Instructional Design Guidelines for Procedural Instruction Delivered via Augmented Reality

Chris W. Wasko

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in

Curriculum and Instruction

Barbra B. Lockee, Chairperson

John K. Burton

David M. Moore

Kenneth R. Potter

Keywords: instructional message design, augmented reality, multimedia learning

May 8th, 2013

Blacksburg, VA

Instructional Design Guidelines for Procedural Instruction Delivered via Augmented Reality

Chris W. Wasko

Abstract

Augmented reality, defined as a real-time direct or indirect view of a physical real-world environment that has been enhanced by adding digital computer generated information to it, is rapidly developing in terms of associated hardware (wearable displays, wireless mobile devices) and software (development platforms). AR enhanced instruction has been shown to provide cognitive and psychomotor support during procedural learning and has been shown to use both words and pictures when delivering instructional content. A set of message design guidelines, created using a design and development research approach, can be used by novice designers to effectively manage the use of words and pictures while developing instructional applications for AR.

Dedication

To Erin, Nancy, and Ron

TABLE OF CONTENTS

List of Appendices

List of Tables

Chapter 1: Introduction and Need for the Study

Brief Introduction to AR

The term "augmented reality" (AR) was first employed in 1992 by Caudell and Mizell who used it to describe the process of overlaying digital content on a view of the real world. The digital content typically associated with AR includes images, text, animations, 3D models, and sound. Generally, the content is overlaid on a direct, real-time view of the world.

Initially, users experienced AR by wearing a complex optical see-through head mounted display hooked up to a computer (Sutherland, 1968). Today, sophisticated wireless mobile devices (WMDs) are capable of delivering AR content without causing the user to carry a cumbersome amount of expensive equipment. WMD hardware and software components are used to provide access to AR content. For example, AR content can be triggered when the device's camera identifies a marker or when the device's accelerometer and GPS system recognize a specific location. While it is more portable and less cumbersome than the initial head mounted displays, AR experienced via a WMD does compromise the use of a user's hand, as they are required to hold the device aloft in their field of vision. The pending release of mass-market wearable displays holds the potential to dramatically shift the manner in which (frequency, full use of hands) the general public will experience AR.

There are many examples of non-instructional applications of AR. For instance, AR is frequently used for advertising and gaming purposes in which digital content is used to solicit or entertain users. However, a review of related literature yields numerous reported examples of informal and formal instruction being delivered with the assistance of AR.

Informally, AR has been used to teach tourists about historic events that happened in a specific location. In the system developed by Vlahakis et al. (2002), models and 3D animations of historic buildings and events were overlaid on a relevant physical space. AR has been used to provide 3D instructions to help users assemble a piece of furniture and in the case of Google Sky Map, as a means to provide just-in-time, as needed information about stars and constellations that a user is viewing.

Formally, there are many notable uses of AR. The instructional delivery approach has been shown to be useful in helping learners complete procedural tasks. Henderson and Feiner (2009) describe an AR system that helps military workers learn to conduct routine maintenance tasks more safely and accurately. Yeo, Ungi, Lasso, McGraw, and Fichtinger (20ll) detailed an AR system designed to help physicians make challenging freehand incisions. Liao, Inomata, Sakuma, and Dohi (2010) used an AR overlay system to assist with complex surgical procedures. Also, AR has been used to create open-ended learning experiences (Dunleavy, Dede, & Mitchell, 2009; Klopfer, 2008; Squire & Jan, 2007; Squire & Jenkins, 2011) in which students must work in teams and navigate a meaningful real-world environment while attempting to solve an ill-defined problem. Interactive digital AR characters and objects, accessed via a WMD, are used to scaffold the experience by providing information that learners can use to solve the central problem of the practice field. Typically, the anchor problem of the AR enhanced learning experiences centers around something that would be logistically impossible, such as a disease outbreak or pollution, for students to investigate in real life.

AR, the concept of overlaying digital information on a direct or indirect view of the physical world, continues to evolve due to advances in related software and hardware. Such advances seem to suggest that in the future, AR will be used to deliver a variety of instructional messages to an increasingly larger amount of learners.

Brief Introduction to Design and Development Research

Richey and Klein (2007) define design and development research as

> the systematic study of design, development, and evaluation processes with the aim of establishing an empirical basis for the creation of instructional and non-instructional products and tools and new or enhanced models that govern their development. (p. 1)

Design and development research aims to answer a call that researchers in the field of instructional design and technology should be "studying design, development, and evaluation as well as doing it" (Richey & Klein, 2005, p. 6).

Researchers can elect to conduct a context-specific product or tool study in which they examine the design and development of an instructional product or program. Conversely, researchers can elect to conduct a more generalized model study with the goal of producing new (or an enhanced) design or development model. Design and development research will contribute to the Design and Development Knowledge Base according to Richey and Klein (2007) who claim that the outcomes will explain, "the role of the designer or the context in which design and development takes place, the designer problem-solving processes, the ultimate value of ID models, ways of reducing design cycle time, and the most effective and efficient ways of using learning objects when designing instruction using advanced technologies" (p. 24).

Need for Study

The New Media Consortium's Horizon Report is a yearly publication that identifies and describes emerging technologies likely to have a large impact on education in the near future. AR was addressed in both the 2011 and 2012 reports as an emerging technology that remains 4-5 years away from widespread adoption. According to the 2012 Horizon Report (*The Horizon Report*, n.d.)

> "AR is a capability that is shifting from what once required rooms of equipment to a set of simple-to-use tools with tremendous potential. The layering of information over 3D space produces a new experience of the world, sometimes referred to as "blended reality," bringing with it new expectations regarding access to information and new opportunities for learning. While the most prevalent uses of AR so far have been in the consumer sector, new uses seem to emerge almost daily, as tools for creating new applications become even easier to use. A key characteristic of AR is its ability to respond to user input. This interactivity confers significant potential for learning and assessment; with it, students can construct new understanding based on interactions with virtual objects that bring underlying data to life" (*The Horizon Report*, n.d., p. 28).

The emergence of AR is partially due to advances in software that can be used by novices to create and share AR enhanced content. One such software tool is the free to use AR browser Layar Vision. The browser uses detection, tracking, and computer vision to overlay digital information on the physical world. A second such tool is BuildAR, which can be used by novices to create and host mobile AR content online. Additionally, ARIS and FreshAiR are two open-source platforms for creating and experiencing mobile games, tours, and interactive stories. Via the platforms, novice users can create hybrid worlds of virtual interactive characters, items, and media placed in physical space. Also, easy to use application creation software such as Game Salad and M.I.T. App Inventor are available for use by those who wish to design and develop their own AR application. Essentially, AR development software designed for non-specialized users is readily available for those with a viable application in mind.

Trends related to the hardware used to experience AR are also contributing to its emergence. The Pew Research Center's 2012 Internet and American Life Project, Summer Tracking Survey indicates that over one half of cell phone owners use their device to access the internet. Such smartphones posses the location-aware components necessary for delivery of digital AR content. According to the Pew Survey, the ownership of such phones has dramatically increased from 25 percent in mid 2008 to 56 percent in 2012. Approximately 77 percent of 18-29 year olds own internet enabled smartphones, compared to 69 percent of 30-49 year olds and 40 percent of 50-64 year olds. Clearly, the ownership of internet enabled smartphones capable of delivering AR content is rapidly increasing, particularly within the 18-29 year old demographic.

More importantly, trends related to wearable computers have the potential to significantly alter the way in which and frequency with which users experience AR content. Recently, the Oakley corporation started selling a pair ski goggles that display information, perceived as a 14 inch display at five feet away, directly in the field of vision of users (Pollicino, 2012). The digital content informs users of performance statistics such as speed and distance traveled and users view the location of other skiers. The pending release of mass-market wearable computing devices such as Google's Project Glass or the Lumus Personal Display potentially means that in the mid-term future, a large amount of users will be able to access hands-free digital AR content in almost any location. The global market for such wearable devices is estimated to jump to $1.5 billion annually by 2014 (Shalvey, 2012).

Clearly AR is a rapidly emerging technology due to advances in related hardware and software. Recent advances suggest that a designer who has an idea for an instructional application of AR has access to the tools required to develop the application and will be able to share the instruction with learners who already own, or will be willing to purchase the equipment necessary to access the content. The instructional applications of AR seem limitless: teaching a user how to maintain the engine of his car, write calligraphy, tie a tie, execute a motor skill such as putting a golf ball, assemble a piece of furniture, or learn about the architectural elements of surrounding buildings. However, entrepreneurs, lacking a background in instructional design, who attempt to develop an AR application, will be unable to find any assistive literature and/or guidelines to support their efforts.

A review of literature indicates that augmented reality has not been directly addressed from an instructional design perspective. Sound instructional design procedures and techniques, while known to experts in the field, will likely be unknown by novices. As such, the novice designers may make decisions during the ID process that will negatively affect learning outcomes.

Purpose of Study

The purpose of this design and development research was to create a set of theoretically grounded and empirically supported guidelines that can help users effectively design instruction to be partially or entirely delivered via AR. The guidelines aim to help design instruction with procedural learning goals, given AR's documented ability to facilitate such learning. The guidelines were built on theoretical components, selected as a result of a comprehensive literature review, from the multimedia learning research base. In order to be of maximum use, the guidelines aim to help designers understand the human learning process and to help them make reasonable decisions about how to use pictures and words in their instruction, how to arrange content with regards to space and time, and how to avoid cluttering the instruction with unnecessary information that can interfere with learning.

In addition to addressing the addressing the aforementioned need, this research contributes to the Instructional Design Knowledge Base (Richey & Klein, 2007) by producing generalizable knowledge that aims to improve the process of instructional design. The completed set of guidelines is presented in document form. For each guideline, users are provided with a description, an explanation of the underlying connections to the cognitive information processing system, and a contextualized example. The initial version of the guidelines was shared with experts who evaluated them based on a researcher created rubric. Expert suggestions were used to revise the guidelines. Ultimately, this research will informs the Knowledge Base about how well the core message design principles for multimedia instruction can be applied to AR instructional design and about the usefulness of the expert review technique during the design and development process.

Chapter 2: Review of Literature

Introduction

The following review of literature is presented to provide information related to the design, development, and validation of the proposed guidelines. The review addresses:

1. Overview of augmented reality including history, associated attributes, entry points, learning outcomes, and sample uses,
2. Diagnostic review of reported instructional applications of AR,
3. Overview of multimedia learning including related theories, limitations, principles, and evidence,
4. Explanation of design and development research including history, goals, and notable examples.

Augmented Reality

Introduction

As explained by Azuma (1997), augmented reality is a variation of virtual reality. Virtual reality technologies completely immerse a user inside a synthetic environment in which they can not see the outside world. In contrast, AR allows the user to see the real world, with virtual objects superimposed upon or composited with the real world. The term *augmented reality* (AR) is attributed to Boeing employees Caudell and Mizell who used it in 1992 to describe how they helped workers assemble wires and cables for an aircraft. AR is sometimes inaccurately referred to as mixed reality. AR is indeed a point on the reality-virtuality continuum but the two terms are not synonymous as AR but one type of mixed reality (Milgram & Kishino, 1994).

Azuma (1997) defined AR as the real time combination of real and virtual information registered in 3D. Carmigniani and Furht (2011) stated "augmented reality is a real-time direct or indirect view of a physical real-world environment that has been enhanced/augmented by adding virtual computer generated information to it" (p.1). The main distinction between the two definitions is the emphasis that Azuma places on the concept of spatial registration. Azuma believes that if the digital and physical objects were not aligned, the utility of the AR application would be compromised.

The term AR still endures today though it is not without its critics. Hughes, Fuchs and Nannipieri (2011) posited the philosophical question, "What is augmented in augmented reality?" Hughes et al. claimed "if reality is the definition of everything that exists, then reality cannot be augmented since it is already everything" (p. 47). Hughes et al. (2011) claim that the user's perception of reality is augmented, implying that AR is used not to augment reality but to either augment the perception of reality or to create an artificial environment.

Taxonomy

Hughes et al. (2011) recently developed a taxonomy of AR in which they sort based on functionalities such as *augmented perception of reality* or *proposing an artificial environment* that represents some past or future reality or even an impossible reality. In the first functionality, augmenting a perception of reality, AR is a tool for assisting decision-making and it can provide information that will optimize our action on reality. Hughes et al. (2011) stated that, "AR enables objects, beings, or relations which exist in reality but which cannot be perceived by users to be visualized" (p. 51). This functionality is "a prisoner of the present" (Hughes et al., 2011, p. 60), and thus all types of augmentation that fall under this category deal with augmenting a perception of reality as it exists presently. One of the most basic types of augmentation in this category deals with adding semantic information like titles, keys, or symbols more or less visually close to real objects. For example, Baird (2000) examined how effectively someone using a wearable computer could be aided in performing an assembly task. The participants used a wearable AR delivery device (head-mounted display or HMD) and were presented with instructions on how to assemble a computer motherboard. In order to help the user select a specific part, the AR instructions placed digital arrows in the user's field of view. A second arrow directed users to install the part in a specific location.

The second functionality of AR, immersing users in a artificial environment, helps users imagine reality as it was in the past or as it could be in the future. Hughes et al. (2011) stated, "if the first type of AR is a prisoner of reality, this type goes far beyond it" (p. 60). One example of this type of augmentation would be the use of a smartphone to overlay pieces of virtual furniture in an empty room to help imagine what a particular arrangement would look like. This functionality of AR can work backwards in time as well. Here, AR associates virtual objects that no longer exist with a real environment. An example of this functionality would be the Berlin Wall 3D (Gardeya, 2010) mobile application that allows users to walk around Berlin and view a digital model of the Wall in the exact location where it used to stand.

Associated Attributes

In an article about selecting instructional media, Levie and Dickie (1971) explained that

> the attributes of a medium are the capabilities of that medium to show objects in motion, objects in color, objects in three dimensions, to provide printed words, spoken words, simultaneous visual and auditory stimuli; to allow for overt learner responses or random access to information. (p.860)

To be used as an instructional medium, AR requires a computer and output device, traditionally a laptop and a HMD but more frequently, given increases in computing power and decreases in size (Kaufmann & Csisinko, 2011), a smartphone that serves as both computer and output device. The primary technical components of an AR delivery system enable the two main attributes commonly associated with AR: (a) the ability to overlay a variety of digital information (text, images, video, 3D models and even sounds) on the real world and (b) the ability to perform the overlaying in almost any location. As will be discussed later in this chapter, the fact that AR can overlay detailed digital information on relevant locations allows for AR to provide useful assistance related to both the cognitive and psychomotor aspects of procedural learning.

Entry Points

Digital AR content can be triggered by the location of the user or by a marker that has been scanned by the camera on the user's device. Location related AR is typically accessed on a mobile phone and takes advantage of the device's camera, GPS, compass, accelerometer, and Internet connection. As a user moves around a physical space, the phone will use GPS to determine the location while the compass and accelerometer will determine the field of view. These technical features, used in conjunction with each other, allow the device to make an approximation of the user's location and appropriate content will be displayed based on that location. Location aware AR is considered less reliable because it is dependent on the device being able to pinpoint an exact location.

Vision related AR is displayed when a marker is scanned by the camera on the user's device. QR codes are frequently used as a marker but their usefulness is limited by the infeasibility of posting QR codes in real world locations. However, QR codes are only one type of vision based entry point. Software such as the AR ToolKit can be used to produce markers of varying sizes and shapes.

Some vision based AR does not require a code or a marker as an entry point. For example, De Crescenzio et al. (2011) developed a vision based AR delivery system that projected maintenance instructions directly onto the cockpit and wings of an airplane. The AR system targeted features of the plane that "consist of visual patterns (such as patches, circular blobs and arbitrarily shaped regions) that can be detected and matched in natural images and exhibit invariance or robustness" (De Crescenzio et al., 2011, p. 98). Offline, the researchers took pictures of the cockpit and wing in order to acquire reference images for the objects to be subjected to augmentation. Software extracted the local invariant features from the images and stored them in a database. A user could then view the cockpit through an online HMD and the system would match up the incoming feed with the stored images and display specific augmentation in an appropriate location without the need for a physical marker. This vision based system proved to be quite effective for delivering maintenance instructions and managed to avoid having to use markers. However, the system was limited in that it is not widely accessible and users had to have access to an expensive HMD in order to experience the augmentation.

Recently, the Autonomy Corporation released a product that enables a marker-less, widely accessible vision based AR. This augmentation is experienced through the Aurasma browser on a mobile phone and is powered by a scaled down version of Autonomy's IDOL pattern recognizer. As explained by documentation from the Autonomy company, the IDOL system allows for “automatic categorization, tagging, linking, retrieval, and profiling of all forms of unstructured information in real time (Autonomy Corp, 2006, p.1). Essentially, this software can be used to augment virtually any physical object ranging from a landmark to a picture. Users take a picture of an object and the IDOL system scans, tags, and links it for future retrieval. Once the target of augmentation is selected, users determine what content (possibilities include an image, video, or 3D animation) will be displayed over the object. Such software allows virtually any physical object to be an entry point for augmentation.

Location related AR is quite common but is only as accurate as the features of the device (GPS, compass, accelerometer, wireless signal) will allow for. Comparatively, vision related AR can involve markers or can be marker-less and is more accurate. Markers such as QR codes or those designed by software such as the ARToolKit can serve as entry points but it is not always possible to place such markers in a physical setting. Marker-less vision related AR allows for virtually any physical object to be an entry point and is becoming increasingly available for mass usage and requires increasingly less cumbersome equipment.

Review of Instructional Applications of Augmented Reality

A review of reported instructional applications of AR, discussed in academic publications, was necessary to ascertain a clearer picture of how exactly the technology was being used instructionally by researchers. Given that AR is a relatively new technology, it was not surprising that the reported instructional uses had gone largely unanalyzed. Table 1 contains summaries of twenty articles that address instructional uses of AR. Each entry is analyzed in terms of the type of knowledge taught, learner traits, the learning environment, and the modalities (words or pictures) in which instruction was delivered.

In order to clarify the nature of the articles discussed in Table 1 (Appendix A), two representative examples will be summarized here. A recent study (De Crescenzio et al., 2011) examined the ability of AR to deliver instructions to workers in an authentic setting. The instructions attempted to reduce procedure violations and misinterpretation of facts on the part of the workers. Using the system via a HMD, workers performed tasks on a real plane while being guided by overlaid augmentations that included 2D arrows that directed their attention to specific parts of the plane and, when completing an oil change, a 3D animation that would show a model of the oil cap being twisted in the proper direction and removed. Users provided qualitative data that suggested the AR system was able to reduce the mental and physical workload associated with the maintenance task.

Researchers from Germany (Reif, Günthner, Schwerdtfeger, & Klinker, 2010) developed an AR system that instructed factory workers to select specific items from warehouse storage shelves. Reif et al. (2010) attempted to enhance the productivity of workers by allowing them to complete item selection quicker and more accurately. The Pick-A-Vision system required workers to wear a HMD and small portable computer as they navigated a factory floor. The system displayed text information indicating the aisle for next stock location and as the worker approached the correct storage compartment, an identifying arrow overlaid on their field of vision changed colors to indicate that the correct storage compartment had been identified (Reif et al., 2010). By using the ability to overlay text and images as a means to guide workers to a specific location, The Pick-A-Vision system was able to lower the selection error rate as compared to workers using a traditional paper list (Reif et al., 2010).

Table 1: *Summary of publications in which AR is used as an instructional delivery method*

Appendix A.

Analysis of Reported AR Instructional Uses

An analysis of the twenty articles described in Table 1 identifies several relevant trends. First, in recent related publications, AR has been used equally with novice learners and with expert learners. AR has been used almost equally in controlled lab settings and in real-world settings. These two trends suggest that AR can be used to deliver instruction to a variety of learners in a variety of settings.

Two additional observations, both of which will be discussed in greater detail, are suggested by the analysis of Table 1. First, while in a few reported cases (Dunleavy et al., 2008; Squire, 2010; Squire & Jan, 2007; Squire & Klopfer, 2007) AR has been as a delivery tool to facilitate open-ended, constructivist style learning experiences that require students to assume roles and interact with digital characters and objects, the overwhelming instructional use of AR is to deliver content that helps users learn to complete procedural tasks. Second, despite the visual emphasis of AR, in nearly eighty percent of the articles summarized, both words and pictures are used to deliver content.

Observation 1: AR and procedural learning.

Overview.

AR has been used to teach a variety of procedural tasks. In several studies (De Cresenzio et al., 2011; Haritos & Macchiarella, 2004; Henderson & Feiner, 2009; Henderson & Feiner, 2011a, Henderson & Feiner, 2011b, Neumann & Majoros, 1998, Tang et al., 2001), AR was used to train participants to complete maintenance or assembly tasks in which a specific set of procedures was to be followed. In two studies (Liao et al., 2010; Yeo et al., 2011), AR instruction was used to train specialized medical professionals to perform complex surgical procedures. In other cases (Cakmacki et a1., 2003, & Gandy et al., 2005), AR was used to help teach arts based procedural tasks such as playing a bass guitar and acting out a movie scene.

