memory wastage resulting from randomised search processes [14]. They aim to eliminate the unnecessary mental process of cross-referencing the port labels identified in the VLAN table with the physical port located in the switching rack.

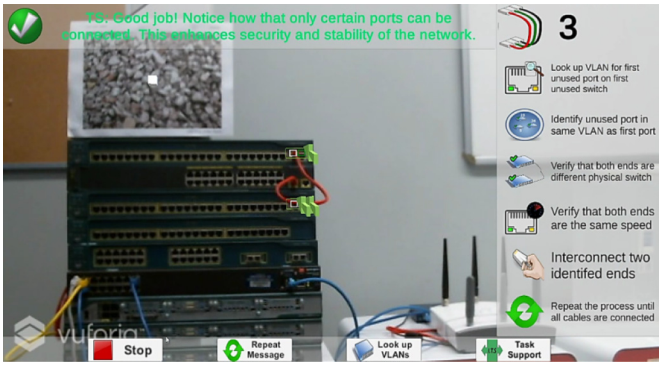



**Fig. 5.** Shows AR arrows over ports to be connected. One or more incorrect ports are also shown at random to prevent the user from blinding following the instructions.

#### 3.3.1. AR-based worked-examples

The first type of instructional support is AR-based worked example instruction. Prior to plugging in the cable, the system showed green arrows to indicate a series of recommended ports to connect (Fig. 5). The arrows shown differed according to the mode and task number in the experiment. For the control condition, the arrows shown did not vary across tasks.

This mechanism supports novices who lack an integrated knowledge base to solve the problem without assistance [35,36].

#### 3.3.2. Check list worked-examples

The second type of instructional support is checklist-based worked-example instruction. Prior to plugging in a cable, the system showed the items that the user should check to ensure that a cabling connection is valid. When an error is made, another check list is displayed, showing the diagnostic steps that the user should follow to identify and fix the error. This mechanism helps novices acquire the appropriate knowledge (Fig. 1(a)).

#### 3.3.3. Remediation

The third type of instructional support is remediation, which is personalised feedback provided during problem solving. It also reduces redundant mental activity that can impede learning by providing immediate feedback after a participant connects or disconnects a cable [45]. Remediation relies on the AI techniques provided by the ITS.

Either the remediation will be positive (i.e. it will affirm that the user has done it satisfactorily) or it will be negative (i.e. it will provide information that the user needs to know to fix the problem). For example, in the case of an incorrect cabling action, virtual red arrows appear over each of the incorrect ports to indicate to the learner which port(s) should be disconnected. After plugging in or removing a cable, AR EtherGuide shows the steps that the learner needs to follow to correct any outstanding problems, if any (Fig. 1(a)).

#### 3.3.4. Fading

The fourth type of instructional support is fading, which is only included in the experimental condition. AR EtherGuide performs three types of fading: (a) worked-example fading, which gradually reduces the detail in the check list pane (Fig. 4(a)); (b) AR-based fading, which gradually reduces the amount of arrows and AR details displayed and finally (c) remediation fading which removes the feedback (Fig. 1(a)).

First, the check list pane transitioned from listing all items in detail that need to be completed, to listing the steps briefly and finally, to removing all of the items. Second, the AR arrows were shown for two ports that can be interconnected together. Then, starting with the second task, only one arrow was shown, so the second port would need to be resolved through problem solving. For the third task, both arrows were removed and no recommendations were given. Arrows were still shown if the user made an error. For the fourth task, no arrows were shown even when the user made an error but a tick and/or a cross still appeared above the two ports. Third, remediation fading transitioned from giving direct and explicit feedback, to simply advising the user of the type of error made, to advising the user that an error has been made and finally, to complete removal of the feedback. After each task, the system presumed that sufficient expertise was obtained.

### 3.4. Development

AR EtherGuide was developed using the Microsoft C# programming language and .NET 4.7. The AR components were developed in Unity3D [90]. The ITS components were developed using Visual Studio [91].

#### 3.4.1. Augmented reality components

The AR front-end was developed using Unity3D 2018.1f gaming engine (Unity) [90]. The models themselves were sourced from the public domain and resized in an open-source modelling suite called Blender [92] before being imported into Unity.

Tracking functionality was implemented using the Vuforia 8 AR tracking library [93]. An image marker with a stones pattern was printed out and attached to the switching rack in a fixed location. Pilot tests were undertaken to test the annotations for occlusion and registration errors. An acceptable level of performance was obtained after tweaking the software. We found that registration errors were not noticeable on the tablet but were noticeable on the PC during testing (as can be seen in Fig. 1(b)). We used white frames to highlight surrounding ports to compensate for potential misalignment.

#### 3.4.2. Intelligent tutoring components

The ITS components were implemented as three separate agents using C# .NET 4.7. Each agent was a .NET console instance that shared the same local PC hardware resources. Each agent was as follows:

- **Assessment agent:** The Assessment agent determined which constraints were violated (or not). It performed the domain modelling. It determined what kind of mistake the user made (if any) and communicated with the Pedagogical agent.
- **Pedagogical agent:** The Pedagogical agent managed what feedback should be displayed. It then communicated with the User agent to determine what type of feedback should be displayed in the AR application. Based on this decision, the Pedagogical agent updated the user model with information so that the user's progress could be tracked in real-time.
- **User agent:** The User agent stored information about the user and their progress throughout each of the tasks.
- **Core agent:** The Core agent interconnected all of the agents together and forwarded messages between the ITS agents and the AR application.
- **Dispatcher agent:** The Dispatcher agent coordinated communication between the Assessment, the Pedagogical and the User agent. The dispatcher used Transmission Control Protocol (TCP) to communicate with the Core agent, which used a raw TCP socket to communicate with our AR application.

## 4. Methodology

We performed a between-participant, randomised-controlled experiment comparing two variants of our AR EtherGuide ITS: one which operates in AR non-faded mode and one which operates in AR faded mode. This was done to understand how varying the instructions produces effects on cognitive load and learning in a network cabling training environment. A total of 30 participants (10 self-reported females and 20 self-reported males) participated in the experiment. Half of the participants were assigned to the AR non-faded condition and the other half were assigned to the AR faded condition in a random manner.

### 4.1. Participants

Participants were randomly assigned to one of the two AR training conditions (AR non-faded or AR faded) to control for selection bias and prior knowledge. We opted out of using diagnostic tests to avoid creating biased stereotype threat effects or test anxiety [94]. Also, tests are only useful for evaluating known factors [95]. In experimental studies where novel technology is used, it is not known what factors are relevant and which ones are not. This makes random assignment more effective than prior screening assignment [95].

The age range for the entire cohort was 25–40 years old ($M =$ 30 years old, $SD = 7.80$). The average age range for the non-faded (control) group was 33 years old ($SD = 8$) and 27 years old ($SD =$ 6) for the faded (experimental) group ($U = 61$; $p = .03$; $d = .85$). For each age range (e.g. 18–24), we calculated the median value (e.g. 21 years) and calculated the mean of each of these values to get an approximate age mean.

Most participants reported prior educational experience, which allowed us to control for variability in learning skills as much as possible. They reported minimal or no cabling and VLAN experience using a 7-point Likert scale with 7 being very high knowledge and 1 being very low knowledge ($M = 1$; $SD = 1$) (Appendix A). Their pre-training task knowledge score percentages were captured at the start of the experiment ($M = 37\%$; $SD = 26$) using a multiple choice exam (Appendix A) followed by a practical assessment task. The average pre-training total knowledge score for the non-faded (experimental) group was 36% ($SD = 27$) and 39% for the faded (experimental) group ($SD =$ 21). These scores did not significantly differ, suggesting a prior knowledge effect was not present ($t = 0.27$; $p = .79$). Finally, during the pre-training practical task, the system recorded each cabling error made by participants. The non-faded (control) group made 3 errors and the faded (experimental) group made 4 errors. The differences were not significant and hence, effects were unlikely due to group differences ($t = 1.21$; $p = .24$).

The mental demand scores were measured using the National Aeronautics Space Administration Task Load Index (NASA-TLX) [96] both before and after the training task. The pre-training task scores were within average ranges [97] ($M = 51$; $SD = 32$). This suggests that the task could benefit from instructional procedures and it also suggests that the task was not too easy (i.e. the mental demand was not below the average). The mean mental demand scores between the two groups on the pre-assessment task were not significant ($U = 107$; $p = .82$). This suggests that the learner's knowledge prior to the intervention had no effect on the mental demand.

### 4.2. Experimental procedure

The experiment and the procedures used were approved by our institution's human ethics group. A health and safety officer then approved the study. Fig. 6 shows the experimental process.

#### 4.2.1. Recruitment

Participants were recruited from a combination of undergraduate/postgraduate university classes and from administration with a diploma to reduce variability in educational knowledge and expertise. Recruitment was done via an email circulated within the computer science department of the university. Participants were paid a $10 Gift voucher for their participation. Potential participants contacted the researcher by email if they were interested.

A questionnaire was used to assess and manage Coronavirus Disease 2019 (COVID-19) risk by asking them a series of questions about their travel and if they had any respiratory symptoms such as a running nose, cough, fever or sore throat within the last 14 days. High risk was determined if the potential participant indicated one or more symptoms or had travelled to a high risk location. Those identified of being at high-risk were asked to fill-out the questionnaire 14 days later. If they continued to report symptoms and no cause for those symptoms were identified, they were excluded from the study (Fig. 6). Only one person reported symptoms, waited 14 days and reported symptoms again, so they were excluded from the study.

#### 4.2.2. Pre-experiment

Data was collected using a learner survey UI shown on the tablet (Fig. 7) to avoid potentially spreading COVID-19 by touching the paper and transferring the virus to face, eyes or mouth.

Since the experiment was conducted during the COVID-19 pandemic in 2020, participants sanitised their hands upon arrival and maintained at least 1.5 metres (6 ft) from the experimenter at all times. The experiment space was shared by a maximum of two people (a facilitator running the experiment and the participant).

Participants completed a demographics questionnaire that asked them to provide their gender, their self-perceived experience with network cabling, VLAN concepts and prior experience with AR/VR technologies (Appendix A.1).

Participants then completed a multiple-choice knowledge test (Appendix A.2). The experimenter explained that the experiment was not a test of intellectual ability, what each of the system's buttons did, what feedback to expect and to make sure that they connect all four cables into the rack. They were told that there were principles that had to be learned and that the tablet would provide the information needed to learn the principles assessed by the multiple-choice knowledge test.

#### 4.2.3. Experimental tasks

The experimental tasks aimed to improve participants' knowledge, skills, transfer and internalisation of VLAN switching concepts in a series of similar network cabling tasks, which were as follows:

- Both ports must be assigned to the same VLAN ID.
- Both ports must be of the same physical speed to maximise efficiency.
- Both ports must reside on different physical switching devices (bridging loops must be avoided).

In addition, learners should also learn:

- Appropriate terminology (e.g. a bridging-loop and how to identify it).
- Port naming conventions (e.g. Fa0/1 means FastEthernet0/1)
- Port identification (e.g. knowledge of how to locate ports on the rack).

Ports within the switching rack are mapped to a VLAN ID, which can be looked up in the table (Fig. 4(b)). Each port in the rack is unique, so port *Fa0/3* on *S3* is not the same as port

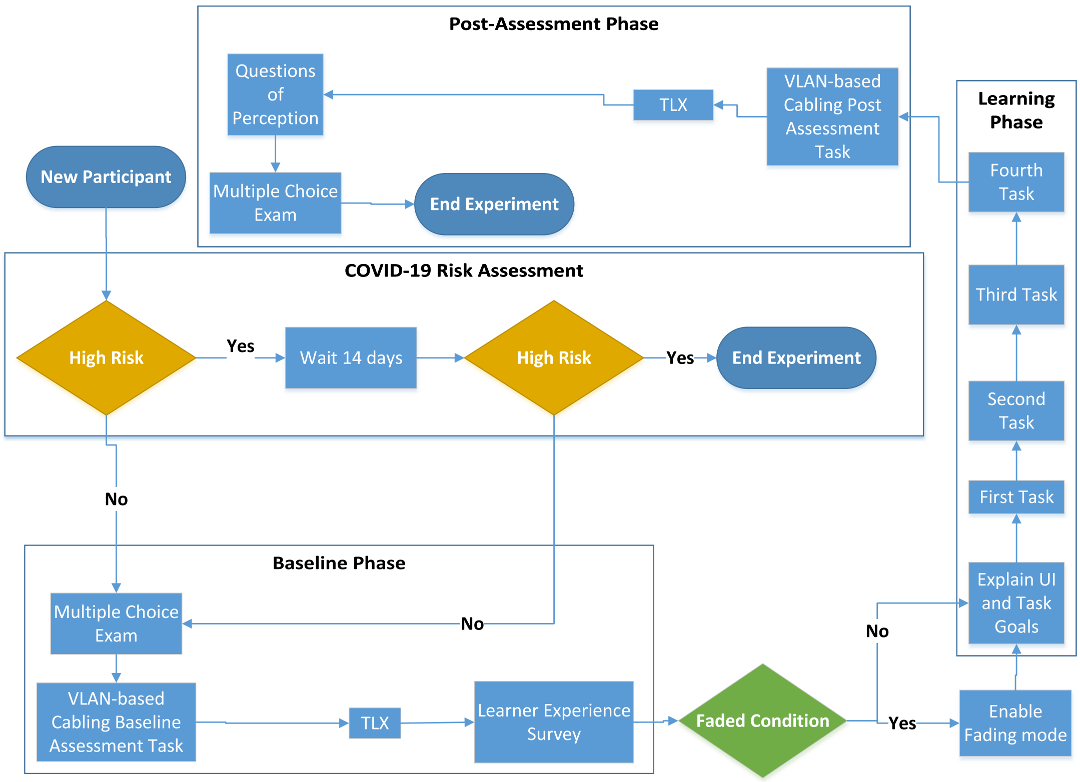



**Fig. 6.** Shows the flow through the phases in the experiment.

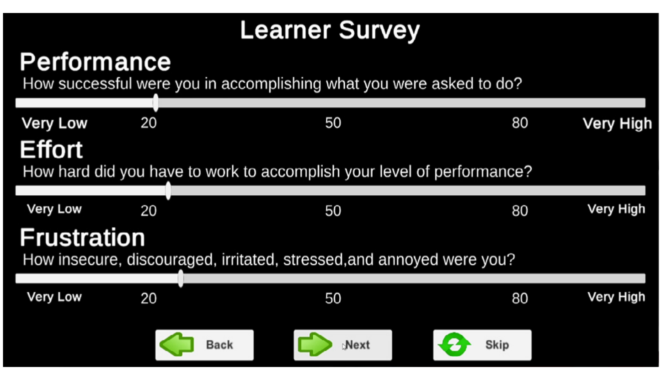



**Fig. 7.** The learner survey UI seen on the hand-held tablet screen.

*Fa0/3* on *S1*. Each port could be mapped to a different VLAN ID. Fa is an abbreviation for *FastEthernet* and corresponds with the speed of the link. A single VLAN ID represents a group of ports. It can have multiple ports mapped to it. Only ports mapped to the same VLAN ID can be interconnected together. The participants must learn these principles, which can be difficult because the mappings often produce errors under high load environments.

The participant began by attempting to complete a practical network cabling without any learning assistance except for looking up the VLAN table. Participants had 15 min to try and complete the task. This step was necessary to measure the learner's baseline knowledge so that any subsequent knowledge gained could be compared. The participant had to interconnect five cables between the switches in the rack.

Next, the participant performed four consecutive training tasks, which were not timed. The VLANs and the ports were randomised to control for memorisation, to reduce observer bias, to reduce potential confounding factors by using the same combination of ports and VLANs and to provide a chance to adapt the knowledge learned to slightly different port layouts.

After the fourth cabling task, participants in both conditions once again repeated a 5-cable network cabling VLAN task without AR or text-based guidance to see if their knowledge and skills in the domain had improved. Participants were allowed up to 15 min to complete this task.

The independent variables were the two groups (AR non-Faded v. AR faded). Each task was equal in difficulty and design with the noteworthy exception of randomising which ports were assigned to which VLAN. This was done to reduce subconscious memorisation from the previous task and to prevent the observer from potentially unconsciously communicating the answer using non-verbal cues [98]. The differing port combinations required the participant to transfer their knowledge to the new situation, which should be difficult if cognitive load impedes learning [14,99]. Therefore, errors during this transfer step should demonstrate a cognitive load effect on learning.

### 4.2.4. Post assessment

Finally, the participant completed another knowledge exam with a 15-minute time limit. It was identical to the first exam to reduce potential differences due to question interpretation. Participants were not told their answers until after the study, so it was not possible to remember the answers from the first test.

Participants were informally asked questions about the system such as: *Overall, what were your impressions of using our system today?'* or *'What were some weaknesses with using the system?'*. This allowed the researcher to ask follow-up questions to clarify their initial responses.

### 4.3. Research tools

The following research tools were used:

#### 4.3.1. Galaxy Tab S5e Wi-Fi 64GB

Our research used a Galaxy Tab S5e Wi-Fi 64GB tablet with a 10.5 inch screen to provide sufficient screen space to display the content.

Our rationale for using a tablet instead of a head-worn AR device, such as the Microsoft HoloLens, were as follows. First, previous work [56] found that they can impose a higher cognitive load in AR tasks. Second, unfamiliarity with using a head-worn AR device may bias the results due to a novelty effect. Third, it may create an unplanned tertiary task that imposes a cognitive load by splitting working memory resources between figuring out how to operate a head-worn AR device and the learning task. Finally, hand held tablets or phones are widespread and likely to be adopted as an industry choice. Their cost effectiveness and ease of use due to familiarity makes them suitable for industry and educational use [21]. We reduced potential split attention effects resulting from perceived mismatches by using spoken audio [38,89].

#### 4.3.2. Networking equipment

The network switches used in this experiment were three Cisco Catalyst 2950 switches and one Cisco Catalyst 2960 switch. Different models of switches were used to create variety in the training scenario so that learners could learn skills across both models. All switches were reset to their default (factory) settings and no further configuration of the switches was done.

The switches were interconnected via a serial (console) connection to an OpenGear Infrastructure Manager 4216 console server appliance [100]. This appliance reads console messages in real-time and is used to detect when cables are plugged or unplugged. This appliance connects with our system over Telnet to provide real-time response to cabling connections. Up to 24 switches can be interconnected via this appliance. In our experiment, four switches were used.

A Cisco Aironet LWAPP (AIR-LAP1252AG-N-K9) wireless access point was used to provide wireless network connectivity between our hand held tablet and the backend infrastructure. The access point used the 802.11N standard for communication.

#### 4.3.3. Desktop computer

An Intel i7 desktop gaming computer with an Nvidia GTX 950 series graphics card, two 256 gigabyte solid state drives (SSDs), an Intel i7-6700 processor and 16 GB of system memory was used to run the backend components. The computer was interconnected with the networking equipment via an unmanaged NetGear switch.

#### 4.3.4. Data collection equipment

Data collection was performed by the OpenGear appliance which recorded when cables were plugged or unplugged. This information was then provided to our backend system running on the gaming computer, which logged this information. Survey information was collected electronically using the survey questionnaire in the tablet. This UI was developed and integrated into our Unity application so that the user did not have to manually switch to a different system. Time and errors were recorded automatically by the system and did not require manual intervention.

### 4.4. Experimental measures

We evaluated the effectiveness of AR EtherGuide using a series of measures including: (1) user perceptions (i.e. NASA-TLX [96]; learning experience surveys (Appendix A) and asking participants about their learning); (2) knowledge acquisition (i.e. knowledge scores of the exam (Appendix A)) and (3) performance measures (i.e. error count and task completion time in seconds). Measuring knowledge acquisition follows from the information store principle and cognitive load. Higher errors and less knowledge indicates higher cognitive load in a CLT framework [14].

#### 4.4.1. NASA-TLX

The NASA-TLX was used to measure subjective task load during the task [101,102]. The instrument sub-scales allowed us to isolate physical exertion from mental effort caused by high cognitive load, so was more suitable over other workload instruments [103,104].

#### 4.4.2. Problem-solving dual-task

Problem solving was added as a secondary task to intentionally impose a load on working memory. So, if performance in the secondary (problem-solving) task diminishes, then it demonstrates a cognitive load effect [102]. For example, reduced performance is measured by counting the errors made. In the network cabling tasks, users use mental resources to mentally search for the correct port and map it to VLANs. Users must also diagnose any invalid connections that they may make. Any errors in these processes would indicate reduced problem-solving capability. These errors would worsen if a cognitive load effect, such as an expertise-reversal effect, was present.

#### 4.4.3. Knowledge retention

Knowledge retention was measured using a multiple-choice knowledge exam (Appendix A) before and after the experiment task to compare how much the knowledge had changed. Higher knowledge scores indicate that the instruction imparted some knowledge since other sources of knowledge were eliminated. If differences between groups were detected, it should show that the imparting of that knowledge was impeded. Therefore, we would expect to see a difference between groups if there was a cognitive load effect present.