Cognitive and psychomotor phases.

Building on earlier work that was done by Neumann and Majoros (1998), Henderson and Feiner (2011a) explain that activities related to procedural learning can be classified as cognitive or psychomotor and offer evidence that AR can assist with both aspects. Henderson and Feiner (2011a) explain that cognitive phase activities include directing attention, comprehending instructions, and transposing information from the instructions to the actual task environment. Psychomotor phase activities are said to include performing physical manipulations such as comparing, aligning, and adjusting. Essentially, the two components involve the learner determining an effective course of action and then acting upon it.

In two studies that examined the ability of AR to assist with procedural learning, Henderson and Feiner (2011a, 2011b) provided specific examples of AR design elements that were used to assist with both the cognitive and psychomotor phases. The design elements are summarized in Table 2.

Table 2: *Cognitive and psychomotor aspects of AR.*

Procedural Task Phase	**Design Elements**
Cognitive	* attention directing information in the form of 3D and 2D arrows * text instructions describing the task and accompanying notes and warnings * registered labels showing the location of a target object *3D models of related components (tools and other objects) registered at their current or projected locations in the physical environment
Psychomotor	* 3D arrows that suggest motion and provide feedback about a current orientation compared to a desired end state * color-coded highlight effects designed to help align and connect two parts * labels that move as objects are moved

In Table 2, we see that the cognitive phase elements are employed to help learners correctly orient themselves to the task that needs to be completed. Here, the digital content is meant to foster comprehension of the next step in the procedure. Comparatively, the psychomotor tasks aim to promote effective action once the learner understands what to do. The psychomotor elements help ensure that learners are efficient in their psychical movements and that objects in the physical world are manipulated correctly.

Supporting evidence.

Strong evidence of the ability of AR to help with both the cognitive and psychomotor aspects of procedural learning is best exemplified in five of the articles addressed in Table 1. Henderson and Feiner (2011) designed and tested a wearable AR system that attempted to help a mechanic perform 18 tasks such as installing and removing fasteners and connecting cables in an authentic field setting. The AR system was designed to assist with cognitive aspects of the procedural task. This assistance was provided in the form of attention directing arrows and text instructions describing the task. Compared to more traditional instructions delivered via an LCD monitor, the AR system significantly helped mechanics locate tasks more quickly and reduced head and neck movements during the repair process. The reduced task location times and physical movements suggest that using AR as an instructional delivery method can minimize the demands of the cognitive phase of a procedural task.

Henderson and Feiner (2011b) conducted a study that illustrated how AR can assist with the psychomotor aspect of procedural learning. In the study, a wearable AR display was used to overlay instructions related to the assembly of an airplane engine. This AR system provided cognitive phase support in the form of 2D text that communicated the desired action, and helped direct attention to salient objects and physical locations. Three specific techniques were used to provide support for the psychomotor phase including 3D arrows that provided feedback about the correct orientation of an object compared to the desired end state, highlighting effects that helped users connect and align two objects, and dynamic labels that responded to the motion of the user and the labeled objects. All of the psychomotor techniques attempted to help users correctly perform the actual tasks once they determined an appropriate course of action. Findings indicated that the AR system was able to help users perform the psychomotor aspects of the engine assembly more quickly than when using a more traditional (LCD monitor) delivery method. Also, a statistically significant amount of participants preferred the AR treatment and indicated that it was more intuitive than the LCD treatment. The AR system also allowed participants to complete alignment tasks more accurately that the LCD treatment. Essentially, the AR system was found to be more accurate and faster for the psychomotor phase activities and was overwhelmingly preferred by participants.

De Crescenzio et al. (2011) used of an AR system to help workers maintain airplanes. This system was designed to deliver instructions that would reduce errors due to procedure violations, misinterpretation of facts or insufficient training. While the study did not use the terms cognitive phase and psychomotor phase, the AR system clearly tried to assist with tasks related to both phases. Cognitive phase assistance came in the form of overlaid 2D arrows that directed worker attention to specific parts of the plane. Psychomotor assistance came in the form of a 3D animation that showed how an oil cap should be twisted during removal. Participants were asked to fill out a NASA-TLX (Task Load Index) form and results indicated that the AR system led to low ratings for mental workload and effort and high ratings for user satisfaction. Such ratings suggest that the cognitive and psychomotor techniques used in the AR delivery were successful.

Tang, Owen, Biocca and Mou (2001) attempted to empirically test the effectiveness of AR instructions in an assembly task. The researchers selected 75 participants and had them complete a 56-step assembly task. At each step, participants had to perform a cognitive phase activity in the form of picking a part of specific color and size from an unsorted part-bin and a psychomotor phase activity in the form of placing the part in the correct location and position. Researchers compared four treatments (Tang et al., 2001, p. 75) including a printed instructional manual, instructions on an LCD monitor next to the workspace, instructions delivered through a HMD but displayed 4 feet ahead of the workspace and a spatially registered AR display delivered through a HMD.

The findings (Tang et al., 2001, p. 77) indicated that the task was completed in the shortest mean time with the AR treatment and that the paper treatment was almost 4 minutes longer. The AR treatment was significantly better than paper manual as were the other two treatments. The AR treatment led to significant gains in accuracy compared to the other approaches in terms of total errors and dependent errors (errors that are the result of a prior mistake).

Participants had to fill out the NASA-TLX (Task Load Index) form, on which they indicated that the AR instructional delivery helped decrease frustration level, mental workload, and physical workload. Tang et al. (2001) argued that the findings provide evidence to support the claim that AR systems improve task performance and can relieve mental workload on assembly tasks. They stated "the ability to overlay and register information on the workspace in a spatially meaningful way allows AR to be a more effective instructional medium" (Tang et al., 2001, p. 79). Here again, AR was shown to help with both cognitive and psychomotor aspects of procedural learning.

Finally, Yeo et al. (2011) attempted to determine if AR could assist medical trainees in learning the correct placement of a needle for a very specific type of joint injection that had to be done in an exact spot at an exact depth, a task that involved both cognitive and psychomotor demands. Researchers selected 40 volunteers and randomized them into a treatment and control group. The control group participated in a training session that included six freehand insertions into a synthetic model spine and the treatment group received a training session that included guidance from an AR experience. To offer support for cognitive phase activities, the AR treatment overlaid a guiding image on the model spine and for psychomotor support, used a digital marker to indicate the proper insertion point of the needle. After training, each member made two insertions on a spine phantom and researchers recorded data related to total procedure time, the time inside the phantom, the path inside the phantom, and potential tissue damage, (Yeo et al., 2011).

The results were strongly in support of the AR training method suggesting that the cognitive and psychomotor support was effective. While the time of the total procedure and the time inside the phantom spine were not significantly different between the two groups, the treatment group had a larger number of successful placements and took a significantly more efficient path inside the phantom and also had a lower amount of potential tissue damage. Yeo et al. (2011) claimed that “the overlay system helped avoid the initial period of high errors and lengthy procedures and also improved the overall accuracy and efficiency after the training session and would likely decrease the amount of practice required for medical students to become eligible for clinical procedures, or to master this specific needle insertion technique" (p. 2035).

Observation 2: words and pictures in AR.

Given the visual emphasis of AR, those unfamiliar with sound instructional design procedures may assume that AR instruction only involves visuals. However, in accordance with cognitive benefits of delivering information both presentation modes, it is not surprising that both words and pictures are used in nearly eighty percent of the articles discussed in Table 1. In the few studies where content was presented using only visuals, verbal content was likely omitted because a. learners were highly skilled experts (Liao et al., 2010; State et al., 1996; Yeo et al., 2011) with background knowledge that would likely have made the additional information redundant or b. words were absolutely unessential (Cakmacki et al., 2003; Robertson et al., 2008; Tang et al., 2003) for that specific type of procedural learning.

Implications

For the design of the guidelines that were developed during this study, the analysis of Table 1 had two major implications. First, in the related literature, AR has primarily been used to assist with the delivery of content that has procedural learning goals. The ability of AR to present appropriate digital content co-located on a view of the real world has been shown to mitigate the demands of cognitive and psychomotor aspects of procedural learning. Also, compared to the current constraints imposed by handheld smart phones and tablets, the release of mass-market wearable AR displays will allow for unfettered use of both hands. Given these pieces of supporting evidence, it seemed practical to focus the guidelines on procedural learning, given how helpful AR has proven with such learning goals and how well the new equipment will support such learning.

Second, the fact that both words and pictures are used to deliver content in the most of AR instructional applications discussed in Table 1 provided guidance as to where to seek theoretical support regarding message design fundamentals. The classic definition of multimedia instruction written by Mayer (2005a) states that, "multimedia instruction involves presenting words and pictures that are intended to promote learning" (p.2). Based on this definition, AR enhanced instruction can be reasonably associated with other well-known forms of multimedia instruction (a professor lecturing while referring to images on a projector, an educational video, a self-paced computer module using both pictures and words) because in most cases, such instruction involves both pictures and words. Given this association, the body of literature that addresses multimedia learning, informed by scholars such as Clark (2005), Fletcher (2005), Low (2005), Mayer (2005a, 2005b, 2005c, 2005d), Tobias (2005), and Sweller (1991), will be the focus of the next section of the literature review.

Cognitive Theory of Multimedia Learning

Introduction

According to Mayer (2005b), "the fundamental hypothesis underlying research on multimedia learning is that multimedia instructional messages that are designed in light of how the human mind works are more likely to lead to meaningful learning than those that are not" (p.31). To that end, Mayer has developed a cognitive theory of multimedia learning (CTML) that is based on the three underlying assumptions: 1.humans have separate channels for processing visual/pictorial information and verbal/auditory information 2.the channels are limited in capacity 3.learning is an active process that requires coordinated cognitive processes. According to Mayer (2005b), CTML posits that in order for learning to occur, humans must select relevant words from text or narration, select relevant images, organize both the words and images into coherent verbal and pictorial representations, and integrate the representations with new knowledge. Mayer (2005b) claims that CTML is consistent with cognitive science principles of learning, yields predictions that can be tested, is consistent with empirical research, and is relevant to educational needs for improving the design of multimedia instructional messages.

Underlying Assumptions

In a recent interview for the AECT History Makers Project, Mayer (2012) explained that CTML emerged out of a desire to contribute to the science of learning by developing a theory of learning based on authenticity. Mayer stated that CTML, "essentially borrows ideas from other people" and it reflects the belief that, "in order to learn, humans must effectively select relevant information, mentally organize that information and then integrate that information with long term memory and other incoming information"(Mayer, 2012).

Dual channel assumption.

The dual-channel assumption, largely influenced by the work of Paivio (2007) and Baddeley (1999), implies that humans have separate information processing channels for visual material and for auditory material. Paivio (2007) explains that dual coding theory (DCT) is built on the idea that cognition involves the cooperative activity of two functionally independent but interconnected systems, a nonverbal system specialized for dealing with linguistic objects and events and a verbal system specialized for dealing directly with language.

Paivio (2007) explained an initial experiment that simulated further research related to DCT. In this experiment, participants were shown a rapid sequence of pictures and a rapid sequence of words. When later asked to recall the words and pictures in any order, the participants were better at recalling images. However, the participants more often recalled the sequential order of the words rather than the pictures. The results led Paivio to believe that visual and verbal information was processed differently, which led to an extensive research program and the eventual development of DCT.

Logogens, one of the representational units of DCT, are used in all language related phenomena such as recognition, memory, and production (Paivio, 2007). Imagens, the other representational unit, are used in perceptual recognition, memory, drawing, and other cognitive processing of nonverbal objects. Paivio claims that the representational units are dormant until they are connected to the outside world and to each other. An example of imagen activation would involve a subject being shown a picture of Columbus' sailing vessel and being able to recognize it as a ship or more specifically, the Mayflower. According to Paivio, the units are said to be transformable as with asking a person to visualize and N turned on its side and the subject being able to identify it as a Z. Paivio claims that there are processing constraints related to the representations as evidenced by the fact that humans struggle to recite the alphabet in reverse, whereas they are more able to recall a movie scene in reverse.

Similarly, the work of Baddeley (1999) also suggests a model of the human mind that assumes short, working, and long-term memories and that working memory processes words and images in different ways. Baddeley claims that working memory has a phonological loop system that evolution has developed to aid language acquisition. Baddeley feels that the loop is used for a subvocal rehearsal process that provides access to a fading memory trace before it decays and becomes inaccessible. Baddeley also assumes that working memory had a visual sketchpad that plays a role in processing visual information and is used for maintaining and manipulating visual images. Ultimately, the two mechanisms of working memory are said to be controlled by a central executive that uses the two systems to transfer relevant information to long-term memory.

As previously indicated, the work of both Paivio (2007) and Baddeley (1999) has contributed to the CTML's assumption that people have separate information channels for visually represented material and auditorily represented material. Mayer (2005b) claims that the channels can work together to create cross-channel representations. For example, on-screen text may initially be processed in the visual channel while the auditory channel may initially process a narration describing the same event but that in either case, the learner may form a corresponding mental image in the other channel.

Limited capacity assumption.

With regards to the limited capacity assumption, CTML borrows mainly from the research of Baddeley (1999), Chandler (1991), and Sweller (1991. 2005a, 2005b). Baddeley (1999) suggested that both the phonological loop and the visual sketchpad are limited in capacity. When discussing the limits of working memory, Sweller (2005a) suggests that working memory can hold about seven elements of information at one time, that it can combine, contrast, or manipulate about two to four elements at one time, and that without rehearsal, almost all the contents of working memory are lost within about 20 seconds. Sweller explains that attending to and organizing new information places either an intrinsic or extraneous load on working memory. Sweller writes that, “intrinsic cognitive load is caused by the natural complexity of the information that must be processed and extraneous cognitive load is caused by inappropriate instructional designs that ignore working memory limits” (p.26-27).

Sweller (2005a) suggests that from an evolutionary perspective, it is helpful that working memory is quite limited. Sweller cites an example of a person encountering a completely novel problematic situation without the benefit of any help. Given that working memory can only combine about four elements at a given time, the person would have to choose from approximately 24 potential solutions. However, if working memory could handle 10 elements, there would be an unworkable amount (3.5 million) of possible permutations. Chandler and Sweller (1991) conducted a series of experiments that aimed to establish the limits of working memory. In the experiments, subjects were provided with instructions in which related but superfluous information was provided. The extra information was found to negatively impact learning because subjects had to devote extra cognitive resources to the unnecessary information.

Active processing assumption.

The third main assumption of CTML is that humans actively engage in cognitive processing. Mayer (2005b) explains that rather than viewing humans as passive processors who seek to add as much information as possible to memory, CTML views humans as processors who actively pay attention, organize information, and integrate it with other knowledge. Mayer says the outcome of active learning is the construction of a coherent mental representation in the form of a knowledge structure such as a process, comparison, or generalization.

How Learning Occurs According to CTML.

CTML (Mayer, 2005b) posits that in order for learning to occur, the learner must select the relevant words for processing by verbal memory (using the limited auditory/verbal channel), must select relevant images for processing by visual working memory (using the limited visual/pictorial channel), must organize the words into a verbal model, must organize the selected images into a pictorial model, and must integrate the representations with each other and with prior knowledge. All five of the cognitive processes involve activity on the part of the learner. Mayer (2005b) states that, "the selection of words and images is not arbitrary. The learner must determine which words and images are relevant for making sense out of presented information" (p.39). Mayer claims that learners are limited in the amount of connections they can make between new information so they must focus on building a simple structure that makes sense to them. Mayer believes that integrating new information with existing knowledge is a demanding process that requires the efficient use of cognitive capacity.

Summary

CTML posits that humans process information using separate channels that allow for pictures and printed text to be held in a visual working memory channel for a very brief time and for spoken words and other sounds to be held in a verbal working memory channel for a very brief period of time. CTML assumes that the information must be actively organized and integrated with existing prior knowledge in order for learning to occur. For instructional design purposes, Mayer (2005b) suggested CTML is useful because the influence of a theory that explains how the human mind works will lead to more meaningful learning than instruction designed without the guidance of such a theory.

Limitations of Multimedia Learning

Multimedia learning on a computer (similar to AR enhanced learning using a WMD) has generated large amounts of excitement and attention in recent years. Clark and Feldon (2005), who define instructional media as a vehicle for presenting and delivering instruction (such as a book, newspaper, or a person) and define multimedia as the capacity of computers to provide real-time representations of nearly all existing media, examine five common misconceptions that educators have about computer based multimedia learning.

The first misconception involves the belief that multimedia instruction produces more learning than live instruction or older media. Clark and Feldon (2005) acknowledge that multimedia instruction on a computer has the ability to provide a broad array of methods such as recorded video or examples in the form of simulations. However, the authors explain that there is no evidence to suggest that, compared to other media, more learning happens using computer based instruction. They suggest that all methods, sensory modes, and components of instruction can be provided in a variety of media with equal learning outcomes.

The second misconception is that multimedia instruction is more motivating than traditional instructional media or live instructors. Clark and Feldon (2005) call this claim, "elusive at best" (p.101). The authors believe that multimedia courses may be more attractive to students, causing them to choose to take them. They suggest that students appear to choose multimedia because of expected flexibility and ease of learning but that such expectations may cause them to reduce effort and ultimately learn less. The authors cite an example of a student having the ability to stop a multimedia lesson when they are tired or bored and then allowing a large gap of time to pass before restarting, in which case learning may be compromised.

The third misconception addressed by Clark and Feldon (2005) is the belief that computer based multimedia instruction can be customized to accommodate different learning styles. They claim that there is no evidence to support the benefits of such customization, citing the unreliability of allowing learners to self-report their learning styles and the ineffectiveness of using diagnostic tests to classify learners as a particular style. The authors suggest that instead of learning styles, such instruction should be customized according to learner prior knowledge.

Clark and Feldon (2005) next address the claim that multimedia instruction can provide active pedagogical agents that increase learning and aid motivation. They cite a mixed review of findings and conclude that in cases where an included agent was shown to enhance learning, the agent employed an instructional method that was not otherwise provided to students. As such, the authors ask instructional designers to consider other, more cost effective ways to employ the methods used by the pedagogical agent.

The final misconception is related to the alleged ability of multimedia instruction to provide learner control and discovery pedagogy to enhance learning. Clark and Feldon (2005) address this misconception by citing a complete lack of empirical validation to support the benefits, in terms of efficiency and impact on successful transfer of skills, of pure discovery education over well-structured, guided instruction. They claim that when less support is available, learners risk wasting valuable cognitive resources on irrelevant information and as such, negatively impact learning outcomes. The authors cite the phenomenon of "thrashing" in which learners, who lack effective strategies to deal with an impasse in learning, will try increasingly maladaptive strategies until they eventually give up and pursue an unrelated learning goal.

Clark and Feldon (2005) address five benefits commonly associated with computer-based multimedia instruction. They suggest that multimedia instruction does not allow for increased learning over other media, that the associated motivational benefits are questionable at best, that there is no evidence to support its effectiveness in supporting a variety of learning styles, that multimedia pedagogical agents often employ an instructional method not otherwise present, and that using multimedia instruction for pure discovery learning is inefficient and ineffective. Ultimately, the authors suggest that the true benefits of computer-based multimedia instruction are found in the cost of instruction, including time savings for students and teachers when instruction is used to reach large numbers of students and increased access to quality instruction by disadvantaged rural groups of students.

Evidence Based Principles Relating to the Design of Multimedia Instruction

The following six principles, all of which are grounded in the cognitive theory of multimedia learning and are well supported by empirical evidence, have been developed through a comprehensive research program by scholars including Fletcher (2005), Low (2005), Moreno (2000, 2004), Tobias (2005), Sweller (1991, 2005a, 2005b), and most frequently, Mayer (2005a, 2005b, 2005c, 2005d). The six principles address designing multimedia instruction in an efficient and effective manner, thereby minimizing the amount of cognitive effort required of learners and in turn, maximizing the learners available cognitive resources.

Multimedia Principle

According to Fletcher and Tobias (2005), “the multimedia principle states that people learn better from words and pictures than from words alone” (p.117) and this position is echoed by Mayer (2008) who stated that, “people learn more deeply when they build connections between a verbal representation and a pictorial representation” (p. 766). Fletcher and Tobias (2005) suggest that the principle is supported by empirically derived theory suggesting that words and images evoke different conceptual processes. The multimedia principle (Mayer, 2005b) assumes that humans possess separate channels for processing visual and auditory information, that the channels are limited in the amount of information that can processed at one time, and that learning is an active process in which humans attend to relevant information, organize that information into coherent representations and integrate those representations with other knowledge.

The multimedia principle was identified in a review, commissioned by the Association of Psychological Science, of twenty-five principles of learning. Halpern, Graesser, and Hakel (2007), listed the principle third, stating, “information is encoded and remembered better when it is delivered in multiple modes than when delivered in only a single mode” (p.2). Clark and Mayer (2011) explain that multimedia presentations containing static or active graphics and words (printed, spoken) give learners the opportunity to make connections between the pictorial and verbal representations whereas instruction containing only words alone offers fewer opportunities to connect the words to other knowledge. Most frequently, the multimedia principle is discussed in terms of improving instruction by adding pictures to instruction containing only words but in some cases (Mayer & Anderson, 1991), the principle has been discussed in terms of adding words to instruction containing only pictures.

Clark and Mayer (2011) stress that graphics must be carefully and intentionally added to text. Clark and Mayer suggest avoiding graphics that decorate a page or represent only a single object. Instead, they advocate for adding graphics that help learners understand material by depicting changes in an object over time and graphics that depict relations among elements of an object. They indicate that in most cases, static animations are no less effective at depicting material compared to active, dynamic animations. For example, Narayanan and Hegarty (2002) attempted to teach students a physical process using both static and dynamic visuals. While students spent more time studying the instructions with animated and interactive graphics, the additional time-on-task did not result in enhanced comprehension. Also, students exposed to the dynamic graphics did not rate the materials as more interesting than students exposed to the static visuals.

Fletcher and Tobias (2005) acknowledge that the multimedia principle is most effective for learners with low levels of prior knowledge and spatial ability but add that additional research is needed to clarify the ways in which these learner characteristics interact with the principle. Clark and Mayer (2011) remind designers to supplement text-based instruction with coordinated graphics for low-knowledge learners and that advanced learners may be able to learn well mainly from text or even mainly from graphics.

Contiguity Principle

According to Mayer (2005c), the contiguity principle attempts to prevent a phenomenon called extraneous overload. Extraneous overload happens when the cognitive processing required to understand material in a multimedia message is overcome by the processing required to process extraneous material or a confusing layout in the multimedia message. Fletcher and Tobias (2005) suggest that verbal and pictorial information should be coordinated in both space (spatial contiguity) and time (temporal contiguity).

Mayer (2005c) explains that spatial contiguity simply attempts to lessen the cognitive effort required on the part of the learner by placing an illustration next to the paragraph that it describes. Integrating the text with the visual, as opposed to placing the text at the bottom of a page or screen, prevents the learner from having to expend cognitive effort scanning back and forth from the text to the visual. Clark and Mayer (2011) suggest that printed words should be placed next to the corresponding part of a graphic to which they refer and that a pointing line should be used to connect the name to the part rather than at the bottom of the area as a caption or legend.

Mayer (2008) states, “the temporal contiguity principle is that people learn better when corresponding narration and animation are presented simultaneously rather than successively” (p.765). Clark and Mayer (2011) stress that this is of particular importance when words are describing actions that are depicted in on-screen graphics. They state that, “when instruction separates corresponding words and graphics, learners experience a heavier load on working memory, leaving less capacity for deep learning” (p.103). They claim that having the pictures and words coordinated allows learners to hold them together in their working memories and therefore make meaningful connections between them.

If words are unnecessarily separated from pictures, either in space or in time, learners may fall victim to what Ayres and Sweller (2005) call split attention. Split attention forces the learner to use limited working memory capacity to coordinate the multiple sources of information. Clark and Mayer (2011) suggest that the contiguity principle may most strongly apply when dealing with low-knowledge learners and when graphics and words are complex.

Modality Principle

Low and Sweller (2005) explain the modality principle by stating that, “under certain, well-defined conditions, presenting some information in visual mode and other information in auditory mode can expand effective working memory capacity and so reduce the effects of an excessive cognitive load” (p.147). Clark and Mayer (2011) recommend that words should be presented in spoken form whenever a graphic is the focus of the words and both are presented simultaneously. Whereas the contiguity principle aimed to reduce extraneous processing, according to Mayer (2008), the modality principle aims to help manage essential processing required by learners.

Like the multimedia principle, the modality principle assumes that learners have both a visual/pictorial channel and an auditory channel. The modality principle assumes that presenting words in printed form requires the eyes to attend to them and therefore prevents them from fully attending to the graphic. Asking the visual channel to process both the graphic and the printed words threatens to overload the channel. The modality principle attempts to avoid this overload by presenting the words as spoken text and therefore having the auditory channel share in the required processing.

Low and Sweller (2005) explain that the logical relation between the two sources of information is critical and explain that the modality principle is not obtainable unless the two sources of information are unintelligible in isolation and therefore must be mentally integrated in order to be understood. Low and Sweller stress that care must be taken to ensure that auditory material is essential and not redundant and that the instructional material is sufficiently complex to warrant the use of a cognitive load reducing technique. Clark and Mayer (2011) caution that if words are technical, unfamiliar or not in the learners native language, they should remain available to the learner in visual form.

Redundancy Principle

According to Sweller (2005b), the redundancy principle suggests that, “redundant material, the same information presented in multiple forms or unnecessarily elaborated, interferes with rather than facilitates learning” (p.159). Sweller suggests that the redundant information causes learners to attend to both forms and to expend working memory resources coordinating the two sources.

Clark and Mayer (2011) caution designers to avoid presenting information in text form when the same information is presented in audio form. They suggest that many designers may be compelled to add the printed text and spoken text in order to cater to students that have visual learning styles and students that have auditory learning styles. This decision is said to be problematic because there is no evidence to support the fact that visualizers learn better with visual forms of instruction and verbalizers learn better with verbal modes of instruction.

Clark and Mayer (2011) note some boundary conditions to the redundancy principle. In situations where there are no graphics on a screen or page, concurrent printed and spoken text proved to be helpful in a problem solving transfer test. This effect was attributed to the fact that the visual channel of the learner was not overwhelmed. They recently revised the redundancy principle based on evidence that indicated the benefit of unobtrusively placing short amounts of printed text next to the corresponding parts of a graphic. Sweller (2005b) reminds that the redundancy principle is counterintuitive. Sweller explains that while designers may assume that presenting the same information in multiple forms can be beneficial for learners, such a decision actually requires learners to unnecessarily coordinate and relate multiple forms of the same information, which causes an extraneous cognitive load that interferes with learning.

Coherence Principle

The coherence principle (Mayer 2005c) states that, "people learn more deeply from a multimedia message when extraneous material is excluded rather than included" (p.183). Clark and Mayer (2011) call it, "the single most important principle" (p.151) and caution against adding any material that does not support the instructional goal. They explain that designers may add extraneous material in an attempt to "spice up" instruction in order to motivate learners. Such extraneous material includes elements such as dramatic stories, pictures or background music.

As early as 1938, Dewey warned about the pitfalls of extraneous information. Dewey (1938) cautioned against adding unnecessary "interesting adjuncts" to an otherwise boring lesson, explaining that such additions will not promote deep learning. As with previous principles, the modern rationale against adding such extra material centers around the assumption that humans have a limited working memory capacity and the additional material can overload and disrupt the limited resources in the auditory or visual channels.

According to Clark and Mayer (2011), adding extraneous visuals can be particularly detrimental to learning because the visuals can guide learner attention away from relevant material, can prevent the learner from building appropriate links among relevant material and may even prime inappropriate existing knowledge, which is then used to organize the incoming content. They explain that simpler visuals, those that present fewer details at one time, may be more helpful than more detailed visuals. They advise designers to carefully examine the words they use in instruction and to eliminate unnecessary words added to increase interest, to unnecessarily expand on key ideas, and to providing unnecessary technical depth. Ultimately, Mayer (2005c) states that, “extraneous material should be ruthlessly weeded out whenever possible so that the core of the essential material is salient to the learner” (p.196).

Personalization Principle

Mayer (2005d) offers that, “people learn more deeply when the words in a multimedia presentation are in a conversational style rather than a formal style” (p.201). Mayer (2008) suggests that a conversational style or polite wording creates a sense of social partnership with the narrator and this causes the learner to try harder to make sense of what their conversational partner is saying.

Clark and Mayer (2011) suggest that some designers may want to avoid putting words in a conversational style because such a style will detract from the seriousness of the message. They suggest that people work harder to understand material when they feel they are in a conversation with a partner rather than simply receiving information. When citing and example of a lesson on weather, Clark and Mayer (2011) suggest straightforward changes such as changing the sentence, “It produces the bright light that people notice as lightning” to “It produces the bright light you notice as a flash of lightning” (p.186). They do note that it is easy to add too much personal material, citing the example that adding material such as, “Wow, hi dude!” and “hang onto your hat, here we go!” (p.187) can easily distract the learner and can set an inappropriate tone for learning. Essentially, the personalization principle assumes what is said and how it is said are important factors to consider when designing a multimedia instructional message.

Summary

The previous six principles pertain to instruction that presents information containing both words and pictures. All six principles, summarized in Table 3, are grounded in the belief that humans have separate channels for processing auditory and pictorial information, that working memory has a limited capacity and that humans are actively involved in the learning process. All six principles directly address the “design” aspect “instructional design”. They are representative of the simple yet effective strategies that designers can use to facilitate learning. Some of the principles address removing or avoiding the addition of non-essential information while others involve taking what is essential and presenting it in a way that makes instruction most accessible for learners.

Table 3: *Summary description of multimedia principles.*

Principle	**Description**
Multimedia	People learn better from words and pictures than from words alone. Graphics must be carefully selected. Adding graphics to words or words to graphics gives learners more opportunities to connect new knowledge with prior knowledge. The multimedia principle is most effective for learners with low levels of prior knowledge.
Contiguity	Designers should attempt to minimize extraneous load caused by unnecessary processing. In order to prevent learners from having to scan from words to graphics, words should be placed next to (spatial contiguity) the part of a graphic to which they pertain. Words should also be presented in conjunction with (temporal contiguity) what is happening in the corresponding graphic. This principle applies most strongly to low knowledge learners.
Modality	In an attempt to manage essential processing and expand the capacity of working memory, whenever the graphic is the focus of words, the words should be presented in spoken form. The two forms of information must be unintelligible in isolation.
Redundancy	Designers should avoid adding repetitive information to their instruction. There is no need to include words in both printed and spoken form. This superfluous information will place an unnecessary cognitive burden upon learners by creating extraneous processing. This principle is said to be counterintuitive as designers may assume that more information is better for learners.
Coherence	Designers must avoid adding any information, such as dramatic pictures, background or background music, that is not directly related to the learning goal.
Personalization	Put words in a conversational style rather than a formal style and thereby create a sense of social partnership. Designers must remain professional when selecting words and avoid adding slang or jargon.

Research Based Evidence

A broad quantitative empirical research base supports all the six principles described in Table 3. For each principle, several illustrative studies have been summarized. All of the studies summarized as follows employed sound research methods (random assignment to control and experimental groups) and reported results using appropriate statistical measures (probability, effect size).

Evidence for the Multimedia Principle

Mayer (2008) detailed the findings of research related to the principle that people learn better from words and pictures than from words alone. Mayer stated that in 11 of 11 experiments involving paper-based lessons on brakes, pumps, generators, and lightning and computer based lessons of the same topics, learners who received corresponding graphics with words performed better on transfer tests than learners who received words alone. The median effect size was 1.39.

For example, Mayer and Gallini (1990) attempted to determine the value of text and visuals compared to text only and attempted to investigate the types of visuals that were most useful for helping learners understand scientific devices. Mayer and Gallini had groups of low-knowledge students read expository passages concerning how scientific devices work. The control group were not provided with any illustrations while other groups were given either static illustrations of the device with labels for each part, static illustrations with labels for each major action (steps), or dynamic illustrations showing the “off” and “on” states of the device along with labels for each part and each major action. Mayer and Gallini referred to the type of illustrations provided as, “explanative illustrations” (p.715) indicating they were aimed at helping students interpret text. The results indicated that the parts-and steps illustrations consistently improved performance on recall of conceptual information and creative problem solving. Students who were not given illustrations did poorly on measurements of recall of explanative information and developed fewer creative answers to transfer problems.

A second example is a series of three experiments by Mayer, Bove, Bryman, Mars, and Tapangco (1996). Researchers found that students who were given a summary containing a sequence of short captions with simple illustrations depicting the main steps in the process of lightning recalled the steps and solved transfer problems as well as or better than students who received only text. In subsequent experiments, Mayer et. al (1996) found that taking away the illustrations or the captions eliminated the effectiveness of the text summary and that adding additional words to the summary reduced its effectiveness. Mayer suggested that the annotated illustrations encouraged the learner to organize the material into a cause-and-effect system and to integrate the verbal and visual representations of the system. Mayer concluded that a verbal summary of a scientific process is not as effective as a multimedia summary that "is concise, in that only a few illustrations and sentences are provided; coherent, in that the images and sentences are presented in cause-and-effect sequence; and coordinated in that the images were presented contiguously with their corresponding sentences" (p.72).

In a third example, Mayer (1989) conducted two experiments with students who lacked prior knowledge about car mechanics. The students read a passage about breaking systems that contained labeled illustrations of the systems, illustrations without labels, labels without illustrations, or no labeled illustrations. Students who received passages that contained labeled illustrations recalled more explanative information than other groups and performed better on problem solving transfer tests. Ultimately, Mayer set out to determine exactly what can be done to improve the understandability of expository text and concluded that labeled illustrations can help guide learner attention in a way that will effect the aforementioned measures.

Certainly, instruction involving words and not pictures is far more common than instruction involving pictures but not words. As such, evidence based demonstrations/examinations of the multimedia principle most frequently involve adding pictures to words. However, Mayer and Anderson (1991) conducted an experiment in which they examined the comparative ineffectiveness of instructions containing pictures but not words. In the experiment, novices were randomly assigned to groups (no instruction, words-only, pictures-only, words and pictures) that received instructions about how a bicycle pump works. Participants were given a problem solving test and a recall test. On the problem solving test, the words and pictures group was superior to the other groups and the pictures only treatment was found to be only marginally better than the words only treatments. On the recall test, the words and pictures group outperformed all other groups. However, the words only group outperformed the pictures only group. As such, there is some evidence to suggest that the multimedia principle is applicable to both adding pictures to words and adding words to pictures.

Designers must be aware that the multimedia principle has been shown to lose its effectiveness with high experience learners. Kalyuga, Ayres, Chandler, and Sweller (2003) labeled this the expertise reversal effect. Kalyuga et. al reviewed empirical studies that demonstrated the expertise reversal effect and found that when fully guided instructional material is presented to more experiences learners, a part or all of the provided instructional guidance may be considered redundant. For expert learners, the redundant information is said to place an excessive and unnecessary load on limited working memory resources.

Evidence for the Contiguity Principle

Mayer (2008) summarized the research related to both the spatial and temporal contiguity principles. In each of the five tests related to the spatial contiguity principle involving paper-based lessons on brakes and lightning and computer-based lessons on lightning, learners who received integrated presentations (with the text located in close proximity to the part of the visual to which it referred), performed better on transfer tests than did students who received separated presentations. The median spatial contiguity effect size was 1.12. In each of the eight experiments related to the temporal contiguity principle, learners performed better when exposed to simultaneous presentation of animation and narration rather than successive presentation. The median temporal contiguity effect size was 1.31.

An example of the research addressing the spatial contiguity principle is the work done by Moreno and Mayer (1999). Moreno and Mayer had students (college age, low-knowledge learners) view a computer animation depicting the process of lightning. Students concurrently viewed on-screen text presented near the animation or far from the animation. Researchers measured learning by retention, transfer, and matching tests. The results of the experiment revealed that students learned better when visual and verbal materials were physically close. Moreno and Mayer concluded that, “learning is impaired when on-screen text is spatially separated from visual materials” (p.366).

Chandler and Sweller (1991) conducted an experiment to test the spatial contiguity principle. They surmised that according to cognitive load theory, effective instructional material facilitates learning by directing cognitive resources toward activities that are relevant to learning rather than toward those that are not. Chandler and Sweller posited that instruction requiring learners to mentally integrate disparate sources of mutually referring information such as separate text and diagrams would generate a heavy cognitive load because the material must be integrated before learning can occur and in turn, this extra processing would inhibit learning. Chandler and Sweller tested this hypothesis by using electrical engineering and biology related instructional materials, some of which were conventional and some of which were integrated. The results favored integrated instructions. Participants exposed to integrated instructions performed better in both written test and practical skill demonstrations.

Mayer and Anderson (1992), examining the temporal contiguity effect, conducted a series of two experiments where students studied an animation depicting the operation of a bicycle tire pump or an automobile braking system. Students studied the topics along with concurrent oral narration, animation alone, narration alone, or no instruction. Students who received animation before or after narration were able to solve transfer problems no better than students who had received no instruction. In contrast, when animation was presented concurrently with narration, students demonstrated a large improvement in problem-solving transfer over the control group. Mayer and Anderson (1992) concluded that, “the temporal contiguity principle encourages learners to build connections between their verbal and visual representations of incoming information, which in turn supports problem-solving transfer” (p.450). Mayer and Anderson suggest that the instructional implication of the study is that pictures and words are most effective when they occur contiguously in time and space.

A second example of research related to the temporal contiguity principle is the work done by Mayer and Sims (1994) who conducted 2 experiments in which they exposed high and low spatial ability students to a computer-generated animation while they listened to either simultaneously or successively presented narration explaining the workings of either a bicycle tire pump (in experiment 1) or the human respiratory system (in experiment 2). The findings indicated that the concurrent group generated more creative solutions to subsequent transfer problems. According to Mayer and Sims, results supported the dual coding theory of multimedia learning in that

> concurrent presentation of verbal and visual descriptions of a system increases the likelihood that students will be able to build connections between their mental representation of visually and verbally presented explanation and that a successive presentation makes it more difficult for learners to form referential connections. (p.399).

Evidence for the Modality Principle

Mayer (2008) explained that in 17 of 17 tests involving computer-based multimedia lessons on science and engineering topics, learners scored higher on transfer tests when the words were presented in spoken form rather than printed form. The median effect size for the studies was 1.02. Mayer emphasized that the modality principle helps manage essential processing by offloading the cognitive processing from the visual channel to the verbal channel.

Allport, Antonis, and Reynolds (1972) conducted two experiments related to the modality principle. In one of the experiments, students had to repeat an auditory prose passage (referred to as shadowing by the researchers) while memorizing verbal or nonverbal material. The material to be memorized was presented as either words in oral form, words in written form, or photographs. Results showed that when shadowing the auditory passage, memorization of the orally presented words declined significantly. Allport et. al concluded that the performance of two tasks in the same modality impaired learner performance.

Mayer and Moreno (1998) demonstrated the modality principle in a study in which participants viewed computer-generated animations depicting either lightning formation or a cars braking system. Students received either concurrent narration describing the major steps or concurrent on-screen printed text information using same words and timing. The results indicated that the narration group outperformed the printed text group in recalling the steps in the process on a retention test, in finding the named element in in a matching test, and in generating correct solutions to problems on a transfer test.

Craig, Gholson, and Driscoll (2002) conducted an experiment using a multimedia learning environment where they measured the effects of printed text, spoken narration, and spoken narration with printed text. A pedagogical agent within the multimedia presentation delivered the narration or text. The spoken narration only condition outperformed the other treatments in retention, matching, and transfer problems, with no differences between the printed text and the printed text with spoken narration. Craig et. al concluded that the findings were expected because of the additional processing requirements the printed text placed on the visual channel.

In another example of a study demonstrating the modality effect, Mayer, Dow, and Mayer (2003) had students learn about electric motors by asking questions and receiving answers from an on-screen pedagogical agent who stood next to an on-screen drawing of an electric motor. The students performed better on a problem-solving transfer test when the explanations were presented as narration rather than on-screen text. Researchers concluded that, “when designing a multimedia presentation that is intended to explain how something works, instructional designers should annotate the animation with spoken rather than printed text” (p.811).

In a related study, Schmidt-Weigand, Kohnery, and Glowalla (2010) examined visual attention distribution in learning from text and pictures. Researchers varied the modality of a multimedia presentation and recorded participant eye movement. Results showed that learners spent more time studying visualizations with the spoken text than with the written text. When presented with written text, learners consistently started reading before alternating between the text and visualization and they spent more time inspecting the text. The findings imply that if there is printed text on a screen, learners will spend time reading it at the expense of examining the visual.

Evidence for the Redundancy Principle

The redundancy principle, which states that people learn better from animation and narration than from animation, narration, and on-screen text, has been examined by several researchers who concluded that the on-screen text creates extraneous processing that diminishes cognitive capacity. Clark and Mayer (2011) cite four studies designed to demonstrate the redundancy principle and state that the median effect size for the four experiments was greater than 1. In the four studies, one group of participants viewed an animation and listened to a concurrent narration explaining the formation of lightning. Other participants received the same presentation, but with concurrent, redundant on-screen text. The participants in the nonredundant groups produced more solutions on a problem-solving transfer test than did the participants in the redundant groups.