Each question of the pre-knowledge and post-knowledge exams were assigned to two categories: (i) total knowledge and (ii) domain, application, comprehension or metacognition. For example, a correct answer worth 4 points of a question assigned to the metacognition category would result in a metacognitive score of 4. All the questions in each category were summed to produce a total value for each category and an overall knowledge score.

#### 4.4.4. Emotional regulation dual-task

Emotional regulation has often been considered an extraneous dual task in cognitive load studies (see Plass and Kalyuga [80] for a review). So, if a cognitive load effect was present, then we should also expect to see reduced capacity for users to self-regulate their emotions in the learning task. Self-regulatory skill involves the use of certain thoughts to maintain self-control when the primal emotions are triggered.

We observed the participants' behaviour for indications of impeded emotional regulation and also asked them to comment on the learning experience. The latter was free-flowing so that potential emotional responses could be identified using a qualitative approach.

### 4.5. Data analyses

The following types of analyses on the data were performed.

The knowledge dimensions, completion seconds and error count data were intervals, so were tested for normality using a Kolmogorov–Smirnov test ($\alpha = 0.05$). We assumed normality if $a$ was greater than 0.05. For normally-distributed data, we performed a parametric student t-test for dependent means between the pre-assessment and post-assessment scores for each category ($\alpha = 0.05$). For normally distributed inter-group data, we performed a parametric student t-test for independent means ($\alpha = 0.05$).

For the data, which was not normally distributed (e.g. ratio data), we performed a pairwise comparison between the pre-assessment and post-assessment tasks using a non-parametric Wilcoxon Signed Rank test ($\alpha = 0.05$). For non-normal error count data between groups, we performed a non-parametric Mann Whitney-U test ($\alpha = 0.05$). We analysed the NASA-TLX data by decoupling each individual dimension to provide diagnostic information on which dimensions were being shaped by the condition. However, due to a recording error, only the performance and mental demand sub-scales of the NASA-TLX were usable. Inter-group analysis was performed using a non-parametric Mann Whitney-U test ($\alpha = 0.05$).

### 4.6. Hypotheses

We made the following hypotheses based on the AR instructional principles and effects derived from the literature in Section 2.

1. **H1: the mental demand score of the post NASA-TLX will be greater in the AR non-faded (control) group:** The AR non-faded (control) group is predicted to have a higher mental demand score on the post NASA-TLX because of higher cognitive load resulting from the AR Novice effect. The problem-solving task will subjectively feel harder to accomplish because of reduced cognitive resources being available.
2. **H2: the AR non-faded (control) group will commit more errors than the faded (experimental) group:** As cognitive load increases, the ability to carry out the secondary (problem-solving) task will diminish. Therefore, the participants in the AR non-faded (control) group will commit more errors due to the AR conditioning effect.
3. **H3: Knowledge retention of the AR faded group will be greater than the AR non-faded group:** We predict that due to the CLT expertise-reversal effect [39], learners in the AR faded (experimental) group will show greater knowledge retention and acquisition compared to the AR non-faded (control) group.
4. **H4: Self-Regulatory ability will be impaired in the AR non-faded (control) group:** The AR non-faded (control) group will exhibit behaviours that indicate poor self-regulatory ability and will not be able to suppress primal emotional responses.

### 4.7. Research questions

This study investigated the following research questions:

1. **RQ1** What are the instructional implications of an AR-based ITS in a network cabling task?
    - **RQ1.1** Does adding a fading mechanism to an AR-based ITS improve knowledge acquisition and recall?
    - **RQ1.2** Does adding a fading mechanism to an AR-based ITS reduce unconscious errors performed in a network cabling VLAN learning task?
    - **RQ1.3** Is there an interaction between worked-example, problem-solving, AR and fading?
2. **RQ2** What cognitive mechanisms do users use when problem-solving network cabling tasks?
    - **RQ2.1** What self-regulatory mechanisms do users use in a network cabling task?
    - **RQ2.2** What unconscious cognitive mechanisms are employed in a network cabling task?

**Table 1**
Comprehension scores.

| Item | AR non-faded | AR faded |
|---|---|---|
| Baseline comprehension score | 36% | 36% |
| Post comprehension score | 58% | 64% |
| Comprehension score gain | +22% | +28% |
| **Baseline SD** | 29 | 26 |
| **Post SD** | 27 | 28 |
| **Gain SD** | 9 | 38 |
| **Baseline/Post p** | .02* | .01* |
| **Baseline/Post t** | 2.74 | 2.79 |
| **Gain p** | .65 | |
| **Gain t** | −0.46 | |

*Significant result ($p < 0.05$).

**Table 2**
Uncorrected error counts.

| Item | AR non-faded | AR faded |
|---|---|---|
| Baseline errors | 2 | 3 |
| Post errors | 1 | 1 |
| Error reduction gain | 1 | 2 |
| **Baseline SD** | 1.74 | 2.02 |
| **Post SD** | 1.78 | 1.31 |
| **Gain SD** | 1.33 | 2.10 |
| **Baseline/Post p** | .04* | < .01* |
| **Baseline/Post t** | −2.26 | −3.91 |
| **Gain p** | .04* | |
| **Gain t** | 2.11 | |

*Significant result ($p < 0.05$).

## 5. Results

### 5.1. Knowledge acquisition

The average domain knowledge gain score for the AR non-faded (control) condition ($M = 7\%$; $SD = 25$) did not significantly differ from the AR Faded (experimental) condition ($M = 9\%$; $SD = 33$), $t(28) = -0.2$; $p = .84$. Table D.4 provides more detail.

Both conditions facilitated knowledge comprehension, $M_{\text{AR non-faded gain}} = 22\%$; $SD_{\text{AR non-faded gain}} = 9$) and $M_{\text{AR faded gain}} = 28\%$; $SD_{\text{AR faded gain}} = 38$). Table 1 provides more detail.

Application knowledge did not differ between the baseline and post-assessment measures or between the conditions. Table D.5 summarises the findings.

### 5.2. Performance

Table 2 shows the amount of uncorrected errors that were recorded by the system during the tasks (Fig. 8).

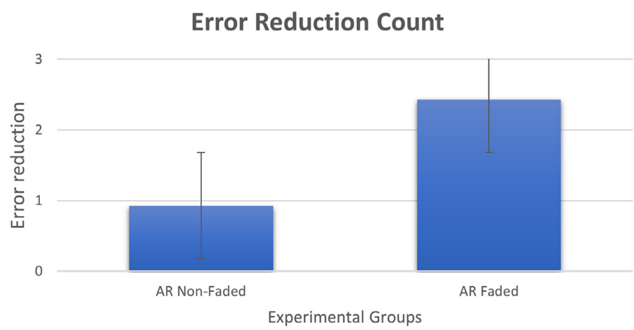



**Fig. 8.** (Uncorrected) error reduction count between the two post assessments.

**Table 3**
Metacognitive scores.

| Item | AR non-faded | AR faded |
|---|---|---|
| Baseline metacognitive score | 39% | 42% |
| Post metacognitive score | 36% | 36% |
| Metacognitive score gain | −3% | −6% |
| **Baseline SD** | 21 | 22 |
| **Post SD** | 23 | 20 |
| **Gain SD** | 23 | 28 |
| **Baseline/Post p** | .60 | .44 |
| **Baseline/Post t** | −0.54 | −0.79 |
| **Gain p** | .80 | |
| **Gain t** | 0.26 | |

* Significant result ($p < 0.05$).

### 5.3. User perceptions

#### 5.3.1. Interview and observations

H4 was not supported because there were no definitive indicators of limited self-regulatory abilities. Nevertheless, eight participants reported that they found the system's AR cues to be useful but that they would become too reliant on using them. One of these participants also gave feedback that they were not giving their full attention because they perceived that AR was telling them what to do.

Two participants also reported that they felt confident in the AR Faded (experimental) condition when answering the exam scores on the post-assessment questionnaire. This seems consistent with similar reports of a confident perception that they had improved their skills using the system. For instance, one participant described their experience as knowing nothing before they did the experiment but felt that they understood the basic concepts. Conversely, AR non-faded (control) participants did not seem to express such feelings.

During the training task(s), there were observations of users trying random things. Sometimes, learners felt that two options could be true and were able to eliminate the incorrect option as a result of getting feedback from the system. This enabled improvement on subsequent attempts. However, each subsequent task was a transfer task with different VLAN combinations. These combinations sometimes revealed misconceptions. For example, in the first task, users did not realise that the switch ID and port ID differed.

#### 5.3.2. Learner experience survey findings

The results from the learner experience survey (Appendix A.3.1) were not significantly different between tasks or groups. The results from the learner experience survey produced an overall score of the learner's metacognition (knowledge about effective learning strategies). Table 3 summarises the findings.

## 6. Discussion

Users were found to have committed more errors in the non-fading (control) group. There was no indication that the mental demand, knowledge retention or self-regulatory abilities differed (H1, H3 and H4 not supported). The lack of support for H1, H3 and H4 were surprising given the cognitive load effect in the AR non-faded (control) group. One explanation might be a limitation with the measurement tools used (e.g. the measured cognitive load fluctuated and was not captured by the NASA-TLX (mental demand) score [105]). Another explanation is that the fading mechanism is intended to facilitate skill development and automation of the knowledge. Experts already have the base knowledge but their performance declines when such knowledge becomes redundant. It is also plausible that our system lacked an appropriate mechanism to facilitate sound knowledge acquisition. Finally, H4 has some support but it is insufficient to conclude that it is supported.

### 6.1. Instructional implications (RQ1)

This experiment provides several instructional implications as predicted by CLT. We found evidence of a worked-example effect; an AR novice effect and finally, an AR conditioning effect. Therefore, as discussed below, user expertise has implications on the design of instruction.

#### 6.1.1. AR transfer principle

The AR transfer principle likely played a key role in the comprehension score increases across both groups (Table 1). AR cues seemed to help the users initially, since they were observed reacting to the AR cues even when they appeared to have limited or no knowledge of the domain. They followed the AR cues in the first instance.

An effect on knowledge recall was not reflected in the post-knowledge exam scores (RQ1.1). Nevertheless, we observed reduced reliance on random problem-solving processes. For example, learners were observed gradually changing their learning strategies. This suggests that knowledge was present in long-term memory that was not there at the start of the experiment [34] (RQ1.1). If so, they must have acquired the knowledge during the experiment. Therefore, there is some support for H3 but the effect needs to be stronger for one group before claiming that it is supported.

Unlike some prior works (e.g. Thees et al. [71]) which found limited evidence of knowledge differences between pre/post conditions, we found knowledge gains in the comprehension area. Usually, ITSs are not very effective for facilitating this kind of knowledge transfer [44,45]. Despite this, our system improved this knowledge. This was likely due to the worked-example effect used across the two groups. It explains why the gain scores did not differ between the groups as predicted by H3. Future work would be required to isolate worked-example instruction from an AR-based ITS to confirm this working hypothesis.

#### 6.1.2. AR interpretation principle

The AR interpretation principle explains the lack of support for H3 (knowledge acquisition) differences between groups. Humans can interpret their environment effortlessly because they have evolved to do it automatically [32]. Therefore, knowledge transferred through AR, would be effortless to process, especially when it comes to intuitive conceptualisation of the port mapping in 3D space. AR likely reduces the extent to which cognitive load interferes with knowledge acquisition. This is based on the knowledge gains recorded in both groups. Once again, future studies are needed to confirm this working hypothesis but there are theoretical explanations in the literature (e.g. Geary and Berch [32]).

#### 6.1.3. AR conditioning effect

The first type of expertise-redundancy effect was the AR conditioning effect. As users' learning progressed, eight participants started to ignore the AR cues. One participant was observed trying a different (non-highlighted) port just to see what happened. Others reported that they felt that the AR cues would create negative behavioural patterns of becoming overly reliant on the system and fail to learn how to do it without help. These beliefs likely stemmed from the AR conditioning effect and explains the attention lapses and errors recorded by participants in the non-faded (control) group. This supported H2. The attention lapses also provided some support for H4 but it is too weak to claim that it is supported.

#### 6.1.4. AR novice effect

The second type of expertise-redundancy effect was the AR novice effect, which suggests that worked-examples and guidance provided by an ITS reduces with growing expertise. For example, as reliance on external aids reduced and more practice was needed, we would expect to see errors increase (or fail to decrease) [49]. Our results support this view (RQ1.2). H2 was supported.

Therefore, adding a fading mechanism can help reduce (unconscious) errors performed by users (RQ1.2). This suggests that the worked-example instructional effectiveness can diminish as skills are developed as per our predictions from the AR novice and CLT expertise-reversal effects [39]. As problem-solving tasks compete for cognitive resources, the ability to effortlessly carry out the problem-solving task diminishes. This explains why greater errors were seen in the non-faded (control) group. This supports H2 and the theoretical explanations of CLT [12].

### 6.2. Cognitive mechanisms (RQ2)

An interaction effect between biologically-primary cognitive mechanisms and biologically-secondary mechanisms may explain why some users did not feel overwhelmed by the novel information (RQ1.3). We identified the following three potential cognitive mechanisms used in our experiment.

#### 6.2.1. Suppression of Irrelevant behaviour (RQ2.1)

Users were observed trying random processes at first but gradually changed their behaviour in response to cues from the environment. Two participants were observed plugging in a cable, stopping themselves and making a correction before proceeding to connect the cable. This shows that there was some use of self-regulatory abilities by users in the task. H4 was not supported.

One participant failed to suppress irrelevant behaviour, verbally commenting about having no idea what they were doing. One other participant plugged in cables by memorising the previous solution and failing to notice that their solution from the previous task was not generalisable to this task. This shows errors in transfer. However, since the effect was not isolated to a particular group, this finding does not support H4.

#### 6.2.2. Suppression of cognitive load awareness (RQ2.1)

Participants seemed to employ self-regulation mechanisms to suppress awareness of cognitive load to avoid fatigue. This may explain why the NASA-TLX findings were not significant in self-reports (H1 not supported). As a result, the participants may not be self-aware of the complex interactions to be able to describe it accurately in an introspective report [102].

#### 6.2.3. Suppression of relevant behaviour (RQ2.1)

Suppression of relevant behaviour occurs when appropriate and expected behaviour does not occur because self-regulatory mechanisms are used to suppress it. For example, one participant failed to notice that their behaviour was not producing desired outcomes (i.e. repeatedly pressing the repeat message button).

Learned metacognitive skills govern when to select these pre-existing skills versus learning new skills. Skills can be internalised and used inappropriately due to conditioning [106].

#### 6.2.4. Unconscious automation mechanisms (RQ2.2)

There was evidence of impaired unconscious automated mechanisms, which suggests that cognitive load may have interfered with schema automation. Both groups were able to acquire the knowledge initially because of the worked-example and AR effects. Practice does not aim to facilitate knowledge acquisition but schema automation. Cognitive load effects at this point would not impede the knowledge since it was already acquired. Rather, it would impede the ability to practice the skills. This may explain why H3 was not supported.

Another theoretical explanation for the lack of support for H3 might be because, according to Cooper and Sweller [14], increased errors in the AR non-faded group, can indicate limited ability to transfer knowledge due to high cognitive load. Even though the knowledge was known, it was not transformed into a higher form of knowledge that would be required for the kind of diagnostic work carried out on network equipment. Conversely, if transfer skills were present, the errors should be less because the learner would be able to adapt what they have learned so far to the new task [14].

### 6.3. Limitations

We identified the following limitations and assumptions with our experiment, which can help frame the basis for future work.

#### 6.3.1. Multifaceted nature of learning

The first limitation is the multifaceted nature of learning [107]. It has been known for over two decades that learning involves an interaction between cognitive and affective elements within the learner. This interaction can influence learning outcomes. For example, one limitation of CLT as a model is that it does not integrate learner emotional factors [80]. It is not necessarily easy to determine how these dimensions interact with one another. For example, perceptions of an environment can influence where one directs their attention, which could also account for attention lapses [108]. However, a recent study found that working memory and not subjective factors like task repetition, accounted for attention lapses [79].

In our experiment, it would be ideal to completely isolate all the potential factors that could influence learning whether they be usability limitations, subjective factors like emotional states or prior knowledge and experience. Further, we could have designed a system that focused just on the AR/ITS components. However, such a system is likely to be impossible to design because there are always going to be factors that interact with each other. This limitation arises because a training system comprises of a collection of components that are difficult to decouple.

### 6.3.2. Other theories

Our theoretical basis was grounded in Geary and Berch [32] and Sweller [27] evolutionary cognitive theories (i.e. cognitive load). However, there are other theories that could explain the increase in errors performed by users. For example, if AR was unusable due to registration errors, then removing AR would result in greater performance. These assumptions could also be explained under a cognitive load model. As working memory capacities reduce due to use, the ability to process and correct for misaligned AR cues would diminish, reducing working memory space even further. The AR Bias effect would explain this interpretation and is compatible with the guidance fading effects. Perhaps, more importantly, is that we would also expect to see degradation (or differences) in knowledge retention because the misalignment would interfere with the worked-example effect (see Chandler and Sweller [37],Ward and Sweller [109]). Yet, no such effect was found in our study.

Theories of flow and motivation could have also played a key role in whether someone regulated their attention and invested cognitive resources needed to solve the problem [110]. For example, task monotony and attention failures could be explained using flow theory. Our study did not decouple motivation from cognitive load effects.

Head and Helton [79] found that high cognitive load rather than boredom accounted for the errors in task environments. Similarly, the regulation of flow is governed by sense of presence. It is possible that presence was modified by AR and this led to deeper immersion in the task. These mechanisms are best explained by our unconscious cognitive mechanisms and are examples of the AR unconscious-stimulation effect.

These potential factors should not be conceived as alternate explanations for the effects observed in this experiment. Rather, they should be interpreted as supplementing the experimental findings. For example, it is possible that both motivation and expertise effects were present. Fading mechanisms could not only improve knowledge retention as per predictions made by CLT, but could also remove other factors such as redundant stimuli that interferes with performance. These explanations would provide weight for a new multifaceted learning framework in which cognitive load is but one dimension.

### 6.3.3. Limitations of the cognitive load model

CLT explains the effects observed in our experiment. Nevertheless, CLT is not a *"theory of everything"* [111]. It is not intended to apply to non-educational contexts [112]. In some domains which focus more on skill development, the theoretical explanations of CLT may not be useful for explaining how these skills are developed. In networking cabling, we made the assumption that instructional procedures were important because users must build up a knowledge base in long-term memory. CLT operates under these assumptions. Other factors outside of CLT were identified such as implicit learning [113], behaviour [106] and motivation [110].

### 6.3.4. Usability

The experiment did not use the System Usability Scale (SUS) [114] to measure UI usability. This meant that we could not isolate the usability from instructional design. For example, we could have correlated usability with performance and cognitive load. Since we did not measure the usability using the SUS, we cannot compare the design with similar UIs. The deciding factor for not using the SUS was based on a belief that too many questionnaires could bias the results due to a spacing effect [115].

### 6.3.5. Medical conditions

Our experiment did not screen for medical conditions (other than COVID-19). Even though we found no significant differences between groups, it could have been the case that if a difference was detected, it may have been due to an undetected medical condition. For example, colour blindness may have reduced the ability to differentiate between the arrows. Perhaps, one of the limitations is that even if we had screened participants, there would have been no guarantee that all possible medical conditions would be found. Further, screening is only useful insofar for detecting known medical conditions [95].

### 6.3.6. Measurement errors

During the study, four of the six NASA-TLX dimensions had to be discarded because the wrong data was recorded due to a software bug. This suggests important lessons of using paper when appropriate to record the results. Nevertheless, the limited NASA-TLX data does not harm the ideas presented in this paper because we draw conclusions based on their learning and performance, which was the aim of the paper rather than measuring cognitive load directly.

## 7. Conclusion and future work

From the literature, it was postulated that adding a guidance-fading mechanism to an AR-based ITS should improve its effectiveness as per predictions made by the CLT guidance-fading effect [49]. We implemented such an AR training system in the network cabling domain and tested the effectiveness of it by comparing it to an identical system without fading. Our experiment showed that, although adding a fading mechanism reduced errors, it had a limited effect on knowledge acquisition. Future work should explore the effectiveness of the worked-example instruction in novices to understand how it can better transfer knowledge.

Whereas, AR-based ITS have been shown to improve knowledge retention up to 25% in novices [18], experts benefit more from an AR-based training system without an ITS. Furthermore, both the AR detail and the amount of remediation steps should be reduced as expertise increases. Novices move along a continuum to gradually become experts [34]. Therefore, novices quickly become experts, so AR and ITS should adapt to the changing expertise. Without these considerations, an AR novice effect and AR conditioning effect is likely to result, reducing problem-solving performance and potentially interfering with learning.

This paper provided support for the AR Novice effect, which suggests, that the effectiveness of an ITS reduces in AR environments with increasing expertise; the AR conditioning effect, which suggests, that AR results in inattention for experts causing errors and finally, evidence for the AR integration principle, which suggests that AR likely facilitated knowledge acquisition but more work is needed to validate the repeatability and generality of this finding. Users reported increased reliance and perceived distraction by AR as they gained experience, further supporting an AR conditioning effect in the network cabling domain.