In a related study, Mayer, Heiser, and Lonn (2001) had college students view an animated multimedia presentation on the formation of lightning. Students in the control group had to listen to a concurrent narration describing what was happening in each animation. In addition to the annotations and narration, one treatment group also received concurrent on-screen text that summarized the narration. A second treatment group was exposed to on-screed text that duplicated the narration. The two treatment groups performed worse on tests of retention an transfer compared to the control group. Mayer et. al (2001) explained that the additional on-screen text overloaded the visual information-processing channel, causing learners to split their visual attention between two sources.

Kayluga, Chandler, and Sweller (2004) conducted an experiment that verified the redundancy principle. One group of technical apprentices was exposed to simultaneously presented written and auditory textual information about how to operate cutting machinery. Another group was exposed to auditory only text. The participants were provided with four different sections of text (around 300-400 words each), the comprehension of which required the inclusion of pictorial information. Participants took a multiple choice test after receiving the treatments and the group exposed to the auditory-only condition performed better on the test with an associated effect size of .8.

More recently, Jamet and Le Bohec (2007) conducted an experiment, related to the redundancy principle, in which participants viewed a multimedia presentation. One group of students was presented with only spoken text. The second group was given redundant written sentences together with spoken information that were progressively presented on the screen. A third group was given the the written sentences all at once. The results indicated that both the sequential and static presentation of duplicated text led to poorer performance on retention and transfer tests, as well as in a task where the memorization of diagrams was evaluated.

Evidence for the Coherence Principle

The coherence principle, which states that people learn better when extraneous material is excluded rather than included, has been thoroughly researched. According to Mayer (2008), in 13 out of 14 experiments involving both paper and computer based lessons on scientific topics, students performed better on problem solving tests when they received a concise lesson rather than an expanded lesson. According to Mayer, the median effect size for the studies was .97.

Adding unnecessary audio is one of the ways that extraneous information can be added to a multimedia presentation. Moreno and Mayer (2000) conducted an experiment to measure this effect. In the experiment, all students viewed a multimedia presentation containing animation and concurrent narration. Students groups received either no additional audio, background music, background sound effects such as lightening striking, or both the music and the sounds. On transfer tests, the group receiving both sounds and music performed worse than groups not receiving superfluous audio and groups receiving sounds performed worse than groups not receiving sounds.

Adding unnecessary visuals is another way that designers can add extraneous information to a multimedia presentation. To test this concept, Mayer, Heiser, and Lonn (2001) had a group of students view a 3 minute narrated animation on lightning formation. A second group of students viewed a presentation containing an additional 1 minute of video clips intended to make the presentation more entertaining. The video clips were related to the topic (lightning) but were not germane to the instructional goal of teaching how lightning forms. Mayer, Heiser and Lonn found that students who viewed the presentation without the extra video clips performed better on transfer problems with an effect size of .86 . Similarly, Harp and Mayer (1997) had one group of students read a paper booklet about lightning formation. The booklet contained six paragraphs of text and six captioned illustrations. Other students read the same illustrated passage but with added color visuals related to the topic but not the instructional goal. Students who viewed the booklet without the added color photos performed better on retention and transfer tests, with an effect size over 1, compared to the group receiving the extraneous visuals.

Finally, unnecessary wordiness can be a source of extraneous information. Mayer, Heiser and Lonn (2001) had one group of students view a three minute narrated multimedia presentation on lighting formation. Another group was presented with an additional 1 minute of extraneous narrated text related to the topic but unrelated to the learning goal. Students who received the presentation without the extraneous information performed better on transfer tests with an associated effect size of .66.

Mayer, Bove, Bryman, Mars, and Tapangco (1996) conducted an experiment in which college students read a summary that contained 5 illustrations accompanied by a 50-word caption explaining the illustration. One group of students viewed the presentation with the standard captions while another group viewed the presentation with an additional 550 words of description for each visual. Students were asked to recall the steps of lightning formation and to solve transfer problems related to the topic. The summary group performed significantly better than the plus 550 words group on transfer problems. Mayer et al. (1996) concluded that, “the results are consistent with the idea that adding a lot more text to the captions can reduce students efficiency in abstracting the core verbal explanation and in connecting it with the visual explanation.” (p.72).

Evidence for the Personalization Principle

According to Mayer (2008), in 11 of 11 experiments involving computer-based lessons on lungs, lightning, botany, and industrial engineering, learners who received words in a conversational style performed better on transfer tests than did learners who received words in a formal style. The median effect size for the experiments was 1.11.

One of the eleven experiments, conducted by Mayer, Fennell, Farmer, and Campbell (2004, exposed one group of students to a personalized narrated explanation of a scientific topic and exposed a second group of students to a non-personalized explanation of the same topic. The narration for the non-personalized version was in a formal style and the narration in the personalized version was in a conversational style in which “the” was changed to “your” in 12 places. In 3 experiments, students receiving the personalized version scored significantly higher on transfer tests than students who received the non-personalized version. Mayer et al. (2004) concluded that, “making a seemingly minor change to 12 words in a multimedia lesson had a large effect on students' subsequent performance on tests of transfer, yielding a median effect size of .81 across three experiments” (p.393).

Moreno and Mayer (2004) conducted a second study in which they attempted to teach college students a science lesson in an agent-based virtual reality educational game. Students were presented with narration delivered in personalized speech (e.g., including I and you) or non-personalized speech (e.g., 3rd-person monologue). The information was presented via a desktop computer or a head-mounted display. Moreno and Mayer (2004) found that across both levels of immersion, students who received personalized messages performed better on retention and problem-solving transfer tests with an associated effect size of 1.64. Moreno and Mayer concluded that students learn more deeply from a personalized style of speech rather than a non-personalized style.

Recently, Kartel (2010) exposed participants to multimedia lessons on scientific topics. The lessons included illustrations and animation along with printed and spoken words. Words were presented either in a formal style, such as "The white dwarf cools down slowly in time" (p. 201) or a personal style, such as "The white dwarf cools down slowly in time. Now we know what will happen to our smallest star in the end" (p.201). Students who received the personalized text performed better on a problem-solving test with an associated effect size of .71.

Summary

There is a broad research base to support each of the multimedia learning principles. Supporting evidence for each of the principles is summarized in Table 4.

Table 4: *Summary of research based evidence*.
Appendix B

Design and Development Research

Overview

The developmental research orientation, as applied to education, emerged in the 1970s as the field of instructional technology shifted away from media comparison studies and began to emphasize "the process of systematic product design, development, and evaluation" (Richey, Klein, & Nelson, 2004, p. 1102). Recently, the approach has been renamed design and development research by Richey and Klein (2007), who indicate that such research attempts to create new knowledge grounded in data systematically derived from practice.

Main Procedures

Richey and Klein (2005) detailed three key stages in conducting developmental research. Stage 1 involves identifying and defining the research problem. An identified problem typically emanates from the workplace of the researcher, an emerging technology or from some aspect of design and development theory (Richey & Klein, 2007). Workplace problems can be related to something as small as an individual classroom or as large as an entire educational system. Richey and Klein (2005) urged researchers to focus the problem by giving it a "developmental twist," which focuses it on some particular aspect of the design, development, or evaluation process as opposed to focusing on an learning variable (Richey & Klein, 2005, p, 26).

Before beginning Stage 2, the literature review, researchers ideally will select the category of developmental research that they want to conduct. What was formerly called Type 1 (Richey & Klein, 2005) research is now called product and tool research (Richey & Klein, 2007). Here, researchers can study the design and development of an instructional product or program and have the option of studying the entire process or examining a specific aspect of the process such as formative evaluation or needs assessment. Tool research involves studying the development of a tool that can facilitate the design process.

What was formerly referred to as Type 2 research (Richey & Klein, 2005) is now called model research (Richey & Klein, 2007). Model research relates to the development, validation, and use of design and development models. As indicated by Richey and Klein, "Model studies focus on models and processes themselves, rather than their demonstration. Model research may address the validity or effectiveness of an existing or newly constructed development model, process, or technique" (p. 11).

Once a problem has been identified and a product, tool, or model approach has been selected, researchers should begin to review relevant literature and should write research questions based on the literature review. Product and tool literature reviews commonly focus on "identifying procedural models which might be appropriate for the task at hand, factors which have impacted the use of the target development processes in other situations and characteristics of similar effective instructional products, programs or delivery systems" (Richey & Klein 2005, p. 28). Model studies typically examine topics such as "a description of models similar to the one being studied, including their strengths and weaknesses and research on factors impacting the use of a given model or process" (Richey & Klein 2005, p. 28).

The final stage in the design and development process involves the selection of research procedures. Researchers must select participants, and often there are a number of them in a developmental research project including designers, developers, clients, organizations and learners (Richey & Klein, 2007). Researchers must select the appropriate methodologies for their study. It is not uncommon for researchers to use multiple research methodologies and designs, with different designs again being used for different phases of the project (Richey et al., 2004). Commonly used methods include case study, interview, field observation, evaluation and content analysis. Richey and Klein (2005) stated that generally, product and tool research uses case study and evaluation methods, while model studies commonly employ case studies, expert interviews, and Delphi studies. The use of appropriate methodologies helps researchers address concerns related to validity and causality and will help researchers interpret the findings of the study (Richey & Klein, 2007). Richey and Klein (2005) offered some detail of the common methodological dilemmas that confront design and development researchers. The first is that they often need to account for contextual variables, given the fact that most of this research takes place in a natural work setting. Also, researchers are often a participant in the study and thus must take care to insure objectivity.

The next step in this stage has researchers collecting and analyzing data. According to Richey and Klein (2004), "Data analysis and synthesis in a developmental study is not unlike that of other research projects. There are likely to be descriptive data presentations, and qualitative data analyses using data from documentation, interviews, and observations as well as traditional quantitative techniques" (p. 34). Finally, researchers have to interpret the findings of their study with the aim of developing design and development theory that is a part of the design and development knowledge base (Richey & Klein, 2007). Researchers should describe how findings can be used to expand the knowledge base and should carefully examine the data for lessons learned, to see what processes can be recommended, what specific conditions contribute to efficient design, and also what suggestions can be made to improve the product or tool.

Representative Examples

There are many examples of the design and development research approach being used to study the creation and validation of models (as is the case with this study) that attempt to help meet an authentic need while concurrently informing the overall design and development knowledge base. One such example is the model research conducted by Spector, Muraida, and Marlino (1992) in which the researchers identified the lack of a method for analyzing the relationship between user traits, authoring environments, and the resulting quality of computer-based instruction. To meet the perceived need, researchers designed and validated a cognitively based model of the computer-based instructional design process.

In a second example, Carliner (1998) documented efforts of a design team attempting to create three permanent museum exhibits. The data was used to develop a model of instructional design for informal learning in museums. A third example is the research done by Plass and Salisbury (2002) who found that most existing instructional design models were conceptualized to develop instructional solutions to needs and requirements that remain relatively stable over time. The researchers were forced to design a management system that needed to accommodate changing requirements and in an effort to address this challenge, they developed a design model based on a living-systems approach.

Summary

The design and development research approach is unique to the instructional design and technology field. Design and development research ultimately informs the design and development knowledge base by systematically studying the design, development, and evaluation processes and thereby establishing an empirical basis for the creation of instructional products, tools and models.

Research Questions

Three primary research questions have emerged from the review of literature:

1. What will the identified theoretical components look like when they are applied to AR enhanced learning?

2. How can the guidelines be designed so they are comprehensible by novice instructional designers?

3. How should the design of the guidelines be influenced by the affordances and limitations associated with AR enhanced procedural instruction?

Chapter 3: Research Methodology

Introduction

This study focused on the design and development and validation of guidelines that can be used to assist with the design of augmented reality (AR) enhanced instruction. The essential steps in this guideline development and validation process, discussed in detail as follows, were: (a) conduct a literature review in order to identify the learning outcomes the guidelines would target and to select specific theoretical components around which the guidelines were developed; (b) provide evidence that justified the selection of said components; (c) explain how the components were operationalized in the guidelines; (d) describe the form the finished guidelines would take; (e) explain how the guidelines were validated; (f) revise guidelines based on expert feedback; and (g) report results.

Research Design: Design and Developmental Research

The design and development research approach is unique to the instructional design (ID) field and aims to answer a call to strengthen the fundamental knowledge base of the field (Richey & Klein, 2007, p.2) and to, "study design, development and evaluation as well as do it" (Richey & Klein, 2005, p. 24). According to Richey and Klein (2007), model research, formerly referred to as Type 2 research, leads to more generalized conclusions and has as its ultimate objective, the production of knowledge, often in the form of a new or enhanced design or development model. Richey, Klein, and Nelson (2004) state that, "one of the most frequently addressed issues in this approach is providing evidence of the validity of a particular technique or model"(p. 1113).

When discussing the methodology of design and development research, Richey, Klein, and Nelson (2004, pp.1113-1117) explain that selecting an appropriate research problem is essential. Richey and Klein (2005) suggested that researchers should consider the following questions: Is the research problem common to many designers and developers? Is it currently critical to the profession? Does it reflect realistic situations faced by designers? Does it pertain to cutting edge technologies and processes? (p. 26). The problem this research attempted to address, the lack of a theoretically grounded and expert validated set of guidelines to assist with the design of instruction that uses AR as a delivery mode, seems to mesh well with Richey and Klein's suggestions.

As indicated by Richey, Klein, and Nelson (2004), the next step in the model development research approach was to conduct a literature review of related topics and to identify the learning outcomes the guidelines will address and the specific theoretical components around which the guidelines will be developed. In this case, AR was found to offer cognitive and psychomotor support for procedural learning and AR instructions have been found to use both words and pictures to deliver instructional content. As such, six theoretical components were selected from the multimedia learning theory base. Next, the components were used to design and develop the guidelines which were then shared with reviewers selected based on their expertise in the learning theory, instructional design, or AR domains. Experts validated the guidelines by following a researcher provided rubric. As indicated by Richey, Klein, and Nelson (2004), the expert reviews were used to reiterate and finalize the set of guidelines. Ultimately, design and development research is pragmatic in nature and is an applied method that is used to create a link between practice, research, and theory while solving practical problems (Richey, Klein and Nelson, 1996).

Procedures

Selection of theoretical components.

Six principles related to multimedia learning (Mayer, 2008) were used to guide this design and development research. All six principles have been thoroughly described in Chapter 2 of this document and are summarized as follows.

Multimedia principle.

The multimedia principle is built upon the belief that words and images evoke different conceptual processes. Clark and Mayer (2011) explain that multimedia presentations, containing static (drawings, charts) or active (animations, videos) graphics and words (printed, spoken) give learners the opportunity to make connections between the pictorial and verbal representations whereas instruction containing only words alone offers fewer opportunities to connect the words to other knowledge.

Clark and Mayer (2011, p. 72) stress that graphics must be carefully and intentionally added to text and advocate for graphics that depict changes in an object over time and graphics that depict relations among elements of an object. Fletcher and Tobias (2005) acknowledge that the multimedia principle is most effective for learners with low levels of prior knowledge and spatial ability but add that additional research is needed to clarify the ways in which these learner characteristics interact with the principle. Mayer and Anderson (1991) suggest that the multimedia principle can be applied to improve instructions containing words without pictures or instruction containing pictures without words.

Contiguity principle.

According to Mayer (2005c), the contiguity principle attempts to prevent the extraneous overload that happens when the cognitive processing required to understand material in a multimedia message is overcome by the processing required to process extraneous material or confusing layout in the multimedia message. Fletcher and Tobias (2005) suggest that verbal and pictorial information should be coordinated in both space (spatial contiguity) and time (temporal contiguity).

Mayer (2005c) explains that spatial contiguity simply attempts to lessen the cognitive effort required on the part of the learner by placing an illustration next to the paragraph that it describes and thereby prevent the learner from having to scan back and forth between the text and the visual.

Mayer (2008) states, “the temporal contiguity principle is that people learn better when corresponding narration and animation are presented simultaneously rather than successively” (p.765). Clark and Mayer (2011) explain that, “when instruction separates corresponding words and graphics, learners experience a heavier load on working memory, leaving less capacity for deep learning” (p.103). They (2011) claim that having the pictures and words coordinated allows learners to hold them together in their working memories and therefore make meaningful connections between them and suggest that the contiguity principle may most strongly apply when dealing with low-knowledge learners and when graphics and words are complex.

Modality principle.

Low and Sweller (2005, p. 147) explain the Modality Principle by stating that, “under certain, well-defined conditions, presenting some information in visual mode and other information in auditory mode can expand effective working memory capacity and so reduce the effects of an excessive cognitive load.” Clark and Mayer (2011) recommend that words should be presented in spoken form whenever a graphic is the focus of the words and both are presented simultaneously because requiring the visual channel to process both the graphic and the printed words threatens to overload the channel.

Low and Sweller (2005) explain that the logical relation between the two sources of information is critical and indicate that the modality principle is not applicable unless the two sources of information are unintelligible in isolation. Low and Sweller stress that care must be taken to ensure that auditory material is essential and not redundant and that the instructional material is sufficiently complex to warrant the use of a cognitive load reducing technique. Mayer and Clark (2011) caution that if words are technical, unfamiliar, or not in the learners native language, they should remain available to the learner in visual form.

Redundancy principle.

According to Sweller (2005b, p.159), the redundancy principle suggests that, “redundant material, the same information presented in multiple forms or unnecessarily elaborated, interferes with rather than facilitates learning”. Clark and Mayer (2011) suggest that many designers may be compelled to add both printed text and spoken text in order to cater to students that have visual learning styles and students that have auditory learning styles. Clark and Mayer indicate a lack of evidence to support such a decision. Sweller (2005b) indicates that adding redundant material requires learners to unnecessarily coordinate and relate multiple forms of the same information, which causes an extraneous cognitive load that interferes with learning.

Coherence principle.

Mayer (2005c, p.183) states that, “people learn more deeply from a multimedia message when extraneous material is excluded rather than included”. Clark and Mayer (2011) caution designers against adding any extraneous material, such as pictures or dramatic stories, in an attempt to make instruction more engaging for learners because the extra material will cause learners to devote valuable cognitive resources towards processing the non-essential material. Clark (2005) states that, “extraneous material should be ruthlessly weeded out whenever possible so that the core of the essential material is salient to the learner”(p.196).

Personalization principle.

Mayer (2005d) states, “people learn more deeply when the words in a multimedia presentation are presented in a conversational style rather than a formal style” (p.201) and suggests (2008) that a conversational style or polite wording creates a sense of social partnership with the narrator, which causes the learner to try harder to make sense of what their conversational partner is saying.

Clark and Mayer (2011) warn designers to avoid adding slang and jargon that can distract the learner and set an inappropriate tone for learning. Essentially, the personalization principle assumes what is said and how it is said are important factors to consider when designing a multimedia instructional message.

Selection criteria.

The six theoretical components detailed above have been carefully selected based on the four criteria described as follows.

Selection criteria #1-the components are applicable to instructional uses of AR.

As indicated by Table 1, AR enhanced instruction almost always involves words and pictures. For example, in a study by Neumann and Majoros (1998), AR was used to deliver both verbal and visual information to learners. Henderson and Feiner (2011a, 2011b) also developed and tested AR instructions in words and pictures were used to deliver information to learners. De Cresenzio et al. (2011) designed and tested instruction in which AR delivered visual cues to direct learners to printed material that could be used to manipulate machine parts. In the studies by Liu (2009) and Tsung-Yu et al. (2009) AR was used to deliver both verbal and visual information that attempted to help learners develop their foreign language speaking skills.

As defined by Mayer (2005a), multimedia instruction " involves presenting words and pictures that are intended to promote learning" (p. 3). As indicated by the examples above, words and pictures are used to deliver information in AR enhanced instruction. As such, it seems reasonable to consider AR instruction a form of multimedia learning in turn, to use the most relevant and established principles from the multimedia learning theory base to design a set of ID for AR guidelines.

Selection criteria #2- all components are informed by a theory of human learning.

All six of the theoretical components selected for use in this design and development research have been informed by the cognitive theory of multimedia learning. CTML was discussed at length in the Chapter 2. CTML, which is strongly influenced by the work of Baddeley (1999), Paivio (2007), and Sweller (1991, 2005a, 2005b), is built upon three cognitive science principles of learning: that humans have separate channels for processing visual/pictorial information; the channels are limited in capacity; and learning is an active process that requires coordinated cognitive processes. According to Mayer (2005b) CTML is consistent with cognitive science principles of learning, yields predictions that can be tested, is consistent with empirical research, and is relevant to educational needs for improving the design of multimedia instructional messages.

Learning occurs, according to CTML, when learners select relevant words from text or narration, select relevant images, organize both the words and images into coherent verbal and pictorial representations and integrate the representations with new knowledge. According to Mayer (2005b), use of CTML during the ID process is important because instruction designed in accordance with a model of how the human mind works will increase the chances that meaningful learning will occur.