There were several other AR confounding effects derived in Section 2. These include: the AR unconscious-stimulation effect, the AR bias effect and the AR problem-solving effect. The AR acceleration effect was derived out of the empirical findings of this study to explain how limited training could account for the expertise increase. Future work should explore the robustness of these effects. For example, an experiment might explore reversing the fading by starting without AR and progressively adding it to see if a AR bias effect results.

The first next step should be to decouple worked-example instruction from an AR-based ITS to demonstrate that worked-example instruction can provide knowledge transfer whereas without it, transfer will be limited. AR-based ITS experiments so far have lacked a comparison between worked-example versus no worked-example. Yet, the effects discovered in our experiment motivate the need for such an experiment.

The second next step is the incorporation of adaptive fading support into the ITS itself [49]. This has been done in non-AR domains but not in ITSs that use AR. It is possible that the instruction was faded out too soon, causing some participants to become easily confused. This motivates the next study to focus on adaptive fading. For example, using predictive models to categorise learners into novice and expert groups based on this foreknowledge. There is some work in VR training systems, but no such work in AR training domains [116].

The third next step is to optimise the UI to ensure consistent results across future experiments. In our experiment, the UI itself was novel because it was based on CLT principles. There was no data to which to base our UI design for network cabling. The next experiments could do various comparisons of similar UIs to understand which part(s) need to be optimised.

Finally, this work is of interest to AR educational researchers, Human-Computer Interaction (HCI) researchers, AIED researchers and educators looking to understand how hands-on domains can be affected by the use of the same technology but which uses different instructional methods.

## CRediT authorship contribution statement

**Bradley Herbert:** Conceptualisation, Investigation, Software, Data curation, Formal analysis. **Grant Wigley:** Methodology, Resources. **Barrett Ens:** Methodology, Software. **Mark Billinghurst:** Resources, Writing – review & editing, Supervision.

## Acknowledgements

We would like to acknowledge that the work in this paper was supported by a government-sponsored scholarship under the Australian Research Training Program (RTP). We would also like to acknowledge the reviewers for their time in reviewing this work.

## Funding

Partially funded by an Australian government-supported Research Training Program (RTP) Scholarship.

## Appendix A. Administered questionnaires

These questionnaires were administered after each phase (continued next page).

### A.1. Demographics

## Participant Demographic Information

**Gender:** ☐ Male ☐ Female ☐ Other

**Age Range:** 18-24 25-40 41-60 61+

**Have you used Handheld Augmented Reality interfaces before?**

☐ Never

☐ A few times a year

☐ A few times a month

☐ A few times a week

☐ Everyday

**How familiar are you with using Augmented Reality interfaces?**

| **Very Unfamiliar 1** | **2** | **3** | **4** | **5** | **6** | **Very Familiar 7** |
|---|---|---|---|---|---|---|
| | | | | | | |

**How experienced are you with network cabling?**

| **Beginner 1** | **2** | **3** | **4** | **5** | **6** | **Expert 7** |
|---|---|---|---|---|---|---|
| | | | | | | |

**How experienced are you with VLAN concepts?**

| **Beginner 1** | **2** | **3** | **4** | **5** | **6** | **Expert 7** |
|---|---|---|---|---|---|---|
| | | | | | | |

*A.2. Exam*

## Knowledge Exam

1. A network engineer connects port FastEthernet0/17 on switch 1 to GigabitEthernet0/2 on switch 2, but they fail to communicate. Which of the following is the most likely cause? (**5 points**) (Knowledge Comprehension)

   a) GigabitEthernet0/2 is assigned to VLAN 1 and FastEthernet0/17 is also assigned to VLAN 1. Due to this, there is a conflict and no traffic can flow between the two ports.

   b) An Ethernet straight-through cable has been used to connect the two points, but a cross over cable should be used instead.

   c) GigabitEthernet0/2 operates at 1,000 mb/s and FastEthernet0/17 at 100 mb/s. The speeds must match for successful communication.

   d) FastEthernet0/17 is assigned to VLAN 1, but GigabitEthernet0/2 is assigned to VLAN 5. Ports within different VLANs will not logically communicate.

2. A network engineer is tasked with connecting an Ethernet cable from any port on switch 1 to any port on switch 2. However, the network engineer reports that there is no suitable port on switch 2 to connect to. Which of the following is the most likely cause? (**10 points**) (Domain Knowledge)

   a) An Ethernet straight-through cable has been used to connect the two points, but a cross over cable should be used instead.

   b) Switch only has ports, which support 10/100 mb/s that is not compatible with the 100/1000 ports on switch 1.

   c) There is no port on switch 2 assigned to VLAN 99.

   d) The switch port on switch 1 is assigned to a VLAN that no port on switch 2 is assigned to.

3. Refer to the table. Which port on switch 2 should port Fa0/7 on switch 1 be connected to? (**10 points**) (Domain Knowledge)

| | | |
|---|---|---|
| **Switch 1** | **VLAN 12**<br>Fa0/7, Fa0/8<br>Fa0/9 | **VLAN 98**<br>Fa0/20 |
| **Switch 2** | **VLAN 12**<br>Fa0/11 | **VLAN 98**<br>Fa0/12 |

4. Refer to the table. Which port should Fa0/23 on switch 2 can be connected to. (**10 points**) (Domain Knowledge)

| | | |
|---|---|---|
| **Switch 1** | **VLAN 12**<br>Fa0/7, Fa0/8<br>Fa0/9 | **VLAN 98**<br>Fa0/23,<br>Fa0/24,<br>Fa0/17 |
| **Switch 2** | **VLAN 12**<br>Fa0/16,<br>Fa0/18,<br>Fa0/19, Fa0/7 | **VLAN 98**<br>Fa0/12,<br>Fa0/22,<br>Fa0/24,<br>Fa0/1, Fa0/23 |

5. Which connection listed below is the most optimal to connect **(5 points)**: (Domain Knowledge)

a) FastEthernet0/7 (S1) to FastEthernet0/9 (S4)

b) FastEthernet0/8 (S1) to FastEthernet0/12 (S2)

c) GigabitEthernet0/1 (S2) to FastEthernet0/5 (S1)

d) GigabitEthernet0/2 (S1) to GigabitEthernet0/2 (S3)

6. Refer to the table. Which port should Fa0/24 on switch 2 can be connected to (**20 points**) (Knowledge comprehension)

| | | | | |
|---|---|---|---|---|
| **Switch 1** | **VLAN 1**<br>Fa0/1, Fa0/2,<br>Fa0/3, Fa0/4,<br>Fa0/5, Fa0/24 | **VLAN 2**<br>Fa0/6, Fa0/7,<br>Fa08, Fa0/9,<br>Fa0/10, Fa0/11 | **VLAN 4**<br>Fa0/12, Fa0/13,<br>Fa0/14, Fa0/15,<br>Fa0/16, Fa0/17, | **VLAN 5**<br>Fa0/18, Fa0/19,<br>Fa0/20, Fa0/21,<br>Fa0/22, Fa0/23 |
| **Switch 2** | **VLAN 11**<br>Fa0/1, Fa0/2,<br>Fa0/3, Fa0/4,<br>Fa0/5, Fa0/24 | **VLAN 12**<br>Fa0/6, Fa0/7,<br>Fa08, Fa0/9,<br>Fa0/10, Fa0/11 | **VLAN 14**<br>Fa0/12, Fa0/13,<br>Fa0/14, Fa0/15,<br>Fa0/16, Fa0/17, | **VLAN 15**<br>Fa0/18, Fa0/19,<br>Fa0/20, Fa0/21,<br>Fa0/22, Fa0/23 |

7. List one advantage of creating multiple VLANs across a physical switched network (**10 points):** (Knowledge Application)

8. List one disadvantage of creating multiple VLANs across a physical switched network (**10 points**): (Knowledge Application)

**Total Score: / 80**

### A.3. Subjective measures

#### A.3.1. Learner experience survey

Learner Experience Survey (After Baseline Task)

1. Please rate how much you disagree / agree with the following (Metacognitive knowledge):

a) You can read up on the concepts easily, but struggle to physically translate them when performing network cabling on the real-world.

| **Strongly Disagree 1** | **2** | **3** | **4** | **5** | **6** | **Strongly Agree 7** |
|---|---|---|---|---|---|---|
| | | | | | | |

b) Although I have a reasonable understanding of network cabling, I frequently find myself plugging in the wrong cable, or other minor errors when doing it on real equipment.

| **Strongly Disagree 1** | **2** | **3** | **4** | **5** | **6** | **Strongly Agree 7** |
|---|---|---|---|---|---|---|
| | | | | | | |

c) I feel easily stressed or disorientated when trying to do for real on actual equipment.

| **Strongly Disagree 1** | **2** | **3** | **4** | **5** | **6** | **Strongly Agree 7** |
|---|---|---|---|---|---|---|
| | | | | | | |

d) I can quickly and efficiently connect cables correctly on the real equipment, but given a theory test on it, I struggle to recall what I did, or when to apply the concepts.

| **Strongly Disagree 1** | **2** | **3** | **4** | **5** | **6** | **Strongly Agree 7** |
|---|---|---|---|---|---|---|
| | | | | | | |

e) Even when someone has high theoretical knowledge, they need to be trained on doing it for real on actual equipment, or else, their knowledge will be forgotten.

| **Strongly Disagree 1** | **2** | **3** | **4** | **5** | **6** | **Strongly Agree 7** |
|---|---|---|---|---|---|---|
| | | | | | | |

f) Practicing on the real-world equipment helps some people better grasp the theory behind it but is not for all learners because some can grasp it from the theory alone.

| **Strongly Disagree 1** | **2** | **3** | **4** | **5** | **6** | **Strongly Agree 7** |
|---|---|---|---|---|---|---|
| | | | | | | |

g) I, personally, find that practicing on real world equipment is limited because I already have enough understanding from the theory.

| **Strongly Disagree 1** | **2** | **3** | **4** | **5** | **6** | **Strongly Agree 7** |
|---|---|---|---|---|---|---|
| | | | | | | |

2. Which of the following strategies do you use to learn network cabling (Select only one answer that closest matches your preference) (Metacognitive knowledge)?

a. I immediately look up the information I need in a text book or online.

b. I immediately ask a friend, tutor or teacher for help

c. I experiment and try a variety of things, usually through trial-and-error until I get the feedback or verification that the answer, or solution is correct.

d. I try to solve the problem by myself, using the strategies and techniques I have been taught, and only then, do I ask for help.