Selection criteria #3- components are well supported by research based evidence.

All six of the components have been thoroughly researched and are supported by empirical evidence. An analysis of the supporting evidence is provided in Chapter 2 (pp. 41-53).

Selection criteria #4-all components are adaptable to an AR delivery mode.

According to Mayer (2005a), examples of multimedia learning include listening to a live teacher giving a presentation that includes visuals on a chalkboard, reading a textbook with words and pictures, watching a video that includes narration, and viewing a PowerPoint presentation that contains instructor commentary. All of the aforementioned delivery methods are considered multimedia because they are able to present both verbal and visual information. As explained earlier, there are numerous reported examples of AR being used to deliver instructional messages involving both words and pictures.

The hardware and software on the devices on which AR content is typically accessed allow for words (both printed and spoken) to be provided to learners as well as all types of visuals (static visuals such as images or dynamic visuals such as animations). An example of AR enhanced instruction involving both words and pictures is the study conducted by Neumann and Majoros (1998) who attempted to teach mechanics a multi-step maintenance task. Another example of AR delivering multimedia instruction is the study conducted by Squire and Jan (2007), in which learners worked together to solve a problem while interacting with digital visuals and text experienced on a mobile device. Essentially, because the devices on which AR is experienced are able to deliver both words and pictures, the identified multimedia principles can be applied to AR enhanced instruction.

Summary.

Four specific criteria for the selection of the theoretical components, summarized in Table 5, have been presented. The fact that the components are applicable to AR instruction, are grounded in a theory of human learning, are supported by a broad empirical research base and are adaptable to an AR delivery mode suggests that they are suitable for use in this guideline development research.

Table 5: *Summary of theoretical component selection criteria.*

Criteria	Explanation
1. The multimedia principles are applicable to AR instructional message design.	There are numerous instructional uses of AR in which information is delivered in both visual and verbal form. As such, the multimedia learning principles can be used to inform a set of guidelines aimed at assisting with the design of AR enhanced instruction.
2. The principles are informed by an underlying theory about how humans learn.	Mayer's cognitive theory of multimedia learning, which is itself informed by dual coding theory (Paivio) and cognitive load theory (Baddeley, Sweller) underlies all of the principles. CTML posits that in order to learn, humans must effectively select relevant information, mentally organize that information and then integrate that information with long term memory and other incoming information. CTML assumes that the human information processing system includes dual channels for visual/pictorial and auditory/verbal processing, that each channel has a limited capacity for processing, and that active learning entails carrying out the aforementioned cognitive processes during learning.
3. All of the principles are supported by research based evidence.	Researchers have conducted a wide variety of quality experiments that combine to provide a broad base of evidence based support for each principle.
4. The principles are adaptable to AR.	Currently, AR is almost exclusively experienced on a smartphone or tablet. The devices posses the hardware and software necessary to overlay visual information and to overlay or play verbal information. As such, the technical means by which AR is experienced are able to facilitate multimedia learning.

Operationalization of components in the guidelines.

The initial steps in the guideline development process were determining the appropriate AR learning outcomes to target and selecting the appropriate theoretical components to be adapted in the set of guidelines. Next, the focus shifted to adapting each of the principles in a way that was congruous with the attributes and affordances associated with AR and to explaining the adapted principles in a way that helped novices understand the underlying learning theory. As has been discussed in Chapter 2, AR, when accessed via a mobile device or a wearable display can overlay digital information, ranging from sounds, 3D models, videos, drawings and on a view of the physical world. When being used to teach procedural tasks, this digital information can be used to help minimize the demands placed on the learner during the cognitive and psychomotor phases of the learning event. In the finished set of guidelines, each of the adapted principles was explained with an emphasis on the underlying learning theory. Also, each adapted principle led to the inclusion of an example, consistent with the related attributes and affordances, of AR being used to facilitate procedural learning. To support the development of the examples, the suggestions by Clark and Mayer (2011) about applying the multimedia learning principles in a computer based environment were used for guidance.

Validation of the guidelines.

As indicated by Richey, Klein, and Nelson (2004), qualitative research methods are often employed in Type 2 (model) development studies. Richey et al. note that, "the nature of the conclusions in Type 2 research often pertains to the evidence of the validity and/or effectiveness of a particular technique or model" (p.1113). As such, the developed guidelines were reviewed by experts to determine their validity.

Regarding the review of the set of proposed guidelines, reviewers with backgrounds from one of three distinct but related domains were used to evaluate the product. An instructional design expert focused on how well the instructional material was designed and how well it met its intended purpose, a AR expert provided insight about how well the guidelines meshed with the affordances, benefits, and limitations related to learning with AR, and finally, four experts in the educational psychology/multimedia learning domains provided feedback on how well the guidelines meshed with the underlying learning theory. The following chart details the experts who served as reviewers:

Table 6: *Expert reviewers*

Name	**Institution**	**Primary Area of Expertise**
Dr. Kathy Cennamo	Virginia Tech	instructional design, application of learning theories to instructional products
Dr. Peter Doolittle	Virginia Tech	educational psychology, multimedia cognition
Dr. Steven Feiner	Columbia University	human-computer interaction, augmented reality
Dr. Richard Mayer	University of California, Santa Barbara	educational psychology, cognitive theory of multimedia learning
Dr. Fred Paas	Erasmus University Rotterdam	educational psychology
Dr. John Sweller	University of New South Wales	educational psychology, cognitive load theory

Evaluation rubric.

A researcher developed rubric (Appendix D) was provided to the expert reviewers. The rubric was designed to avoid limiting reviewers to commenting only on their identified domain of expertise. This strategy was employed in an attempt to collect a maximum amount of data during the review process. The rubric was created using the theory selection criteria described earlier in this chapter. The first section of the rubric focused on each of the six principles that were adapted in the initial draft of the guidelines. Here, experts were asked to evaluate the how effectively each selected principle was adapted and how well supporting evidence was explained. Also reviewers were asked to examine the usefulness of the included examples.

The second section of the rubric asked participants to evaluate the overall design and applicability of the guidelines. Once the evaluation rubrics were returned, the expert responses were to reiterate the guidelines. As will be discussed in Chapter 5, the qualitative responses were collected, analyzed, and used to reiterate the guidelines. Ultimately, the review process sought to determine if the proposed guidelines would be of use to novices who are attempting to design instruction that has content delivered via AR.

Presentation of the guidelines.

Primarily out of convenience for the reviewers, the initial version of the proposed guidelines was presented in the form of a multimedia website. The website started with an Introduction page on which the history, definition, instructional uses, and recent advances related to AR were explained. Next, reviewers were presented with a Purpose page that attempted to explain how the guidelines could be used to effectively manage words and pictures while designing AR enhanced instruction. On the following page, reviewers were presented with a summary of the learning theory upon which all of the multimedia instructional principles are built. Finally, each of the adapted principles was explained and a summary of related research was included. For each principle, a misapplication and a successful application were presented.

For example, upon selecting the modality principle, users were taken to a landing page for that specific principle. Here, in order to acquaint them with the principle, users were provided with an explanation of the principle and brief summaries of two research studies that provide the empirical evidence behind the principle. Next, users were presented with a non-example screen-capture that showed a violation of the modality principle, as related to learning via AR. For the modality principle, the non-example was similar to the use of AR in the Neumann and Majoros (1998) study where both a large amount of text and directive visuals were overlaid on an already complex visual environment, thereby overtaxing the visual channel of the subject.

Finally, a proper application how the of the modality principle can be applied to AR learning was presented to users. Here, users were presented with a mocked-up screen capture showing the text information switched over to verbal narration, thereby dividing the cognitive load among two channels. At this point, reviewers were able to select another adapted principle and the same process was replicated for each component.

Conclusion

Six theoretical components (Mayer, 2008), all of which are related to the effective design of multimedia instruction, were selected to influence the development of a set of guidelines that will help users effectively design procedural instruction to be delivered using augmented reality.

Four specific criteria were provided for the selection of the components. The criteria include the suggestion that components are applicable to AR instruction, that they are supported by a theory of human learning, that a broad empirical research base supports them, and that they are adaptable to the ways in which AR can be used instructionally. The components were operationalized into a set of guidelines that were presented in a multimedia website. Six experts from three different domains of expertise reviewed the developed product. As discussed in Chapter 5, the results of the review were used to reiterate the guidelines and the results have been reported in this document.

Chapter 4: Constructing Design Guidelines for Procedural Instruction Delivered via Augmented Reality

Introduction

As indicated by Richey and Klein (2005), design and development research must begin with the identification of a research problem. Richey and Klein suggest selecting a problem that is common to many designers and developers and that pertains to cutting edge technologies and processes. The process and rationale behind the identification and definition of the research problem for this study has been described in detail in the preceding chapters and are summarized here.

Augmented reality, the concept of overlaying digital information on a real-time direct or indirect view of the physical world is considered an emerging technology (*The Horizon Report*, n.d.). Presently, the general public experiences AR by holding a smartphone or tablet, equipped with a camera, directly in their field of vision. Most of the current applications of AR relate to entertainment, marketing, or the presentation of verbal information in the form of labels and facts. The pending release of mass-market wearable displays holds the potential to radically transform how and how much the general public will experience AR. Google, a manufacturer of wearable AR displays, acknowledges (Doerr, 2013) that the ways in which the devices will shape AR usage have yet to be determined and suggest that entrepreneurs and engineers will play a huge role in shaping the evolution of AR.

Regarding how AR may be of use to instruction and learning, the body of research related to AR, informed by scholars and engineers such as Azuma (1997), Caudell (1992), Feiner (2009, 2011a, 2011b), Henderson (2009, 2011a, 2011b), Neumann (1998), Majoros (1998), and Minzell (1992), seems to offer a suggestion that could potentially influence how AR is used by the general public. A review of related literature indicates that AR has been used to assist with the performance/learning of a variety of procedural tasks related to assembly, maintenance, and repair. Building on the work of Neumann and Majoros (1998), Henderson and Feiner (2011a, 2011b) have identified how AR provides benefits related to both the cognitive (activities such as directing attention, comprehending instructions, transposing information from instructions to the physical task environment) and psychomotor (activities such as comparing, aligning, adjusting, and manipulating) aspects of procedural learning. Instructions delivered via AR, when compared to other delivery modes, have been shown to reduce task completion time, error rate, and perceived mental workload.

At present, the use of AR to assist with procedural tasks has generally been limited to higher education institutions and the military (De Cresenzio et al., 2011; Haritos & Macchiarella, 2004; Henderson & Feiner, 2009; Henderson & Feiner, 2011a, Henderson & Feiner, 2011b, Neumann & Majoros, 1998, Tang et al., 2001), likely due to the associated costs of developing the necessary software and purchasing a useable head-mounted-display. However, with relatively affordable wearable displays soon to be available for use by the general public, it is possible that entrepreneurs and engineers, looking to maximize the affordances of the new wearable displays, will be able to apply the benefits of AR to procedural tasks that must be learned/completed by the average person. Regarding the mass-market displays, the key affordance of such devices (Calhoun & McMillan, 1998) is the fact that they user will be able to experience AR with full use of their hands, an affordance that seems to perfectly align with the documented benefits of using AR to facilitate procedural learning.

The way in which the average user experiences AR is about to change. Related research indicates that AR provides cognitive and psychomotor benefits (Henderson & Feiner, 2011a, 2011b) during the procedural learning process. New wearable displays will afford users full use of their hands, thus allowing for the completion of procedural tasks. However, entrepreneurs and engineers, looking to shape the way in which AR is used, may be unaware of the benefits of using AR to facilitate procedural learning. Also, entrepreneurs and engineers will likely be untrained as instructional designers and will be unaware of message design principles that can be applied to the development of instructional AR applications. As such, the guidelines designed and developed as a result of this study aim to raise awareness of the benefits of using AR to facilitate procedural learning and to introduce some principles that novice designers can apply to their AR learning applications during the development process.

A review of AR literature indicates that despite the visual emphasis of AR, when the delivery mode is used instructionally, both words and pictures are involved (DeCresenzio et al., 2011; Gandy et al., 2005; Haritos & Macchiarella, 2004; Henderson & Feiner, 2011a, 2011b; Neumann & Majoros, 1998; Quarles et al., 2008; Reif et al., 2010; Rusch et al., 2012; State et al., 1996;). Therefore, AR instruction can be considered another form of multimedia instruction as defined by Mayer (2005a). The multimedia leaning research base (Mayer, 2008) has produced a set of evidence-based principles that attempt to help designers correctly use words and pictures while designing instruction. In the guidelines produced as a result of this study, the learning theory underlying the multimedia principles has been explained, and applicable multimedia principles have been identified and adapted to an AR delivery mode.

As will be discussed in the forthcoming chapter, once a problem was identified and a model development approach was selected to address the identified problem, an initial version of the guidelines was designed and developed using an established instructional design model. As discussed in Chapter 5, a panel of expert reviewers then evaluated the initial draft of the guidelines and the guidelines were revised based on the expert feedback.

Guideline Design and Development

Model Selection

The Dick, Carey, and Carey (1990) instructional design model was used to assist with the development of the initial draft of the guidelines that aim to provide support to novices who are designing procedural instruction to be delivered via AR. As explained by Gustavson and Branch (2002), "instructional design models function as tools that can be used to visualize, direct, and manage processes for creating high quality instruction" (p.1). The Dick, Carey, and Carey (DC&C) model was selected because, as indicated by Gustavson and Branch (2002), the model has been widely utilized by instructional designers, is the standard to which all other ID models are compared, and is applicable to the product-oriented design activities that took place in this study. Dick and Carey (1990) indicate that the model is based on theory, research, and practical experience. Dick and Carey suggest that instructional designers use their model in the predefined sequence. Below, the development of the initial draft of the guidelines is discussed in order of the predefined steps in the DC&C model.

INSTRUCTIONAL DESIGN GUIDELINES FOR AR

Assess Needs to Identify Goals

Dick and Carey (1990) explain that the first step in the DC&C model is to determine exactly what learners should be able to do after they have been exposed to the instruction. Given the AR trends and research reported earlier, four instructional goals for the set of guidelines were identified. First, upon reading the guidelines, learners should be able to explain the link between AR and procedural learning. Second, learners should posses a fundamental understanding of the information processing (Baddeley, 1999; Paivio, 2007) model of the human mind. This goal is of particular importance because all of the multimedia principles that were adapted are grounded in an information processing perspective. An incomplete understanding of the human mind from the information processing perspective will inhibit learner understating of why and how to apply the AR instructional message design guidelines.

Third, learners must understand the three fundamental assumptions upon which all of the multimedia principles are based (Mayer, 2005b). If learners do not understand the consequences of ineffective use of words and pictures, they may lack motivation to design effective instruction. Finally, upon reading the guidelines, AR application developers should be able to apply each of the adapted multimedia principles to their own instructional designs. In total, the goals attempt to strengthen/clarify the link between AR and procedural learning and to address the problem that most developers and programmers likely lack a basic understanding of instructional message design techniques.

Conduct Instructional Analysis and Analyze Learners and Contexts

As indicated by the DC&C model (1990), the next two steps in the guideline design process, learner/context analyses and an instructional analysis, were completed in conjunction with each other. Guideline readers will likely be application developers that are looking to use AR to deliver content. Given the fact that learners will have sought out the set of guidelines, they will likely be motivated to learn from them and apply the learning to their development process. While the learners may possess sophisticated programming skills, they may know little about human learn theory and instructional message design principles. Therefore, all content related to learning theory and message design was developed with novices in mind.

All of the aforementioned learning goals aim to teach intellectual skills. If the guidelines are effective, learners will be able to take the message design techniques discussed in the guidelines and apply them to their own instructional development. Ultimately, three of the instructional goals serve as subordinate learning that must take place in order to meet the fourth goal. Readers must first understand/value the link between AR and procedural learning (Henderson & Feiner, 2011a, 2011b) and understand the theoretical suppositions of the information processing perspective (Baddeley, 1999; Paivio, 2007) and CTML (Mayer, 2005b). With this subordinate understanding in place, learners will be able to more fully comprehend the adapted principles and will then be better equipped to apply the information in the guidelines to their instructional design efforts.

Write Performance Objectives

Based on the identified goals of the guidelines, drafting performance objectives was a straightforward process. First, upon reading the guidelines, learners should be able to recognize the benefits of teaching procedural content via AR. Second, upon reading the guidelines, entrepreneurs and developers should be able to apply the adapted message design principles to their own AR instructional designs.

Develop Instructional Strategy

The guidelines are meant to be used by novice instructional designers who are in the process of developing an instructional application that uses AR to deliver content. The novices will read the guidelines on their own without the benefit of a live instructor. As such, the instructional strategy that governed the design and development of the guidelines was to present all relevant material clearly and directly in an attempt to avoid misunderstanding or misapplication. This strategy is reflective of a belief that novices learn best when they are presented with direct, explicit instruction that fully explains the background knowledge and message design concepts that they are being asked to learn.

As will be discussed in Chapter 5, the revision of the guidelines led to a slight modification of the instructional strategy. As a form of preinstruction, an illustrative scenario of how AR may be used to deliver instruction to the general public was inserted at the beginning of the guidelines. The manner in which the selected principles were explained was reworded in an attempt to make the underlying concepts easier to understand. Also, misapplied message design principles were included at the end of the guidelines in an attempt to provide an opportunity for learner practice and feedback.

Develop and Select Instructional Materials

Next, the instructional strategy of presenting all relevant information in a clear and direct manner was applied to the development of the instructional materials. For the convenience of the expert reviewers, the initial draft of the guidelines was presented in a website. Given that the instructional materials contained both words and pictures, the very same multimedia principles being discussed in the guidelines were also used to influence the design and development of the guidelines. Based on the recommendations of the multimedia literature base, text was worded using a conversational tone, only essential material was added to the website, and carefully chosen printed words were matched with carefully selected visuals.

Ultimately, drafting the content for the initial set of guidelines was straightforward given the work that had been done during the literature review process. The initial version of the guidelines started with an Introduction page on which the history, definition, instructional uses, and recent advances related to AR were explained. Next, reviewers were presented with a Purpose page that attempted to explain how the guidelines could be used to effectively manage words and pictures while designing AR enhanced instruction. On the following page, reviewers were presented with a summary of the learning theory upon which all of the multimedia instructional principles are built. Finally, each of the adapted principles was explained and a summary of related research was included. For each principle, a misapplication and a successful application were presented.

As a result of the expert feedback discussed in Chapter 5, the final draft of the guidelines is now in document form. The “redundancy” and “coherence” principles have been combined, text explanations have been edited to make them more accessible for those who lack an instructional design background, the same scenario has been used for all examples, the link between each of the adapted principles and the underlying learning theory has been made more apparent, and each of the adapted principles has been explained using a short descriptive sentence rather than the name of the multimedia principle that was being adapted.

Design and Conduct Formative Evaluation

Once a draft of the guidelines was completed, the product was evaluated by a group of expert reviewers. As was discussed in Chapter 3, the reviewers were selected based on their expertise in the learning theory, instructional design, or AR domains.

Reviewers were provided with a rubric that was built around the theory selection protocols detailed in Chapter 3. Essentially the rubric was designed to solicit feedback about: (a) the overall rationale behind the guidelines; (b) how well the link between AR and procedural learning was explained; (c) the applicability and appropriateness of the selected theoretical components; (d) how well the theory underlying the selected components was explained; (e) how well each of the selected components was adapted to an AR instructional delivery mode; and (f) the overall design of the guidelines. The rubric was designed to allow each of the reviewers to comment on any aspect of the proposed guidelines.

Ultimately, six pieces of feedback were received and analyzed. The full analysis and reiteration process has been discussed in Chapter 5. The expert review process yielded several suggestions about how to improve the guidelines and proved to be an effective form of formative evaluation.

Design and Conduct Summative Evaluation

As is discussed in Chapter 5, the initial draft of the guidelines has been carefully revised based on feedback obtained from the expert review process. The guidelines are now ready to be shared with entrepreneurs and programmers who are attempting to develop AR enhanced instructional applications. Such use by independent evaluators will serve as an ideal form of summative evaluation and will help determine the ultimate usefulness and relative value of the guidelines.

Conclusion

Once a viable research problem was identified and the appropriate components from the AR and multimedia learning literature bases were selected, the initial draft of the guidelines, aided by the DC&C model (1990), was relatively straightforward. As described in this chapter, instructional goals were identified, an instructional analysis was conducted as were learner and context analyses, performance objectives were written, an instructional strategy was developed, instructional materials were developed, and a formative evaluation, in the form of an expert review, was conducted. The results of the expert review and the subsequent revisions to the initial draft of the guidelines are discussed in Chapter 5.

Chapter 5: Expert Review

An expert review process was used to validate the proposed guidelines detailed in Chapter 4. The review process addressed in this chapter will include a description of the coding process used to organize the qualitative data provided by the expert reviewers, a description of the three themes that emerged from the expert review, and a discussion of how the feedback associated with each of the themes were applied to the final draft of the AR instructional design guidelines.