3. Which learning strategy would be the least effective to use when learning network cabling?

a. Studying examples of previously cabled topologies to try and infer differences between them and learn the underlying principles.

b. Being instructed what strategies to use by a teacher, tutor or an expert.

c. Being instructed as to what the answer or solution should be and then learning the steps from the answer(s) given.

**NASA TLX (After Baseline Task)**

To complete the instrument, place a handwritten cross, either on or in-between the lines (according to your preference). Avoid overthinking the answer. You should specify what immediately comes to mind.

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low — Very High

## Appendix B. EtherGuide Example domain model

```
<KnowledgeModel>
   <String name="name" default="" value="Unknown" Constraint="IsNotValue" />
   <Int name="vid0" default="0" value="0" Constraint="GreaterThan"
    tag="InvalidVlanID" />
   <Int name="vid1" default="0" value="vid0" Constraint="IsValue"
   tag="vid1" />
   <String name="switch0" default="" value="S" Constraint="IsNotValue"
   tag="InvalidSwitch" />
   <String name="switch1" default="switch" value="switch0"
   Constraint="IsNotValue" tag="switch" />
   <Int name="speed0" default="0" value="0" Constraint="GreaterThan"
   tag="InvalidSpeed" />
   <Int name="speed1" default="10" value="speed0" Constraint="IsValue"
   tag="speed" />
 ...
</KnowledgeModel>
```

**Table D.4**
Domain scores.

| Item | AR non-faded | AR faded |
|---|---|---|
| Baseline knowledge score | 29% | 27% |
| Post knowledge | 36% | 38% |
| Domain knowledge gain | +7% | +9% |
| **Baseline SD** | 21 | 27 |
| **Post SD** | 31 | 34 |
| **Gain SD** | 25 | 33 |
| **Baseline/Post p** | .51 | .36 |
| **Baseline/Post t** | −0.67 | |
| **Gain p** | .84 | |
| **Gain t** | −0.2 | |

* Significant result ($p < 0.05$).

**Table D.5**
Application of knowledge.

| Item | AR non-faded | AR faded |
|---|---|---|
| Baseline application score | 40% | 50% |
| Post application score | 46% | 30% |
| Application score gain | +7% | −20% |
| **Baseline SD** | 45 | 41 |
| **Post SD** | 43 | 36 |
| **Gain SD** | 17 | 36 |
| **Baseline/Post p** | .157 | .058 |
| **Baseline/Post T** | −1.41 | 3.5 |
| **Baseline/Post Z** | −1.41 | −190 |
| **Gain p** | .09 | |
| **Gain z** | 1.68 | |

* Significant result ($p < 0.05$).

## Appendix C. Pedagogical example models

Some example XML structures, which could model pedagogical characteristics.

```
<PedagogicalModel>
  <Pedagogy type="Standard">
    <PedagogyName>Standard</PedagogyName>
    <PedagogyGroup>Standard</PedagogyGroup>
    <VisualizeCorrectActionInitially>
      False
    </VisualizeCorrectActionInitially>
    <VisualizeCorrectActionOnError>
      True
    </VisualizeCorrectActionOnError>
  </Pedagogy>
</PedagogicalModel>
```

## Appendix D. Additional results

Some additional results are reported.

### D.1. Knowledge acquisition

See Table D.4.

### D.2. Performance measures

See Fig. D.9 and Tables D.5–D.8.

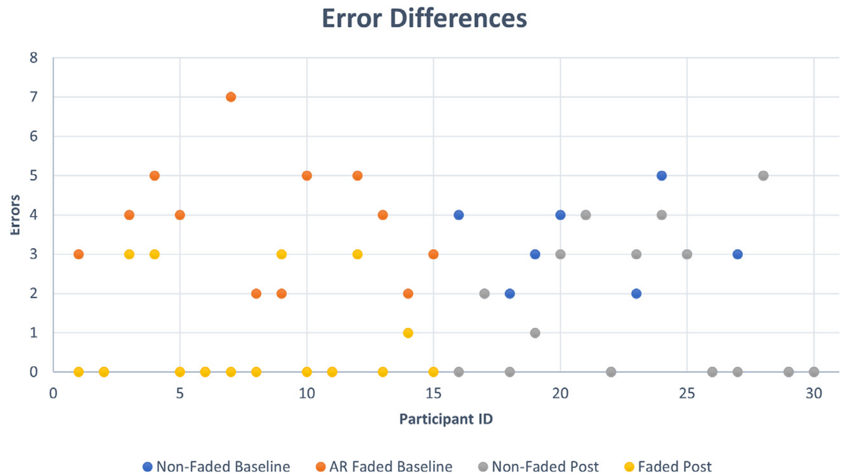



**Fig. D.9.** Shows the Error Differences between the two conditions. Both the baseline and post assessment for each condition is shown.

**Table D.6**
Completion times.

| Item | AR non-faded | AR faded |
|---|---|---|
| Baseline completion seconds | 331 | 384 |
| Post completion seconds | 216 | 175 |
| Completion seconds gain | −115 | −209 |
| **Baseline SD** | 100 | 143 |
| **Post SD** | 159 | 53 |
| **Gain SD** | 182 | 136 |
| **Baseline/Post p** | .03* | < .01* |
| **Baseline/Post t** | −2.37 | −5.73 |
| **Gain p** | .15 | |
| **Gain t** | 1.51 | |

*Significant result ($p < 0.05$).
Decreased time would be preferred.

**Table D.7**
NASA TLX performance scores.

| Item | AR non-faded | AR faded |
|---|---|---|
| Baseline performance | 36 | 23 |
| Post performance | 61 | 65 |
| Performance gain | 25 | 41 |
| **Baseline SD** | 33 | 32 |
| **Post SD** | 35 | 27 |
| **Gain SD** | 31 | 33 |
| **Baseline/Post p** | .02* | $p < .01$ |
| **Baseline/Post T** | 79.5 | 120 |
| **Gain p** | .15 | |
| **Gain z** | 1.45 | |

*Significant result ($p < 0.05$).

**Table D.8**
NASA TLX mental demand scores.

| Item | AR non-faded | AR faded |
|---|---|---|
| Baseline mental demand | 51 | 55 |
| Post mental demand | 46 | 50 |
| Mental demand gain | −4 | −5 |
| **Baseline SD** | 33 | 28 |
| **Post SD** | 19 | 23 |
| **Gain SD** | 43 | 22 |
| **Baseline/Post p** | .73 | .40 |
| **Baseline/Post Z** | −0.34 | −0.84 |
| **Baseline/Post T** | 54 | 33.5 |
| **Gain p** | .95 | |
| **Gain Z** | −0.06 | |

* Significant result ($p < 0.05$).

## References


[1] Ye C, Indra PP, Aspinall D. Retrofitting security and privacy measures to smart home devices. In: 2019 sixth international conference on internet of things: systems, management and security (IOTSMS). 2019, p. 283–90. http://dx.doi.org/10.1109/IOTSMS48152.2019.8939272.

[2] Nishino H, Nagatomo Y, Kagawa T, Haramaki T. A mobile AR assistant for campus area network management. In: 2014 eighth international conference on complex, intelligent and software intensive systems. 2014, p. 643–8. http://dx.doi.org/10.1109/CISIS.2014.93.

[3] Konak A, Clark TK, Nasereddin M. Using Kolb's experiential learning cycle to improve student learning in virtual computer laboratories. Comput Educ 2014;72(Supplement C):11–22. http://dx.doi.org/10.1016/j.compedu.2013.10.013.

[4] Azuma RT. A survey of augmented reality. Presence: Teleoperators Virtual Environ 1997;6(4):355–85.

[5] Dennis M, Masthoff J, Mellish C. Adapting progress feedback and emotional support to learner personality. Int J Artif Intell Educ 2016;26(3):877–931. http://dx.doi.org/10.1007/s40593-015-0059-7.

[6] Westerfield G. Intelligent augmented reality training for assembly and maintenance. New Zealand: University of Canterbury; 2012.

[7] Herbert B, Ens B, Billinghurst M. An adaptive AR tutor for cabling a network topology. In: Eurographics symposium on virtual environments (2017). Adelaide, Australia: The Eurographics Association; 2017, p. 11–2. http://dx.doi.org/10.2312/egve.20171369.

[8] Herbert B, Ens B, Weerasinghe A, Billinghurst M, Wigley G. Design considerations for combining augmented reality with intelligent tutors. Comput Graph 2018;77:166–82.

[9] Retnowati E. Faded-example as a tool to acquire and automate mathematics knowledge. J Phys Conf Ser 2017;824:012054. http://dx.doi.org/10.1088/1742-6596/824/1/012054.

[10] Sweller J. Cognitive load theory and educational technology. Educ Technol Res Dev 2020;68(1):1–16. http://dx.doi.org/10.1007/s11423-019-09701-3.

[11] Sweller J, Sweller S. Natural information processing systems. Evol Psychol 2006;4(1):147470490600400135. http://dx.doi.org/10.1177/147470490600400135.

[12] Sweller J. Cognitive load during problem solving: effects on learning. Cogn Sci 1988;12(2):257–85. http://dx.doi.org/10.1016/0364-0213(88)90023-7.

[13] Sweller J, van Merrienboer JJG, Paas FGWC. Cognitive architecture and instructional design. Educ Psychol Rev 1998;10(3):251–96. http://dx.doi.org/10.1023/A:1022193728205.

[14] Cooper G, Sweller J. Effects of schema acquisition and rule automation on mathematical problem-solving transfer. J Educ Psychol 1987;79(4):347–62. http://dx.doi.org/10.1037/0022-0663.79.4.347.

[15] Kalyuga S. Expertise reversal effect and its implications for learner-tailored instruction. Educ Psychol Rev 2007;19(4):509–39. http://dx.doi.org/10.1007/s10648-007-9054-3.

[16] Akçayır M, Akçayır G. Advantages and challenges associated with augmented reality for education: A systematic review of the literature. Educ Res Rev 2017;20:1–11. http://dx.doi.org/10.1016/j.edurev.2016.11.002.

[17] Haramaki T, Nishino H. A device identification method for AR-based network topology visualization. In: 2015 10th international conference on broadband and wireless computing, communication and applications (BWCCA). 2015, p. 255–62. http://dx.doi.org/10.1109/BWCCA.2015.81.

[18] Westerfield G, Mitrovic A, Billinghurst M. Intelligent augmented reality training for motherboard assembly. Int J Artif Intell Educ 2015;25(1):157–72. http://dx.doi.org/10.1007/s40593-014-0032-x.

[19] LaViola J, Williamson B, Brooks C, Veazanchin S, Sottilare R, Garrity P. Using augmented reality to tutor military tasks in the wild. In Interservice/industry training simulation and education conference, Orlando, FL, USA, Vol. 15050. 2015. p. 1–10.

[20] Menozzi M, Ropelat S, Köfler J, Huang Y-Y. Development of ophthalmic microsurgery training in augmented reality. Klin Monatsbl Augenheilkd 2020;237(04):388–91.

[21] Herbert B, Wigley G, Ens B, Billinghurst M. Perceptions of integrating augmented reality into network cabling tutors. In: 2020 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-adjunct). IEEE; 2020, p. 231–6. http://dx.doi.org/10.1109/ISMAR-Adjunct51615.2020.00068.

[22] Herbert B, Hoff W, Billinghurst M. Usability considerations of hand held augmented reality wiring tutors. In: 2020 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-adjunct). IEEE; 2020, p. 277–82. http://dx.doi.org/10.1109/ISMAR-Adjunct51615.2020.00078.

[23] Kalyuga S, Renkl A. Expertise reversal effect and its instructional implications: Introduction to the special issue. Instruct Sci 2010;38(3):209–15. http://dx.doi.org/10.1007/s11251-009-9102-0.

[24] Kissane M, Kalyuga S, Chandler P, Sweller J. The consequences of fading instructional guidance on delayed performance: The case of financial services training. Educ Psychol 2008;28(7):809–22. http://dx.doi.org/10.1080/01443410802322069.

[25] Seufert T. The interplay between self-regulation in learning and cognitive load. Educ Res Rev 2018;24:116–29. http://dx.doi.org/10.1016/j.edurev.2018.03.004.

[26] Geary DC. Reflections of evolution and culture in children's cognition: implications for mathematical development and instruction. Am Psychol 1995;50(1):24–37. http://dx.doi.org/10.1037/0003-066X.50.1.24.

[27] Sweller J. Cognitive load theory. In: Mestre JP, Ross BH, editors. Psychology of learning and motivation. Vol. 55, Academic Press; 2011, p. 37–76. http://dx.doi.org/10.1016/B978-0-12-387691-1.00002-8.

[28] Sweller J. Human cognitive architecture. In: Handbook of research on educational communications and technology. AECT series, third ed.. Boca Raton, FL: Routledge; 2008, p. 369–81.

[29] Paas F, Sweller J. An evolutionary upgrade of cognitive load theory: using the human motor system and collaboration to support the learning of complex cognitive tasks. Educ Psychol Rev 2012;24(1):27–45. http://dx.doi.org/10.1007/s10648-011-9179-2.

[30] Miller GA. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychol Rev 1956;63(2):81–97. http://dx.doi.org/10.1037/h0043158.

[31] Peterson L, Peterson MJ. Short-term retention of individual verbal items. J Exp Psychol 1959;58(3):193–8. http://dx.doi.org/10.1037/h0049234.

[32] Geary DC, Berch DB. Evolution and children's cognitive and academic development. In: Geary DC, Berch DB, editors. Evolutionary perspectives on child development and education. Evolutionary psychology, Cham: Springer International Publishing; 2016, p. 217–49. http://dx.doi.org/10.1007/978-3-319-29986-0_9.

[33] Sweller J. What human cognitive architecture tells us about constructivism. In: Constructivist instruction : success or failure?. 1st ed.. London, United Kingdom: Taylor & Francis Group; 2009, p. 127–43.

[34] Sweller J. Evolution of human cognitive architecture. In: Psychology of learning and motivation. Vol. 43, Academic Press; 2003, p. 215–66. http://dx.doi.org/10.1016/S0079-7421(03)01015-6.

[35] Sweller J. The worked example effect and human cognition. Learn Instruct 2006;16(2):165–9. http://dx.doi.org/10.1016/j.learninstruc.2006.02.005.

[36] Ashman G, Kalyuga S, Sweller J. Problem-solving or explicit instruction: which should go first when element interactivity is high? Educ Psychol Rev 2020;32(1):229–47. http://dx.doi.org/10.1007/s10648-019-09500-5.

[37] Chandler P, Sweller J. The split-attention effect as a factor in the design of instruction. Br J Educ Psychol 1992;62(2):233–46. http://dx.doi.org/10.1111/j.2044-8279.1992.tb01017.x.

[38] Tindall-Ford S, Chandler P, Sweller J. When two sensory modes are better than one. J Exp Psychol: Appl 1997;3(4):257–87. http://dx.doi.org/10.1037/1076-898X.3.4.257.

[39] Kalyuga S, Ayres P, Chandler P, Sweller J. The expertise reversal effect. Educ Psychol 2003;38(1):23–31. http://dx.doi.org/10.1207/S15326985EP3801_4.

[40] Sweller J, van Merriënboer JJG, Paas F. Cognitive architecture and instructional design: 20 years later. Educ Psychol Rev 2019;31(2):261–92. http://dx.doi.org/10.1007/s10648-019-09465-5.

[41] Nwana HS. Intelligent tutoring systems: An overview. Artif Intell Rev 1990;4(4):251–77.

[42] Latham A, Crockett K, McLean D, Edmonds B. A conversational intelligent tutoring system to automatically predict learning styles. Comput Educ 2012;59(1):95–109. http://dx.doi.org/10.1016/j.compedu.2011.11.001.

[43] D'mello S, Graesser A. AutoTutor and affective autotutor: learning by talking with cognitively and emotionally intelligent computers that talk back. ACM Trans Interact Intell Syst 2013;2(4):23:1–39. http://dx.doi.org/10.1145/2395123.2395128.

[44] VanLehn K, Lynch C, Schulze K, Shapiro JA, Shelby R, Taylor L, et al. The andes physics tutoring system: lessons learned. Int J Artif Intell Educ 2005;15(3):147–204.

[45] Salden RJCM, Aleven VAWMM, Renkl A, Schwonke R. Worked examples and tutored problem solving: redundant or synergistic forms of support? Top Cogn Sci 2009;1(1):203–13. http://dx.doi.org/10.1111/j.1756-8765.2008.01011.x.

[46] Najar AS, Mitrovic A, McLaren BM. Adaptive support versus alternating worked examples and tutored problems: which leads to better learning? In: Dimitrova V, Kuflik T, Chin D, Ricci F, Dolog P, Houben G-J, editors. User modeling, adaptation, and personalization. Lecture notes in computer science, Cham: Springer International Publishing; 2014, p. 171–82. http://dx.doi.org/10.1007/978-3-319-08786-3_15.

[47] Lee HS, Anderson JR. Student learning: what has instruction got to do with it? Annu Rev Psychol 2013;64(1):445–69. http://dx.doi.org/10.1146/annurev-psych-113011-143833.

[48] Leahy W, Sweller J. Cognitive load and the imagination effect. Appl Cogn Psychol 2004;18(7):857–75. http://dx.doi.org/10.1002/acp.1061.

[49] Renkl A, Atkinson RK. Structuring the transition from example study to problem solving in cognitive skill acquisition: a cognitive load perspective. Educ Psychol 2003;38(1):15–22. http://dx.doi.org/10.1207/S15326985EP3801_3.