Introduction

As explained by Richey, Klein, and Nelson (2004), developmental research studies often target the validity and/or evidence for the effectiveness of a particular model (in this case, a set of guidelines) and expert reviews are commonly used validate the proposed product. In this study, the guidelines for helping novices design instruction that attempts to teach procedural tasks via AR were validated by a carefully selected case of expert reviewers.

Reviewers were selected based on their expertise related to either educational psychology/human learning theory, augmented reality, or instructional design. The strategy of triangulating data from experts in three distinct areas was adopted to provide a comprehensive review of the guidelines from all related perspectives. A total of six experts (Table 6) agreed to review the guidelines using a predesigned evaluation rubric. Dr. Kathy Cennamo (Virginia Tech) served as an expert in instructional design practices, Dr. Stephen Feiner (Columbia University) served as an expert in AR use, and Drs. Peter Doolittle (Virginia Tech), Richard Mayer (University of California Santa Barbara), Fred Paas (Erasmus University Rotterdam), and

John Sweller (University of New South Wales) served as experts in cognitive principles of human learning and multimedia learning domains.

All reviewers used the same rubric to evaluate the guidelines. The rubric was designed based on the theory selection criteria described in Chapter 3. The rubric avoided restricting reviewers to commenting only on their domain of expertise. For example, Dr. Doolittle provided valuable feedback on instructional design issues in addition to theory selection/application and Dr. Feiner provided feedback on theory application in addition to AR use.

Qualitative Data Analysis

Six responses (three completed rubrics and three letters), varying in perspective, length, and detail were collected. Given the variety and thoroughness of responses, bringing order to the expert feedback was of paramount importance. As such, a three level coding approach that is frequently used by qualitative researchers (Charmaz, 2006; Hahn, 2008) was applied to the expert reviews. First, first a level of initial/open coding was done by reading through the responses and assigning a brief description to each recommendation. Next, a second level of focused coding reexamined the initial codes and ultimately led to the development of eight categories. Finally, a third level of thematic coding analyzed the identified categories and assigned them to one of the three overall themes that emerged from the review of the feedback.

Overall Perspectives

Broadly, expert feedback indicated that reviewers believed the fundamental premise behind the guidelines, that AR used instructionally to teach procedural tasks is indeed an emerging form of multimedia learning and that such use will place great demands on working memory, was sound. For example, Dr. Paas stated that, “In general I think this is an interesting topic and you did a good job working it out”, Dr. Sweller stated that “The proposed guidelines look great to me”, and Dr. Doolittle explained that, “I think the combination of CTML and AR is powerful and AR without CTML will be a working memory nightmare. I fully support this endeavor and think the project is on the right track, but the devil is in the details”. Dr. Doolittle's comment is indicative of the fact that while the proposed guidelines were generally headed in the right direction, they were in need of several revisions.

Coding Results

An application of the three level coding approach resulted in the emergence of three overall themes for the suggested revisions. The first theme, Guideline Design, is based on feedback related to the overall design of the guidelines. Areas targeted for improvement under this theme include the selection of appropriate visuals and the use of effective cueing strategies. The second theme, Content Level, is based on feedback related to the level at which the guidelines are presented. Areas targeted for improvement related to this theme include the removal of several technical terms and a complete reconceptualization of the way in which the underlying learning theories are presented. The third overall theme that emerged from the coding process, Content Detail, is based on feedback that relates to what is being said in the guidelines. Areas targeted for improvement here include combining the coherence and redundancy principles and better explaining the multimedia principle.

The Table 7 is meant to visually convey how the three overall themes that emerged from the expert feedback are made up of related categories that are, in turn, informed by individual pieces of feedback. For example, note that the theme Guideline Design contains the category Cueing Strategies, which is informed by specific comments such as, “If the colors are designed to employ the cueing/signaling principle, then there are too many and it is unclear why these words are important” and “ I suggest you minimize your use of color highlighting of words”.

Table 7: *Expert feedback overview.*

Theme	**Examples**
1. Guideline Design:	
Layout	* “Reorder the examples”. * “Use colons instead of dashes at the end of each principle of learning”. * “Tells us somewhere that, for each principle, you will present an explanation, a research sample, and examples”.
Visuals	* “The picture is simply window dressing”. * “Present text balloons in such a way that makes it clear to the reader that it is spoken text.”.
Cuing Strategies	* “If the colors are designed to employ the cueing/signaling principle, then there are too many and it is unclear why these words are important.” * “ I suggest you minimize your use of color highlighting of words”.
2. Content Level:	
Explanations	* “Contextualize the guidelines in the first paragraph of the introduction page through a scenario”- * “For each principle, clearly state the principle related to AR”. * “Use a single example across all six principles”.

Theme	Examples
Word Choice	* “Use terms consistently”. * “Do not use the terms effect size and control group”.
Technical Language	* “ Explain the cognitive science principles of learning like you would explain them to your mother or father”. * “These terms will not have any meaning to your average reader”. * “Explain it, cut it out, or otherwise reword so it is clear the novices to which this is targeted”. * “Your users will not be active researchers, but rather, designers. Speak to them”. * “The page on theory selection contains too much jargon”.
3. Content Detail:	
Definitions	* “Tell us somewhere that you are focusing on procedural learning as this is the type of learning outcome for which AR is most often used”. * “When presenting the principles for the first time, present some overview text that ties them together”.
Concept Application	* “One of the key advantages of AR is that the virtual instructional material and the real domain objects can be co-located, eliminating the need for users to switch the attention between the two. I think you need to stress this issue more”. * “Be careful of overly broad generalizations”.
Principle Application	* “Further explain the effect of integrating visual + auditory in the multimedia principle”. * “I think a little high-level preview text is not a bad idea”. * “What you call the coherence principle, we call the redundancy principle as they both have the same effect for the same reason”.

In total, the expert reviewers provided 67 distinct comments. The Design theme received 15 pieces of feedback, the Content Level theme received 13, and 37 pieces of feedback addressed the Content Detail theme with all six of the reviewers making at least one suggestion for improvement in this area.

Resultant Changes

In this section, each of the three identified themes will be addressed individually. For each theme, applicable suggestions will be described and the effect that the suggestions had on the revision process and the final draft of the guidelines will be detailed. Individually addressing each of the three themes helped ensure that all suggested revisions were given the proper attention and played an appropriate role in determining the form of the final product.

Theme 1: Guideline Design Issues

Only one reviewer was identified as an instructional design expert yet five of the six reviewers make at least one suggestion related to the overall design of the guidelines. The choice of visuals for the Introduction and Theory Explanation pages was called into question and reviewers suggested selecting visuals that more accurately aligned with the text. Also, a cuing strategy that was meant to call attention to the components of the definition of AR was cited as ineffective.

Reviewers took issue with the manner in which non-examples were used in the initial draft of the guidelines and suggested rethinking their use. Also, reviewers felt that when the principles were introduced for the first time, a short text-based preview of what was to be shown for each principle (description, research example, enactments) would have been beneficial.

Theme 1: Resolution of Guideline Design Issues

The feedback related to the Guideline Design theme led to some substantial changes in the final form of the guidelines. On the Introduction and Theory pages, visuals were selected that better aligned with the text. For example, the visual of a person wearing AR glasses on the Introduction page was replaced with a picture that better indicated what a person would see if they were wearing an AR display. The unnecessary coloring on the Introduction page was removed. Most importantly, the non-examples were removed from the main section of the guidelines. Given the fact that novice users are likely experiencing these principles for the first time, pedagogically, it seemed more effective to present learners with clear examples first and avoid the risk of confusing them with potentially ineffective non-examples. Some of the non-examples were moved to a quiz-like assessment at the end of the guidelines. Finally, at the start of the principle overview section, a short preview was included to help the readers orient to the forthcoming information.

Theme 2: Content Level Issues

The Content Level theme received the fewest individual pieces of feedback but resulted in arguably the most substantial changes in the final form of the guidelines. Several of the reviewers felt that the guidelines contained writing that was too specialized/technical and that such writing would likely be difficult for novices to comprehend. The AR and Theory Selection sections were identified as being particularly technical and one reviewer emphasized the fact that an incomprehensible Theory Selection page would negatively impact the ability of readers to understand the forthcoming principles. A reviewer suggested using a researcher-created schematic to more effectively explain the cognitive science principles underlying the guidelines. Inconsistency with terms was cited as a factor that contributed to an unnecessarily complex content level. For example, in various places throughout the guidelines, spoken text was referred to as spoken text, auditory information, and narration. It was suggested that one of the terms be adopted and used consistently throughout the guidelines. Additionally, one reviewer requested that each principle be reworded to more directly relate the principle to how it will used in conjunction with AR to teach a procedural task.

Two additional major suggestions emerged from the data related to the Content Level theme. In an attempt to make the guidelines more comprehensible to novices, one reviewer suggested the use of a descriptive scenario to help readers understand the type of AR use that will be discussed in the guidelines. Also, a reviewer suggested using one AR use scenario be used consistently throughout the description of each adapted principle. For example, in the first draft of the guidelines, the multimedia principle was introduced by showing AR helping a user learn to shoot a basketball and the modality principle was introduced by demonstrating how AR could be used to help a person tie a tie. The reviewer expressed concern that switching scenarios created unnecessary confusion and ultimately directed attention away from the principle that was being introduced.

Theme 2: Resolution of Content Level Issues

The Content Level feedback resulted in eight major revisions in the guidelines. First, all of the text was edited with the goal of removing overly technical language. For example, the terms “effect size” and “control group” were removed from the guidelines. Also, the Theory Selection section was completely rewritten in an attempt to make it more accessible to novices. A figurative comparison, equating an overloaded working memory to a computer with too many applications open (a phenomenon that many novices likely have experienced) was used in an attempt to help novices understand how an overloaded working memory negatively impacts learning and performance. A schematic, created with the intent of providing a simple visual representation of the human information processing system, was included. The text was edited to ensure that consistent terminology was used throughout and all of the principles were intentionally reworded to better relate them to AR instruction.

Additionally, a descriptive scenario relating how AR enhanced instructions may help with the procedural task of changing a tire, was included on the Introduction page. This scenario was included in an attempt to provide a clear picture of the type of instruction and learning outcomes to be addressed by the guidelines. Finally, the examples that were presented with each guideline were completely redesigned to include the same scenario, learning how to tie a tie, for all examples. Such continuity will make it easier for readers to focus on the application of the guideline rather than the instructional scenario presented in the example. Some of the procedural learning AR scenarios were repurposed for use in a simple assessment that was included at the end of the guidelines.

Theme 3: Content Detail Issues

The majority of the expert feedback was related to the Content Detail theme. Comments associated with this theme advocated for adding, subtracting, or modifying information conveyed in various sections of the initial version of the guidelines. All of the reviewers made at least one comment related to the Content Detail with Dr. Feiner and Dr. Doolittle providing the most feedback related to this theme.

Dr. Feiner, an AR expert, stressed that the definition of AR be amended to include the explanation of spatial registration, which is a key component of the definition that informs his line of research. Spatial registration relates to the synchronization between the virtual content and the physical world. Dr. Fenier suggested including an explanation of spatial recognition as it relates to AR and illustrating how such registration would occur in the AR scenarios contained in the guidelines. Dr. Feiner also expressed concern that a key affordance of AR, co-location, was not emphasized enough in the guidelines. Co-location, the fact that virtual instructional material and the real domain objects can be co-located thus eliminating the need for users to switch their attention between the two, is considered by Dr. Feiner to be a key leverage point/attribute of AR and one that he suggested that this affordance be more overtly discussed in the guidelines.

The link between AR and procedural learning was discussed at length in the literature review. Numerous studies (De Cresenzio et al., 2011; Haritos & Macchiarella, 2004; Henderson & Feiner, 2009; Henderson & Feiner, 2011a, 2011b; Neumann & Majoros, 1998; Tang et al., 2001) suggest that AR can provide valuable cognitive and psychomotor support during procedural learning. AR instruction has been used to help with the performance and/or learning of a variety of procedural tasks including object assembly, machine maintenance, and surgical procedures. One reviewer noted that despite the strong link between AR and procedural learning, the connection was not expressed clearly and directly in the guidelines. It was suggested that this connection be made apparent via a text explanation included in the Introduction section and also in the section where the individual guidelines are presented for the first time.

Regarding the principles on which the guidelines are built, two reviewers felt that the multimedia principle should be explained in greater detail, with an emphasis on the optimal relationship between the words and pictures that are being used instructionally. Dr. Mayer suggested that the wording in the explanation of the personalization principle be altered to better match the wording cited in the research related to the principle. Finally, Dr. Sweller suggested that combining the redundancy and coherence principles may be helpful for novices. Sweller cited the fact that CLT (cognitive load theory, developed by Sweller) combines the two principles because they have the same effect for the same reasons. Sweller suggested that for the purposes of the guidelines, there was no need for two separate principles. Additionally, one reviewer suggested that for each example, the connection between the principle and the underlying theory be stated clearly and directly.

Theme 3: Resolution of Content Detail Issues

The feedback related to this theme resulted in seven major revisions to the initial draft of the guidelines. First, the definition of AR included in the guidelines was revised to include an explanation of spatial registration and a description of how such registration typically occurs. Also, the concept of co-location, a key affordance of AR, was explained in both the opening scenario and in the AR section of the guidelines.

The fact that the guidelines are best suited for assisting with the design of instruction for procedural learning was stated clearly in the Introduction section and at the point where the guidelines are introduced for the first time. This modification better aligns the guidelines with the body of related research on instructional applications of AR.

The description of the multimedia principle was altered to better explain the optimal interplay between words and pictures. The changes were made with novices in mind and the writing was not overly technical. As per the request of Dr. Mayer, the wording in the description of the personalization principle was altered to better align with the wording in the research related to the principle.

Based on the suggestion of Dr. Sweller and in an attempt to reduce the overall number of principles, thereby making the amount of content more manageable for novices, the redundancy principle was subsumed within the coherence principle. Given that the two principles have the same effect for the same reason (placing an unnecessary burden on working memory) the duplicated text associated with the redundancy principle was simply listed as another type of superfluous material that, when included in instruction, could potentially waste valuable cognitive resources. Finally, for each guideline, the underlying cognitive theory was detailed in a more direct manner so that readers would be able to better connect the theory to the guideline.

Revision Summary

Table 8 summarizes the recommendations related to each theme and details the action taken to address the recommendations. In total, four major alterations related to the Guideline Design theme were made, eight alterations related to the Content Level theme were made, and seven modifications related to the Content Detail theme were made.

Table 8: *Summary of guideline revisions*.

Appendix C.

Conclusion

With wearable displays (such as Google's Glass) poised to hit the market, AR use by the general public is likely to increase dramatically. Certainly, some of the ways in which AR is currently used will persevere. For example, when exploring a new area, users will perhaps activate their wearable displays and be directed to the most highly reviewed restaurants. A more compelling use of AR on wearable displays may take advantage of the fact that users will now have unfettered use of both hands and may build upon the identified benefits of using AR to assist with procedural learning/performance (De Cresenzio et al., 2011; Haritos & Macchiarella, 2004; Henderson & Feiner, 2009; Henderson & Feiner, 2011a, 2011b; Neumann & Majoros, 1998; Tang et al., 2001). Ideally, entrepreneurs and programmers will identify many ways in which AR can assist with procedural learning/task performance in everyday life.

As explained by Richey and Klein (2005), developmental research can produce context specific knowledge that serves a problem solving function. Certainly those with a background in instructional design will understand how to effectively use words and pictures when designing instruction. Those familiar with the multimedia research base will understand that when AR is used instructionally, it is another form of multimedia learning as defined by Mayer (2005). However, many entrepreneurs and developers who will work with AR in the near future may not be familiar with instructional design fundamentals and how they may apply to AR learning. The guidelines that were designed and developed as a result of this study attempt to connect sound multimedia design principles with the type of learning outcome for which AR seems most useful.

Ideally, this connection will help novice designers avoid simple design mistakes they may otherwise make and will lead to the development of more effectively designed AR applications for use on new wearable displays.

Ultimately, this set of guidelines aims to bridge a gap between theory and practice, a goal that is at the heart of the design and development research approach (Richey & Klein, 2007). The guidelines discussed here have emerged from a comprehensive literature review, an initial development stage guided by the Dick, Carey, and Carey model (1990), an expert review by six scholars from three related domains of expertise, and a thorough redesign process that has been detailed in this chapter. The guidelines are now ready to be shared with a larger audience and the next round of evaluation will ideally see them tested in practice by actual developers who use them to guide the development of an AR application.

Chapter 6: Observations

Purpose of the Study

The purpose of this study was to develop a set of guidelines that can be used to assist with the design of instruction delivered via AR. AR is a rapidly emerging technology and the pending release of wearable displays means that in the near future, mobile, hands-free AR usage is likely to increase dramatically. Developers and entrepreneurs, looking to capitalize on this new delivery approach will seek to design applications that take advantage of the new wearable displays. The developers will likely have a viable idea and will possess the requisite programming knowledge but may lack an understanding of sound instructional design procedures. The guidelines developed as a result of this study have been adapted from carefully selected instructional message design principles that are applicable to AR instruction. The guidelines are meant to serve as a tool to assist novice developers make informed instructional design decisions as they create new AR applications that will ideally build upon AR's documented benefits related to procedural learning and that will maximize the affordances of the hands-free delivery supported by wearable displays.

Study Synopsis

After the aforementioned design problem was identified, a comprehensive literature review was conducted. The review of literature related to AR led to two essential conclusions. First, a majority of the studies addressing AR involve procedural learning goals. AR was repeatedly shown (De Crescenzio et al., 2011; Henderson & Feiner, 2011a, 2011b; Tang et al., 2001; Yeo et al., 2011) to assist with both mental and physical aspects of such learning due to the fact that it allows for relevant information to be overlaid directly on a view of the real world. Second, despite the visual emphasis of AR, the majority of reported uses employed the use of both pictures and words, meaning that AR, used to facilitate learning, falls under the umbrella of Mayer's (2005a) definition of multimedia learning.

Given that connection between AR instruction and multimedia learning, the literature review was extended to the multimedia learning research base and the most applicable principles, including the multimedia, modality, contiguity, redundancy, coherence, and personalization principles, were selected for use in the guideline design process. With the principles in place, an initial set of guidelines, adapted from the selected principles and written specifically for novice instructional designers, was developed using the Dick, Carey, and Carey model (1990). Connections to underlying learning theory were made apparent and scenario specific examples were included. The initial draft of the guidelines and a researcher created rubric were provided to expert reviewers. Expert feedback was collected, systematically coded, and used to redesign the initial version of the guidelines.

Contribution of Study

This study established a set of guidelines that can be used inform instructional design decisions during the development of AR enhanced procedural learning applications. Information gathered from this study had both theoretical and practical implications.

Theoretical Implications

As indicated in the review of literature, AR (used instructionally) has been shown to employ both words and pictures to deliver content. As such, AR can reasonably be included as another form of multimedia instruction as defined by Mayer (2005a) who states that, "multimedia instruction involves presenting words and pictures that are intended to promote learning" (p.2). Given this connection, the instructional design principles that have emerged from the multimedia learning theory base can reasonably be applied to the design of AR instruction. In this study, applicable principles were selected and adapted to an AR delivery approach. This adaptation provides additional evidence of the quality and versatility of the principles that have emerged from the multimedia theory base. Essentially, all of the selected principles are evidence-based, straightforward techniques that can be used to successfully manage words and pictures when designing any form of multimedia instruction, including AR delivered instruction.

Design and development research attempts to contribute to the design and development knowledge base, according to Richey and Klein (2007). This study contributes to the knowledge base by presenting further support for the effectiveness of the expert validation technique. One contributing factor that led to a successful review process was the manner in which expert reviewers were selected. The overall topic of the guidelines was analyzed and three related domains (learning theory, instructional design, AR) were identified. Experts were selected from each of the three domains and this decision resulted in a more complete, triangulated amount of feedback.

A second factor that contributed to a successful review process was the open-ended design of the evaluation rubric. The rubric presented to reviewers was designed to allow for the experts to comment on all aspects of the proposed guidelines, not just those related to their identified domain of expertise. As described in Chapter 5, reviewers took advantage of this freedom as most provided feedback related to multiple domains. This behavior indicates that when applicable, evaluation materials should be designed in a way that allows for a maximum amount of freedom on the part of reviewers.

Finally, the three level coding process that is popular with qualitative researchers (Charmaz, 2006; Hahn, 2008) proved useful in identifying the essential revisions that needed to be made to the initial draft of the guidelines. Arranging feedback into categories and then themes helped to identify the essential improvements that needed to be made and helped ensure that all relevant expert feedback was operationalized in the revision process.