[50] Renkl A, Atkinson RK, Große CS. How fading worked solution steps works – a cognitive load perspective. Instruct Sci 2004;32(1):59–82. http://dx.doi.org/10.1023/B:TRUC.0000021815.74806.f6.
[51] Billinghurst M, Clark A, Lee G. A survey of augmented reality. Found Trends Hum-Comput Interact 2015;8(2–3):73–272. http://dx.doi.org/10.1561/1100000049.
[52] Billinghurst M. Augmented reality in education. New Horiz Learn 2002;12.
[53] Dayagdag CV, Catanghal RA, Palaoag TD. Improving vocational training in the Philippines using AR. In: Proceedings of the 8th international conference on informatics, environment, energy and applications. IEEA '19, Osaka, Japan: Association for Computing Machinery; 2019, p. 253–7. http://dx.doi.org/10.1145/3323716.3323755.
[54] Craik K. The nature of explanation. CUP Archive; 1952.
[55] Al-Diban S. Mental models. Encycl Sci Learn 2012;2200–4. http://dx.doi.org/10.1007/978-1-4419-1428-6_586.
[56] Baumeister J, Ssin SY, ElSayed NAM, Dorrian J, Webb DP, Walsh JA, et al. Cognitive cost of using augmented reality displays. IEEE Trans Vis Comput Graphics 2017;23(11):2378–88. http://dx.doi.org/10.1109/TVCG.2017.2735098.
[57] Bruner JS. The act of discovery. Harvard Educ Rev 1961;31:21–32.
[58] Duffy TM, Cunningham DJ. Constructivism: implications for the design and delivery of instruction. In: Jonassen D, editor. Handbook of research for educational communications and technology. New York, NY, USA; 1996, p. 170–98.
[59] Piaget J. Piaget's theory. In: Inhelder B, Chipman HH, Zwingmann C, editors. Piaget and his school: a reader in developmental psychology. Springer study edition, Berlin, Heidelberg: Springer; 1976, p. 11–23. http://dx.doi.org/10.1007/978-3-642-46323-5_2.
[60] Ibáñez M-B, Delgado-Kloos C. Augmented reality for STEM learning: A systematic review. Comput Educ 2018;123:109–23. http://dx.doi.org/10.1016/j.compedu.2018.05.002.
[61] Henderson SJ, Feiner S. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In: 2009 8th IEEE international symposium on mixed and augmented reality. 2009, p. 135–44. http://dx.doi.org/10.1109/ISMAR.2009.5336486.
[62] Vargas González AN, Kapalo K, Koh SL, LaViola JJ. Exploring the virtuality continuum for complex rule-set education in the context of soccer rule comprehension. Multimodal Technol Interact 2017;1(4):30. http://dx.doi.org/10.3390/mti1040030.
[63] Morillo P, García-García I, Orduña JM, Fernández M, Juan MC. Comparative study of AR versus video tutorials for minor maintenance operations. Multimedia Tools Appl 2020;79(11):7073–100. http://dx.doi.org/10.1007/s11042-019-08437-9.
[64] Lin H-CK, Chen M-C, Chang C-K. Assessing the effectiveness of learning solid geometry by using an augmented reality-assisted learning system. Interact Learn Environ 2015;23(6):799–810. http://dx.doi.org/10.1080/10494820.2013.817435.
[65] Radkowski R, Herrema J, Oliver J. Augmented reality-based manual assembly support with visual features for different degrees of difficulty. Int J Hum Comput Interact 2015;31(5):337–49. http://dx.doi.org/10.1080/10447318.2014.994194.
[66] Henderson S, Feiner S. Exploring the benefits of augmented reality documentation for maintenance and repair. IEEE Trans Vis Comput Graphics 2011;17(10):1355–68. http://dx.doi.org/10.1109/TVCG.2010.245.
[67] Hou Lei, Wang Xiangyu, Bernold Leonhard, Love Peter ED. Using animated augmented reality to cognitively guide assembly. J Comput Civ Eng 2013;27(5):439–51. http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000184.
[68] Tsuruzoe H, Odera S, Shigeno H, Okada K. MR work supporting system using pepper's ghost. In: Proceedings of the 26th international conference on artificial reality and telexistence and the 21st eurographics symposium on virtual environments. ICAT-EGVE '16, Goslar Germany, Germany: Eurographics Association; 2016, p. 61–8. http://dx.doi.org/10.2312/egve.20161436.
[69] Henderson SJ, Feiner SK. Augmented reality in the psychomotor phase of a procedural task. In: 2011 10th IEEE international symposium on mixed and augmented reality. 2011, p. 191–200. http://dx.doi.org/10.1109/ISMAR.2011.6092386.
[70] Funk M, Kosch T, Schmidt A. Interactive worker assistance: Comparing the effects of in-situ projection, head-mounted displays, tablet, and paper instructions. In: Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing. UbiComp '16, Heidelberg, Germany: Association for Computing Machinery; 2016, p. 934–9. http://dx.doi.org/10.1145/2971648.2971706.
[71] Thees M, Kapp S, Strzys MP, Beil F, Lukowicz P, Kuhn J. Effects of augmented reality on learning and cognitive load in university physics laboratory courses. Comput Hum Behav 2020;108:106316. http://dx.doi.org/10.1016/j.chb.2020.106316.
[72] Wiedenmaier S, Oehme O, Schmidt L, Luczak H. Augmented reality (AR) for assembly processes design and experimental evaluation. Int J Hum Comput Interact 2003;16(3):497–514. http://dx.doi.org/10.1207/S15327590IJHC1603_7.
[73] Gavish N, Gutiérrez T, Webel S, Rodríguez J, Peveri M, Bockholt U, et al. Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. Interact Learn Environ 2015;23(6):778–98. http://dx.doi.org/10.1080/10494820.2013.815221.
[74] Norman DA. Categorization of action slips. Psychol Rev 1981;88(1):1–15. http://dx.doi.org/10.1037/0033-295X.88.1.1.
[75] Schomaker J, Meeter M. Short- and long-lasting consequences of novelty, deviance and surprise on brain and cognition. Neurosci Biobehav Rev 2015;55:268–79. http://dx.doi.org/10.1016/j.neubiorev.2015.05.002.
[76] Lesgold A, et al. Sherlock: A coached practice environment for an electronics troubleshooting job. 1988, https://eric.ed.gov/?id=ED299450.
[77] Baddeley A. Working memory and language: An overview. J Commun Disorders 2003;36(3):189–208. http://dx.doi.org/10.1016/S0021-9924(03)00019-4.
[78] Skinner BF. Superstition in the pigeon. J Exp Psychol 1948;38(2):168–72. http://dx.doi.org/10.1037/h0055873.
[79] Head J, Helton WS. Sustained attention failures are primarily due to sustained cognitive load not task monotony. Acta Psychol 2014;153:87–94. http://dx.doi.org/10.1016/j.actpsy.2014.09.007.
[80] Plass JL, Kalyuga S. Four ways of considering emotion in cognitive load theory. Educ Psychol Rev 2019;31(2):339–59. http://dx.doi.org/10.1007/s10648-019-09473-5.
[81] Schnotz W, Kürschner C. A reconsideration of cognitive load theory. Educ Psychol Rev 2007;19(4):469–508. http://dx.doi.org/10.1007/s10648-007-9053-4.
[82] Kirschner PA, Sweller J, Clark RE. Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educ Psychol 2006;41(2):75–86. http://dx.doi.org/10.1207/s15326985ep4102_1.
[83] Ohlsson S. Constraint-based student modelling. J Artif Intell Educ 1992;3(4):429–47.
[84] Ohlsson S. Constraint-based student modeling. In: Greer JE, McCalla GI, editors. Student modelling: the key to individualized knowledge-based instruction. NATO ASI Series, Springer Berlin Heidelberg; 1994, p. 167–89.
[85] Ohlsson S. Constraint-based modeling: from cognitive theory to computer tutoring – and back again. Int J Artif Intell Educ 2016;26(1):457–73. http://dx.doi.org/10.1007/s40593-015-0075-7.
[86] Ma W, Adesope OO, Nesbit JC, Liu Q. Intelligent tutoring systems and learning outcomes: A meta-analysis. J Educ Psychol 2014;106(4):901–18. http://dx.doi.org/10.1037/a0037123.
[87] Fossati D. The role of positive feedback in intelligent tutoring systems. In: Proceedings of the 46th annual meeting of the association for computational linguistics on human language technologies: student research workshop. HLT-SRWS '08, Stroudsburg, PA, USA: Association for Computational Linguistics; 2008, p. 31–6.
[88] Schröder M, Trouvain J. The german text-to-speech synthesis system MARY: A tool for research, development and teaching. Int J Speech Technol 2003;6(4):365–77. http://dx.doi.org/10.1023/A:1025708916924, https://doi.org/10.1023/A:1025708916924.
[89] Mayer RE. Multimedia learning. In: Psychology of learning and motivation. vol. 41, Academic Press; 2002, p. 85–139. http://dx.doi.org/10.1016/S0079-7421(02)80005-6.
[90] Technologies U. Unity real-time development platform | 3D, 2D VR & AR engine. 2021, https://unity.com/.
[91] Microsoft Inc. Visual studio IDE, code editor, azure DevOps, & app center - visual studio. 2021, https://visualstudio.microsoft.com/.
[92] Blender Foundation, Home of the Blender project - Free and Open 3D Creation Software, https://www.blender.org/, [n.d.].
[93] PTC Inc., Vuforia: Getting Started, https://library.vuforia.com/getting-started/overview.html, n.d..
[94] Spencer SJ, Logel C, Davies PG. Stereotype threat. Annu Rev Psychol 2016;67(1):415–37. http://dx.doi.org/10.1146/annurev-psych-073115-103235.
[95] Harris AD, McGregor JC, Perencevich EN, Furuno JP, Zhu J, Peterson DE, et al. The use and interpretation of quasi-experimental studies in medical informatics. J Am Med Inform Assoc 2006;13(1):16–23. http://dx.doi.org/10.1197/jamia.M1749.
[96] Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock PA, Meshkati N, editors. Advances in psychology. Human mental workload, vol. 52, North-Holland; 1988, p. 139–83. http://dx.doi.org/10.1016/S0166-4115(08)62386-9.
[97] Grier RA. How high is high? a meta-analysis of NASA-tlx global workload scores. In: Proceedings of the human factors and ergonomics society annual meeting. Vol. 59, (1):2015, p. 1727–31. http://dx.doi.org/10.1177/1541931215591373.

[98] Molden DC. Understanding priming effects in social psychology: what is "social priming" and how does it occur? Soc Cogn 2014;32(Supplement):1–11. http://dx.doi.org/10.1521/soco.2014.32.supp.1.
[99] Clark RE. How much and what type of guidance is optimal for learning from instruction? In: Constructivist instruction : success or failure?. First. London, England: Taylor & Francis Group; 2009, p. 158–83.
[100] OpenGear, IM4208, IM4216, IM4232 & IM4248 Quick Start Guide, n.d..
[101] Hart SG. Nasa-task load index (NASA-TLX); 20 years later. In: Proceedings of the human factors and ergonomics society annual meeting. Vol. 50, (9):SAGE Publications Inc; 2006, p. 904–8. http://dx.doi.org/10.1177/154193120605000909.
[102] Sweller J, Ayres P, Kalyuga S. Measuring cognitive load. In: Sweller J, Ayres P, Kalyuga S, editors. Cognitive load theory. Explorations in the learning sciences, instructional systems and performance technologies, New York, NY: Springer; 2011, p. 71–85. http://dx.doi.org/10.1007/978-1-4419-8126-4_6.
[103] Gerjets P, Scheiter K, Catrambone R. Designing instructional examples to reduce intrinsic cognitive load: molar versus modular presentation of solution procedures. Instruct Sci 2004;32(1):33–58. http://dx.doi.org/10.1023/B:TRUC.0000021809.10236.71.
[104] Gerjets P, Scheiter K, Catrambone R. Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? Learn Instruct 2006;16(2):104–21. http://dx.doi.org/10.1016/j.learninstruc.2006.02.007.
[105] Sweller J. Measuring cognitive load. Perspect Med Educ 2018;7(1):1–2. http://dx.doi.org/10.1007/s40037-017-0395-4.
[106] Skinner BF. The behavior of organisms: an experimental analysis. Century psychology series, first ed.. New York: Appleton-Century Company; 1938.
[107] Kraiger K, Ford JK, Salas E. Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. J Appl Psychol 1993;78(2):311–28. http://dx.doi.org/10.1037/0021-9010.78.2.311.
[108] Witmer BG, Singer MJ. Measuring presence in virtual environments: a presence questionnaire. Presence: Teleoperators Virtual Environ 1998;7(3):225–40. http://dx.doi.org/10.1162/105474698565686.
[109] Ward M, Sweller J. Structuring effective worked examples. Cogn Instruct 1990;7(1):1–39. http://dx.doi.org/10.1207/s1532690xci0701_1.
[110] Csikszentmihalyi M. Flow: The psychology of optimal experience. 1990.
[111] Sweller J. Story of a research program. Educ Rev 2016;23. http://dx.doi.org/10.14507/er.v23.2025.
[112] Sweller J, Paas F. Should self-regulated learning be integrated with cognitive load theory? A commentary. Learn Instruct 2017;51:85–9. http://dx.doi.org/10.1016/j.learninstruc.2017.05.005.
[113] Greenwald AG, Banaji MR. The implicit revolution: reconceiving the relation between conscious and unconscious. Am Psychol 2017;72(9):861–71. http://dx.doi.org/10.1037/amp0000238.
[114] Brooke J. SUS: A "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability evaluation in industry. London: Taylor and Francis; 1996, p. 189–227.
[115] Chen O, Castro-Alonso JC, Paas F, Sweller J. Extending cognitive load theory to incorporate working memory resource depletion: evidence from the spacing effect. Educ Psychol Rev 2018;30(2):483–501. http://dx.doi.org/10.1007/s10648-017-9426-2.
[116] Moore AG, McMahan RP, Dong H, Ruozzi N. Extracting velocity-based user-tracking features to predict learning gains in a virtual reality training application. In: 2020 IEEE international symposium on mixed and augmented reality (ISMAR). Virtual Conference: IEEE Computer Society; 2020, p. 881–90. http://dx.doi.org/10.1109/ISMAR50242.2020.00099.