AR can reasonably be listed as another form of multimedia instruction, albeit one that that continues to evolve. Based on how easily they were adapted to the AR instructional design guidelines, there is further evidence to support the usefulness and versatility of the principles that are at the heart of the multimedia learning theory base. Regarding the design and development knowledge base, this study provides additional support for the expert validation technique as an integral part of the design and development process. Carefully analyzing the topic and selecting experts from all related domains and then giving those experts freedom to comment on all areas of the product helped maximize the value of the feedback that was given. Finally, there is the implication that the three level coding process popular with qualitative researchers can be useful to design and development researchers during the reiteration phase when they attempt to organize and operationalize the suggestions made by expert reviewers.

Practical Implications

As has been discussed in this document, the release of wearable displays holds the potential to dramatically alter the manner in which the general public experiences AR. Even the designers of the wearable displays are unsure of how the new displays will influence AR usage. According to the Google Glass Collective, an organization charged with facilitating the adoption and use of wearable displays, (Doerr, 2013)

> This is a platform so new, so unlike anything before, that we can't guess what the killer services will be. But, believe me, they're coming. The best ideas for the Glass platform will come from entrepreneurs — they always do (http://www.glasscollective.com/).

Currently, the general public typically experiences AR when they look through the display of their mobile device and are presented with cursory information related to their immediate location (facts about a building, restaurant reviews, etc.). While such use can be helpful, it seems to fall short of the true potential of AR, particularly when wearable displays allow for unfettered use of both hands. Given the research that outlines the benefits of combining AR and procedural learning, it is possible that such benefits will be applied to the public sector. Entrepreneurs who attempt to figure out how to best take advantage of wearable AR displays will almost certainly lack background knowledge related to instructional design. Hopefully, this set of guidelines, which clearly and directly addresses instructional design for AR, will provide some general oversight to the design process and will help the content developers successfully manage the use of words and pictures in their applications.

Next Steps

As a result of the expert review process, the set of guidelines have been carefully recalibrated to make them more accessible for those who lack a background in learning theory and instructional design. Ideally, the guidelines will now be tested by application developers to determine how well they achieve their intended goal of helping novices make informed instructional message design decisions. To facilitate their use, the guidelines will now be made available to users in both document and website form.

References

Alessi, S. M. (1988). Learning interactive videodisc development: A case study. *Journal of Instructional Development, 11*(2), 2–7.

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *The Quarterly Journal of Experimental Psychology*, *24*(2), 225-235.

Autonomy. (2006). *Technical brief of the intelligent data operating layer*. San Francisco, CA: Author.

Ayres, P.L.. & Sweller, J.. (2005). The split-attention principle in multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 135-146). New York, NY: Cambridge University Press.

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, *6*(4), 355.

Baddeley, A. D. (1999). *Essentials of human memory*. Psychology Press/Taylor & Francis (UK).

Baird, K. M. (2000). *Evaluating the effectiveness of augmented reality and wearable computing for a manufacturing assembly task*. Blacksburg, VA: University Libraries, Virginia Polytechnic Institute and State University.

Cakmakci, O., Bérard, F., & Coutaz, J. (2003). An augmented reality based learning assistant for electric bass guitar. In *Proc. of the 10th International Conference on Human-Computer Interaction.* Crete, Greece.

Calhoun, G. L., & McMillan, G. R. (1998, March). Hands-free input devices for wearable computers. In *Human Interaction with Complex Systems, 1998. Proceedings., Fourth Annual Symposium on* (pp. 118-123). IEEE.

Carmingniani, J., & Furht, B. (2011). Augmented reality: An overview. In B. Furht (Ed.), *Handbook of augmented reality* (pp. 3-45). New York, NY: Springer.

Carliner, S. (2008). How designers make decisions: A descriptive model of instructional design for informal learning in museums. *Performance Improvement Quarterly*, *11*(2), 72-92.

Caudell, T.P., & Mizell, D.W. (1992). Augmented reality: An application of heads-up display technology to manual manufacturing processes. *Proceedings of the Hawaii International Conference on Systems Sciences.* doi 10.1109/HICSS.1992.183317

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London: Sage Publications.

Chen, N.-S., Teng, D. C.-E., Lee, C.-H., & Kinshuk. (2011). Augmenting paper-based reading activity with direct access to digital materials and scaffolded questioning. *Computers and Education*, *57*(2), 1705-1715.

Clark, R. E., & Feldon, D. F. (2005). Five common but questionable principles of multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 97-115). New York, NY: Cambridge University Press.

Clark, R. C., & Mayer, R. E. (2011). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. San Francisco, CA: Pfeiffer.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. *Journal of Educational Psychology*, *94*(2), 428.

De Crescenzio, F., Fantini, M., Persiani, F., Stefano, L. D., Azzari, P., & Salti, S. (2011). Augmented reality for aircraft maintenance training and operations Support. *IEEE Computer Graphics & Applications*, *31*(1), 96-101.

Dewey, J. (1938). Experience and education. New York, NY: Macmillan.

Dick, W., & Carey, L. (1990). The systematic design of instruction. Glenview, Ill: Scott, Foresman/Little, Brown Higher Education.

Doerr, J. (2013). *Google glass collective.* Retrieved from http://www.glasscollective.com/

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning. *Journal of Science Education & Technology*, *18*(1), 7-22.

Fletcher, J. D., & Tobias, S. (2005). The multimedia principle. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 117-133). New York, NY: Cambridge University Press.

Gardeya, M. (2010). Berlin wall 3d. Hoppala mobile augmented reality projects. Retrieved from http://www.hoppala-agency.com/article/berlin-wall-3d/

Gandy, M., MacIntyre, B., Presti, P., Dow, S., Bolter, J., Yarbrough, B., & O'Rear, N. (2005). AR Karaoke: Acting in your favorite scenes. In *Mixed and Augmented Reality. Proceedings. Fourth IEEE and ACM International Symposium on* (pp. 114-117). IEEE.

Gustafson, K. L., Branch, R. M., ERIC Clearinghouse on Information & Technology., & United States. (2002). Survey of instructional development models. Syracuse, N.Y: ERIC Clearinghouse on Information & Technology.

Graesser, A. C., Hakel, M., & Halpern, D. F. (2007). Life-long learning at work and at home. *APS Observer*, *20*, 17-21.

Halpern, D.F., Graesser, A., & Hakel, M. (2007). *25 Learning principles to guide pedagogy and the design of learning environments.* Washington, DC: Association of Psychological Science Taskforce on Lifelong Learning at Work and at Home.

Hahn, C. (2008). *Doing qualitative research using your computer: A practical guide*. London: Sage Publications Limited.

Hari-Narayanan, N., & Hegarty, M. (2002). Multimedia design for communication of dynamic information. *International Journal of Human-Computer Studies*, *57*(4), 279-315.

Haritos, T. & Macchiarella, N. D. (2005). A mobile application of augmented reality for aerospace maintenance training. In *The 24th Digital Avionics Systems Conference, DASC 2005*, pp. 5.B.3–5.1-9.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*(3), 414.

Henderson, S. J., & Feiner, S. (2009). Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In *Mixed and Augmented Reality, 2009. 8th IEEE International Symposium on* (pp. 135-144). IEEE.

Henderson, S. J., & Feiner, S. K. (2011a). Augmented reality in the psychomotor phase of a procedural task. In *Mixed and Augmented Reality, 2011. 10th IEEE International Symposium on* (pp. 191-200). IEEE.

Henderson, S., & Feiner, S. (2011b). Exploring the benefits of augmented reality documentation for maintenance and repair. *IEEE Transactions on Visualization and Computer Graphics 17*(10), 1355-1368.

Hughes, O., Fuchs, P. & Nannipieri, O. (2011). New augmented reality taxonomy: technologies and features of augmented environments. In B. Furht (Ed.), *Handbook of augmented reality* (pp. 47-63). New York, NY: Springer.

Jamet, E., & Le Bohec, O. (2007). The effect of redundant text in multimedia instruction. *Contemporary Educational Psychology*, *32*(4), 588-598.

Kartel, G. (2010). Does language matter in multimedia learning? Personalization principle revised. *Journal of Educational Psychology,* 102, 615-624.

Kaufmann, H., & Csisinko, M. (2011). Wireless displays in educational augmented reality applications. In B. Furht (Ed.), *Handbook of augmented reality* (pp. 157-175). New York, NY: Springer.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23-31.

Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(3), 567-581.

Klopfer, E. (2008). *Augmented learning: Research and design of mobile educational games* . Cambridge, MA: MIT Press.

Levie, W.H., & Dickie, K.E. (1971). The analysis and application of media. In R.M.W. Travers (Ed.), *Second handbook of research on teaching*. Chicago, IL: Rand McNally.

Liao, H., Inomata, T., Sakuma, I., & Dohi, T. (2010). 3-D augmented reality for MRI-guided surgery using integral videography autostereoscopic image overlay. *Biomedical Engineering, IEEE Transactions on*, *57*(6), 1476-1486.

Liu, T.-Y. (2009). Context-aware ubiquitous learning environment for language listening and speaking. *Journal of Computer Assisted Learning*, *25*, 515-527.

Liu, T. Y., Tan, T. H., & Chu, Y. L. (2009). Outdoor natural science learning with an RFID-supported immersive ubiquitous learning environment. *Educational Technology & Society*, *12*(4), 161-175.

Lockee, B. (Interviewer) & Mayer, R. (Interviewee). (2012). *Interview with Dr. Richard Mayer* [Online Video]. Retrieved from AECT HistoryMakers Project Web Site Project Web site: http://www.vtidt.com/aect/

Low, R., & Sweller, J. (2005). The modality principle in multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 147-158). New York, NY: Cambridge University Press.

Mayer, R. (2005a). Introduction to multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 1-18). New York, NY: Cambridge University Press.

Mayer, R. (2005b). Cognitive theory of multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 31-49). New York, NY: Cambridge University Press.

Mayer, R. (2005c). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity and temporal contiguity principles. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 183-200). New York, NY: Cambridge University Press.

Mayer, R. (2005d). Principles of multimedia learning based on social cues: Personalization, voice, and image principles. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 201-214). New York, NY: Cambridge University Press.

Mayer, R. E. (2008). Applying the science of learning: evidence-based principles for the design of multimedia instruction. *American Psychologist*, *63*(8), 760.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, *83*(4), 484.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, *84*(4), 444.

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology*, *88*(1), 64.

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology*, *95*(4), 806.

Mayer, R. E., Fennell, S., Farmer, L., & Campbell, J. (2004). A Personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology*, *96*(2), 389.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*(4), 715.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, *93*(1), 187.

Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*(1), 107-119.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, *86*(3), 389.

Milgram, P., & Kishino, F. (1994). Taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems,* 1321-1329.

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*(2), 358.

Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology*, *92*(1), 117.

Moreno, R., & Mayer, R. E. (2004). Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology*, *96*(1), 165.

Moussa, G., Radwan, E., & Hussain, K. (2012). Augmented reality vehicle system: Left-turn maneuver study. *Transportation Research Part C: Emerging Technologies*, *21*(1), 1-16.

Nadolski, R., Kirschner, P., van Merrienboer, J., & Hummel, H. (2001). A model for optimizing step size of learning tasks in competency-based multimedia practicals. *Educational Technology Research & Development, 49*(3), 87–103.

Neumann, U., & Majoros, A. (1998). Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. *Proceedings of IEEE, Virtual Reality Annual International Symposium*, pp. 4-11.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. Mahwah, N.J: Lawerence Erlbaum Associates.

Park, H., & Park, J.-I. (2010). Invisible marker-based augmented reality. *International Journal of Human-Computer Interaction*, *26*, 829-848.

Pew Research Center's Internet and American Life Project. (2012). *Summer tracking survey*. (1st ed.). Washington, DC: Author.

Plass, J. L., & Salisbury, M. W. (2002). A living-systems design model for web-based knowledge management systems. *Educational Technology Research and Development*, *50*(1), 35-56.

Pollicino, J. (2012). Oakley unveils airwave ski goggles equipped with Recon's MOD Live heads up display. Retrieved from http://www.engadget.com/2012/10/27/oakley-unveils-airwave-ski-goggles-equipped-with-recons-mod-live

Quarles, J., Lampotang, S., Fischler, I., Fishwick, P., & Lok, B. (2008, March). Tangible user interfaces compensate for low spatial cognition. In *3D User Interfaces, 2008. 3DUI 2008. IEEE Symposium on* (pp. 11-18). IEEE.

Reif, R., Günthner, W. A., Schwerdtfeger, B., & Klinker, G. (2010). Evaluation of an augmented reality supported picking system under practical conditions. *Computer Graphics Forum*, *29*(1), 2-12.

Richey, R., & Klein, J. (2005). Developmental research methods: creating knowledge from instructional design and development practice. *Journal of Computing in Higher Education*, *16*(2), 23-38.

Richey, R., & Klein, J. (2007). *Design and development research: methods, strategies, and issues*. Mahwah, NJ: Erlbaum.

Richey, R., Klein, J., & Nelson, W.A. (2004). Developmental research: Studies of instructional design and development. In D.H. Jonassen (Ed.), *Handbook of research for educational communications and technology: A project of the association for educational communications and technology* (2nd ed., pp. 1099–1130). Hillsdale, NJ: Erlbaum.

Robertson, C. M., MacIntyre, B., & Walker, B. N. (2008, September). An evaluation of graphical context when the graphics are outside of the task area. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality* (pp. 73-76). IEEE Computer Society.

Rusch, M. L., Schall Jr, M. C., Gavin, P., Lee, J. D., Dawson, J. D., Vecera, S., & Rizzo, M. (2013). Directing driver attention with augmented reality cues. *Transportation Research Part F: Traffic Psychology and Behaviour*, *16*, 127-137.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system-and self-paced instruction in multimedia learning. *Learning and Instruction*, *20*(2), 100-110.

Shalvey, K. (2012). Google glasses, wearable computer sales seen swelling. Retrieved from http://news.investors.com/technology/103112-631578-juniper-research-says-wearable-computer-sales-to-double.htm

State, A., Livingston, M., Garrett, W., Hirota, G., Whitton, M., Pisano, E., & Fuchs, H. (1996). Technologies for augmented reality systems: Realizing ultrasound-guided needle biopsies. *Medical Image Analysis, 6*(3), 313-320.

Spector, J. M., Muraida, D. J., & Marlino, M. R. (1992). Cognitively based models of courseware development. *Educational Technology Research and Development*, *40*(2), 45-54.

Squire, K. (2010). From information to experience: Place-based augmented reality games as a model for learning in a globally networked society. *Teachers College Record*, *112*(10), 2565-2602.

Squire, K. & Jan, M. (2007). Mad city mystery: developing scientific argumentation skills with a Place-based augmented reality games on handheld computers. *Journal of Science Education & Technology*, *16*(1), 5-29.

Squire, K. & Jenkins, H. (2011). *Video games and learning: Teaching and participatory culture in the digital age*. New York, NY: Teachers College Press.

Squire, K., & Klopfer, E. (2007). Augmented reality simulations on handheld computers. *Journal of the Learning Sciences*, *16*, 371-413.

Sutherland, I. E. (1968). A head-mounted three-dimensional display. *1968 Fall Joint Computer Conference, AFPIS Conference Proceedings*. 33:757-764.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction*, *8*(4), 351-362.

Sweller, J. (2005a). Implications of cognitive load theory for multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 19-30). New York, NY: Cambridge University Press.

Sweller, J. (2005b). The redundancy principle in multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 159-168). New York, NY: Cambridge University Press.

Tang, A., Owen, C., Biocca, F., & Mou, W. (2003). Comparative effectiveness of augmented reality in object assembly. 73-80. doi: 10.1145/642611.642626

The horizon report. (2012). New Media Consortium NMC. Retrieved from http://www.nmc.org/publications/horizon-report-2012-higher-ed-edition

van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker, R. M. Branch, K. Gustafson, N. Nieveen,& T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 1-14). Boston: Kluwer Academic.

Vlahakis, V., Ioannidis, M., Karigiannis, J., Tsotros, M., Gounaris, M., Stricker, D., & Gleue, T. (2002). Archeoguide: an augmented reality guide for archaeological sites. *Computer Graphics and Applications, IEEE*, *22*(5), 52-60.

Yeo, C. T., Ungi, T., U-Thainual, P., Lasso, A., McGraw, R. C., & Fichtinger, G. (2011). The effect of augmented reality training on percutaneous needle placement in spinal facet joint injections. *IEEE Transactions on Biomedical Engineering*, *58*, 2031-2037.

APPENDIX A

TABLE 1

Table 1: *Summary of publications in which AR is used as an instructional delivery method.*

Authors	**Title**	**Summary**	**Type of Knowledge Taught**	**Learner Traits**	**Learning Environment**	**Instructional Modality**
Cakmakci, O., Berard, F., & Coutaz, J. (2003).	An Augmented Reality Based Learning Assistant for Electric Bass Guitar	An AR system was used to help users find and play the correct notes on a bass guitar.	procedural task	novice/non-specialized	controlled lab setting	visual only
De Cresenzio, F., Massimilliano, F., Persiani, F., Di Stefano, L., Azzari, P., & Salti, S. (2011).	Augmented Reality for Aircraft Maintenance Training and Operations Support	AR system was developed and tested to evaluate its effect on the performance of 7 successive airplane maintenance tasks.	procedural task	expert-specialized	authentic real world setting	verbal and visual
Dunleavy, M., Dede, C. & Mitchell, R. (2008).	Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning	Middle and high school students participated in an AR simulation designed to teach math, language arts, and scientific literacy skills.	verbal information	novice/non-specialized	authentic real world setting	verbal and visual
Gandy, M., MacIntyre, B., Presti, P., Dow, S., Bolter, J., Yarbrough, B., & O'Rear, N. (2005)	AR Karaoke:Acting in Your Favorite Scenes	AR was used to facilitate an experience where users performed their favorite dramatic scenes with virtual actors.	procedural task	novice/non-specialized	authentic real world setting	verbal and visual

Authors	**Title**	**Summary**	**Type of Knowledge Taught**	**Learner Traits**	**Learning Environment**	**Instructional Modality**
Haritos, T. & Macchiarella, N. (2004).	Effectiveness of Video-Based Augmented Reality as a Learning Paradigm for Aerospace Maintenance Training	An AR based training system was tested to determine its effectiveness on the assembly and maintenance of aerospace parts.	procedural task	novice/non-specialized	authentic real world setting	verbal and visual
Henderson, S., & Feiner, S. (2009).	Evaluating the Benefits of Augmented Reality for Task Localization in Maintenance of an Armored Personnel Carrier Turret	An AR system was used to overlay information to assist with machine maintenance.	procedural task	novice/non-specialized	authentic real world setting	verbal and visual
Henderson, S., & Feiner, S. (2011)	Augmented Reality in the Psychomotor Phase of a Procedural Task	With the aid of an AR interface, users were able to complete the psychomotor aspects of an assembly task significantly faster and with greater accuracy than when using a traditional set of instructions.	procedural task	expert-specialized	controlled lab setting	verbal and visual

Authors	**Title**	**Summary**	**Type of Knowledge Taught**	**Learner Traits**	**Learning Environment**	**Instructional Modality**
Henderson, S., & Feiner, S. (2011)	Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair	Military mechanics, guided by instructions delivered via AR performed an 18 step maintenance procedure under field conditions. AR allowed for quicker task location and in some instances, resulted in less head movement.	procedural task	procedural task	authentic real world setting	verbal and visual
Liao, H., Inomata, T., & Dohi, T. (2010).	3-D Augmented Reality for MRI-Guided Surgery Using Integral Videography Autostereoscopic Image Overlay	To assist with surgical procedures, specific 3D images are overlaid onto a the view of a body.	procedural task	expert-specialized	controlled lab setting	visual only
Neumann, U., & Majoros, A. (1998).	Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance	A sequenced maintenance task was overlaid on the fuselage of and aircraft.	procedural task	novice/non-specialized	authentic real world setting	verbal and visual
Quarles, J., Lampotang, S., Fischler, I., Fishwich, P., & Lok, B. (2008).	Tangible User Interfaces Compensate for Low Spatial Cognition	An AR system was found to significantly compensate for low user spatial cognition in the domain of anesthesia machine training.	procedural task	expert-specialized	controlled lab setting	verbal and visual

Authors	**Title**	**Summary**	**Type of Knowledge Taught**	**Learner Traits**	**Learning Environment**	**Instructional Modality**
Reif, R., Gunther, W., Schwerdtfeger, B., & Klinker, G. (2010)	Evaluation of an Augmented Reality Supported Picking System Under Practical Conditions	An AR system was designed and tested to measure performance of order picking in a real storage environment.	procedural task	novice/non-specialized	authentic real world setting	verbal and visual
Robertson, C., MacIntyre, B., & Walker, B. (2008).	An Evaluation of Graphical Context When the Graphics are Outside of the Task Area	Registered AR helped participants efficiently complete a Lego block placement task.	procedural task	novice/non-specialized	controlled lab setting	visual only
Rusch, M., Schall, M., Gavin, P., Lee, J., Dawson, J., Vecera, S., & Rizzo, M. (2012).	Directing Driver Attention with Augmented Reality Cues	An AR system was used to provide cues to drivers and to direct their attention to roadside hazards and other cars.	procedural task	novice/non-specialized	authentic real world setting	verbal and visual
State, A. Livingston, W., Garrett, W., Hirota, G., Whitton, M., Pisano, E., & Fuchs, H. (1996).	Technologies for Augmented Reality Systems: Realizing Ultrasound-guided Needle Biopsies	An AR system helped a physician successfully guide a needle into an artificial tumor within a training phantom of a human breast.	procedural task	expert-specialized	controlled lab setting	visual only
Squire, K. (2010).	From Information to Experience: Place-Based Augmented Reality Games as a Model for Learning in a Globally Networked Society	Students experienced an AR game, Sick at South Beach, with the objective of developing scientific argumentation skills.	attitude, problem solving	novice/non-specialized	authentic real world setting	verbal and visual

Authors	**Title**	**Summary**	**Type of Knowledge Taught**	**Learner Traits**	**Learning Environment**	**Instructional Modality**
Squire, K., & Jan, M. (2007).	Mad City Mystery: Developing Scientific Argumentation Skills with a Place-based AR Game on Handheld Computers	An AR simulation called Mad City Mystery was used with students in an attempt to have them "think like scientists" and perform adult scientific discourses.	attitude, problem solving	novice/non-specialized	authentic real world setting	verbal and visual
Squire, K., & Klopfer, E. (2007).	Augmented Reality Simulations on Handheld Computers	An AR simulation called Environmental Detectives was used to to help students understand the socially situated nature of scientific practice.	attitude, problem solving, rule use	novice/non-specialized	authentic real world setting	verbal and visual
Tang, A., Owen, C., Biocca, F., & Mou, W. (2003)	Comparative Effectiveness of Augmented Reality in Object Assembly	AR instructions helped decrease object assembly time and error rate.	procedural task	novice/non-specialized	controlled lab setting	visual only
Yeo, C., Ungi, T., Lasso, A., McGraw, R., & Fichtinger, G. (2011).	The Effect of Augmented Reality Training on Percutaneous Needle Placement in Spinal Facet Joint Injections	An AR system was used to measure the effectiveness of image overlay and laser guidance systems in helping medical trainees learn the correct placement of a needle for a joint injection.	procedural task	expert-specialized	controlled lab setting	visual only

APPENDIX B

TABLE 4

Table 4: *Summary of research based evidence*.

Principle	**Research Summary**	**Notable Examples**
Multimedia	11 of 11 studies indicated that learners who received words with corresponding graphics performed better on transfer tests than those who did not. The 11 experiments had a median effect size of 1.39 (Mayer, 2008).	*Mayer (1996) Low-knowledge learners who were given illustrated descriptions of mechanical systems did better on problem solving and transfer tests than those that did not * Mayer and Gallini (1990) found that illustrations labeling parts of a scientific system and illustrations actions in a scientific were helpful in learner recall and creative problem solving, compared to text only treatments. * Mayer, Bove, Bryman, Mars, and Tapangco (1996) found that a multimedia summary was more effective in teaching a scientific process than a verbal summary. *Mayer and Anderson (1991) found that instruction containing words and pictures, compared to words-only and pictures-only treatments, led to superior performance on recall and problem solving tests. *Kalyuga, Ayres, Chandler, and Sweller (2003) found that fully guided instructional material placed an unnecessary burden on the working memory of expert learners.

Principle	Research Summary	Notable Examples
Contiguity		
<u>Spatial</u>	5 of 5 tests indicated that learners who received integrated presentations performed better on transfer tests than those who received separated presentations. The median effect size was 1.12 (Mayer, 2008).	* Moreno and Mayer (1999) showed that learning is impaired when on-screen text is separated from visual materials. * Chandler and Sweller (1991) demonstrated that learners exposed to integrated instructions, compared to those exposed to disconnected instructions, performed better in both written test and practical skill demonstrations.
<u>Temporal</u>	8 of 8 tests involving computer-based lessons showed that learners performed better when animation and narration was presented concurrently rather than successively. The median effect size was 1.31(Mayer, 2008).	*Mayer and Anderson (1992) showed large increases in problem-solving transfer when narration and animation was presented successively. *Mayer and Sims (1994) found that narration and animation, related to scientific topics, was more effective when the two sources were presented concurrently rather than successively.

Principle	**Research Summary**	**Notable Examples**
Modality	17 of 17 tests involving computer-based multimedia lessons indicated that learners scored higher on transfer tests when the words were presented in spoken form rather that printed form. The median effect size for the experiments was 1.02 (Mayer, 2008).	* Allport, Antonis, and Reynolds (1972) showed that a disruption in a learners audio channel would effect the performance of the recall of orally presented words but not visually presented words. *Mayer and Moreno (1998) showed that students presented with concurrent narration describing the major steps in a process performed better in recall, retention, and problem solving tests compared to students presented with on screen text information. * Craig, Gholson, and Driscoll (2002) measured the effects of printed text compared to spoken narration delivered by a pedagogical agent and found that students performed better on retention, transfer, and matching problems compared to students receiving the spoken-narration-only condition. *Mayer, Dow, and Mayer (2003) found that students learning about scientific topics performed better on a problem-solving transfer test when presented with explanations in narration form. *Schmidt-Weiglans, Konery, and Glowalla (2010) tracked learner eye movement while viewing a multimedia presentation. Researchers found that when text was presented in written form, learners consistently started reading before alternating between the text and a visual and spent more time inspecting the text than the visual.

Principle	**Research Summary**	**Notable Examples**
Redundancy	In 4 of 4 studies where the presentation of animation, narration, and redundant on-screen text was compared to the presentation of animation and narration, participants in the group without the redundant information performed better on a problem-solving transfer test with a median effect size greater than 1 (Clark and Mayer, 2011).	* Mayer, Heiser, and Lonn (2001) conducted an experiment in which the control group had to listen to a narration describing what was happening in an animation. One treatment group was exposed to redundant summarized printed text and another group was exposed to a duplication of the text. The control group outperformed the treatment groups on a transfer test. *Kayluga, Chandler, and Sweller (2004) used only textual information (with no accompanying visuals or animations) to demonstrate the redundancy effect. One group of participants was given only spoken text while another group was given simultaneous redundant printed and spoken text. The auditory-only group performed better on a transfer test. * Jamed and Le Bohec (2007) compared three groups of students who viewed a multimedia presentation. One group was presented with only spoken text, one which was given spoken text and sequentially presented redundant sentences and a third group was given a static presentation of redundant sentences. The duplicated information led to a poorer participant performance on retention and transfer tests.

Principle	**Research Summary**	**Notable Examples**
Coherence	In 13 out of 14 studies involving either computer or paper based multimedia presentations, students performed better on problem solving tests when they received a concise lesson rather than an expanded lesson, with an associated median effect size of .97 (Mayer, 2008).	* Moreno and Mayer (2000) found that extraneous music and/or sound effects interfered with student learning in a computer-based multimedia lesson. * Mayer, Heiser, and Lonn (2001) found that related but nonessential video clips inserted into a multimedia presentation interfered were detrimental to student performance on transfer problems. * Harp and Mayer (1997) found that additional color visuals, inserted into a paper-based multimedia presentation had a negative effect on student performance on retention and transfer tests. * Mayer, Heiser and Lonn (2001) found that additional, unnecessary narrated text was detrimental to student performance on transfer tests. * Mayer, Bove, Bryman, Mars, and Tapangco (1996) found that unnecessary textual information included in a multimedia presentation reduced students ability to abstract the core verbal explanation of the text and connect it with the visual explanation.

Principle	**Research Summary**	**Notable Examples**
Personalization	In 11 of 11 experiments involving computer-based lessons on lungs, lightning, botany, and industrial engineering, learners who received words in a conversational style performed better on transfer tests than did learners who received words in a formal style. The median effect size for the experiments was 1.11(Mayer, 2008).	* Mayer, Fennell, Farmer, and Campbell (2004) found that changing words form a formal style to a personalized style caused and improvement on transfer tests for participants exposed to the personalized style. * Moreno and Mayer (2004) presented students with both personalized and nonpersonalized speech. Students who received personalized messages performed better on retention and problem-solving transfer tests with an associated effect size of 1.64. *Kartel (2010) found that presenting a personalized text style along with animations and illustrations led to an increased performance on a problem-solving test.

APPENDIX C

TABLE 8

Table 8: *Summary of guideline revisions*.

Theme	Recommendation	Resolution
1. Guideline Design:	1. Select visuals that better align with the text	1. Unnecessary visuals were replaced with visuals that better align with the text
	2. Remove ineffective cueing strategies	2. Ineffective cueing strategies were removed
	3. Reconceptualize the use of non-examples	3. Non-examples were removed and some were used in a mini-assessment at the end of the guidelines
	4. Provide a short preview of the content to be shared for each guideline	4. When the guidelines are introduced for the first time, readers are presented with a preview of the content to be included for each
2. Content Level:	1. Edit text to remove overly technical language	1. All unnecessary technical language was removed from the guidelines.
	2. Rewrite specific sections of the guidelines to make the text descriptions more understandable for novice readers	2. All of the text in the guidelines was edited to make sure that the content was presented on a level that would be understandable for users.
	3. Include a schematic to help readers understand the how cognitive information processing occurs.	3. A researcher created schematic was included to help readers better understand cognitive information processing
	4. Use consistent terminology throughout the guidelines	4. The text was edited to ensure that consistent terminology was used throughout
	5. Reword guideline descriptions to better relate them to AR	5. Descriptions were edited to emphasize how they relate to AR
	6. Include a descriptive scenario at the beginning of the guidelines	6. A descriptive scenario related to changing a flat tire was included at the beginning of the guidelines
	7. To avoid confusing readers, use the	7. Examples were edited so that the same

Theme	Recommendation	Resolution
	same scenario when presenting examples for each guideline 8. Reword guideline descriptions to better relate them to AR	scenario was used throughout 8. The text was edited to ensure that consistent terminology was used throughout
3. Content Detail:	1. Include the concept of spatial registration when defining AR 2. Place a greater emphasis on the affordance of co-location as related to AR instruction 3. Explicitly state the fact that the guidelines are intended to assist with the design of instruction that attempts to teach procedural learning. 4. Better explain the multimedia principle; specifically, the optimal relationship between words and pictures 5. Edit the wording in the description of the personalization principle 6. Combine the redundancy and coherence principles 7. When describing each guideline, more clearly explain the underlying theory	1. The stated definition of AR was altered to include spatial registration 2. The benefit of co-location was emphasized in the AR section of the product 3. The applicability of the guidelines for the design of instruction with procedural learning goals was emphasized 4. The explanation of the multimedia principle was edited to better explain the ideal interplay between words and pictures 5. The description of the personalization principle was edited to better match the wording in the related research 6. The redundancy principle was subsumed under the coherence principle 7. For each guideline, the link to the underlying theory was clearly stated

Appendix D

Expert Review Request

Chris W. Wasko, Ph.D. Candidate
50 South Main Street Suite 326
Blacksburg, VA 24061
3369700435
cwwask0@vt.edu

March 5, 2013

Dear Reviewer,

My name is Chris Wasko and I am a Doctoral Candidate in the Instructional Design and Technology program at Virginia Tech. I am mailing to request that you act as an expert reviewer for a set of guidelines that I am designing and developing in partial fulfillment of the requirements of my doctoral program. My goal is to create a set of theoretically grounded and evidence based instructional design guidelines that can be used by designers who elect to use augmented reality (AR) to deliver instruction.

You have been identified as a potential reviewer based on your scholarly reputation in the (Instructional Design/Learning Sciences/Human Computer Interaction) field. Your history of publications including (INSERT PUBLICATIONS HERE) are indicative of your expertise. Such expertise would be useful in ensuring that the developed guidelines are sound and will meet their specified objective.

Should you elect to act as an expert reviewer, you will be provided with a rubric that can be used to evaluate the proposed guidelines. Also, you will be provided with access to a website on which the guidelines will be made available. Both the rubric and the guidelines can be accessed electronically and the completed rubric can be emailed to me. I estimate that your full participation will take between 30-60 minutes.

If you are willing to help review and validate my proposed guidelines, please reply to this email by (INSERT DATE HERE). Once I receive your reply, I will mail an electronic copy of the rubric and a link that can be used to access the guidelines.

I appreciate your consideration of this request. Based on your reputation as a scholar, I believe you have the potential to provide valuable insight to my research.

Sincerely,
Chris W. Wasko

Appendix E

Evaluation Rubric

Name:

The following questions are specific to each of the six principles:

1. Multimedia Principle -

Based on your review of the information contained on the guideline website, is the multimedia principle adequately explained?

Has an adequate amount of research explaining the multimedia principle been provided?

Does the non-example successfully portray a violation of the multimedia principle?

Is the example of an enactment of the multimedia principle reasonable based on the attributes and affordances associated with AR?

Does the example provided reflect how AR may be used instructionally?

2. *Contiguity Principle -*

Based on your review of the information contained on the guideline website, is the contiguity principle adequately explained?

Has an adequate amount of research explaining the contiguity principle been provided?

Does the non-example successfully portray a violation of the contiguity principle?

Is the example of an enactment of the contiguity principle reasonable based on the attributes and affordances associated with AR?

Does the example provided reflect how AR may be used instructionally?

3. *Modality Principle -*

Based on your review of the information contained on the guideline website, is the modality principle adequately explained?

Has an adequate amount of research explaining the modality principle been provided?

Does the non-example successfully portray a violation of the modality principle?

Is the example of an enactment of the modality principle reasonable based on the attributes and affordances associated with AR?

Does the example provided reflect how AR may be used instructionally?

4. *Redundancy Principle -*

Based on your review of the information contained on the guideline website, is the redundancy principle adequately explained?

Has an adequate amount of research explaining the redundancy principle been provided?

Does the non-example successfully portray a violation of the redundancy principle?

Is the example of an enactment of the redundancy principle reasonable based on the attributes and affordances associated with AR?

Does the example provided reflect how AR may be used instructionally?

5. *Coherence Principle -*

Based on your review of the information contained on the guideline website, is the coherence principle adequately explained?

Has an adequate amount of research explaining the coherence principle been provided?

Does the non-example successfully portray a violation of the coherence principle?

Is the example of an enactment of the coherence principle reasonable based on the attributes and affordances associated with AR?

Does the example provided reflect how AR may be used instructionally?

6. *Personalization Principle -*

Based on your review of the information contained on the guideline website, is the personalization principle adequately explained?

Has an adequate amount of research explaining the personalization principle been provided?

Does the non-example successfully portray a violation of the personalization principle?

Is the example of an enactment of the personalization principle reasonable based on the attributes and affordances associated with AR?

Does the example provided reflect how AR may be used instructionally?

The following questions relate to the overall development of the guidelines:

Are the theoretical components appropriately selected?

Is there evidence of a theory of human learning behind each of the principles and is this apparent in the guidelines?

Is the presentation of the guidelines (website) in need of any modifications?

What changes need to be made?

Does the overall product hold the potential to be of use to novice and expert designers who elect to deliver instruction using AR?

Appendix F

Permission Request

From: Steven Feiner
Subject: **Re: image request**
Date: April 30, 2013 11:33:33 AM EDT
To: Chris Wasko

Hi Chris. You are welcome to include the image in your paper. Please credit it to Steven Henderson and me by citing our ISMAR 2011 paper.

Steve

Steven Feiner
Professor
Department of Computer Science
Columbia University
500 W 120th St., 450 CS Building
New York, NY 10027
feiner@cs.columbia.edu

Quoting Chris Wasko <cwwask0@vt.edu>:

> Dr. Feiner,
>
> I am writing to request your permission to use an image from your 2011 study about AR in the psychomotor phase of a procedural task. I would like to include the image in my dissertation because I believe it clearly and directly illustrates all of the essential attributes of AR- 1. spatial registration 2. digital information overlaid on a direct view of the physical world 3. assistance related to both cognitive and psychomotor aspects of a procedural task 4. real-time.
>
> Thank you for considering this request.
>
> Chris Wasko

Appendix G

IRB APPROVAL

Office of Research Compliance
Institutational Review Board
2000 Kraft Drive, Suite 2000 (0497)
Blacksburg, VA 24060
540/231-4606 Fax 540/231-0959
email irb@vt.edu
website http://www.irb.vt.edu

MEMORANDUM

DATE: February 1, 2013

TO: Christopher Warren Wasko, Barbara B Lockee

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires May 31, 2014)

PROTOCOL TITLE: Design and Development of Guidelines for Instruction Delivered via AR

IRB NUMBER: 13-101

Effective February 1, 2013, the Virginia Tech Institution Review Board (IRB) Chair, David M Moore, approved the New Application request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

http://www.irb.vt.edu/pages/responsibilities.htm

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As:	**Exempt, under 45 CFR 46.110 category(ies) 2**
Protocol Approval Date:	**February 1, 2013**
Protocol Expiration Date:	**N/A**
Continuing Review Due Date*:	**N/A**

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution

Appendix H

Guidelines

Guidelines to Assist with the Design of
Instruction Delivered via **A**ugmented **R**eality

Designed and Developed by
Chris Wasko

Scenario

Imagine you are driving down a busy freeway and you get a flat tire. You pull off to the side of the road and recall that you have no idea how to change a tire on your car! You grab the jack and spare tire from the trunk, turn on your AR glasses (in the next section, you will learn all about AR but for now, just know that it essentially involves overlaying digital information on a view of the real world), and prepare to change your tire.

Crouching down near the flat, you see digital instructions overlaid on your field of vision (as shown in Picture 1). The instructions teach you exactly where to put the jack, how to use the jack, how to loosen the lug-nuts, and everything else you need to know in order to change the flat tire. Because the instructions are delivered directly in your field of vision, you do not have to waste extra time or effort learning how to fix the flat by looking back and forth between the owners manual (with small print that may be hard to read on the side of the road while cars are whizzing by) and the car. Following the AR instructions, you quickly change tires and are back on the road in minutes!

In this picture, you see relevant digital information placed (co-located) directly in your field of vision while spoken text provides additional details. Note how the red line (meant to be lined up with the wheel) is used to help the digital content align with the physical world. This alignment ensures that the digital information is presented in the correct location.

Introduction

The aforementioned scenario depicts an example of how Augmented Reality (commonly referred to as AR) can be used to deliver instruction. As you are about to learn, AR involves overlaying digital information on a view of the real world. The information in this set of guidelines will be useful if you are an entrepreneur or a programmer who has an idea for an instructional application of AR. The information presented here can be used to guide the design and development of your application and will help you make decisions that will, in turn, facilitate learning.

As you are about to read, AR related hardware and software have, in the past, been accessible by a relatively small amount of people. However, application programming is easier than ever thanks to software like Game Salad and M.I.T. App Inventor. Also, wearable displays (such as those being prototyped by Google) are about to hit the market in the very near future. Such wearable displays hold the potential to dramatically alter how people experience AR. Given the related advances in necessary hardware and software, anyone with a idea for an instructional application of AR will be able to make their idea a reality and will be able to share their creation with users who own the necessary hardware.

AR

Let's spend some time learning a bit more about AR including the definition, history, typical uses, and associated hardware and software.

Definition

There four key components of the definition frequently used by researchers (Azuma, 1997; Carmigniani & Furht, 2011) who study and write about AR.

AR involves (1) overlaying digital information, in a (2) spatially appropriate location, on a (3) real-time (4) direct or indirect view of the real world. Seems pretty confusing but with a little explanation, it is easy to understand.

Let's look at each piece individually:

overlaying digital information:

This extra digital information is responsible for the "augmented" part of AR. The information, which can come in the form of words, pictures, or sounds, is generally used to help the user better understand their immediate environment. Think back to how the digital arrows and spoken words helped the person in our introductory scenario understand where to place the jack.

in a spatially appropriate location:

This part of the definition simply means that in order for the digital information to be of use, it needs to be in the right place (or as close to the right place as possible!). Think back to the spare tire scenario in our example and imagine the disaster that would ensue if the jack was placed in the wrong location. Placing the digital content (called *spatial registration* by researchers) generally happens in one of two ways. The first is done by relying on the GPS features of the device that is being used to display the digital content (smartphone, tablet, AR glasses...more on this in the next section) to determine the approximate location of the user. Once the location is approximated, the content is displayed. Of course, this type of registration can be a bit unreliable at times because the GPS on your device may not be very accurate. A more reliable way of presenting digital content in the correct location involves using the device's camera to scan a marker. Frequently, QR codes such as the one shown here are scanned to trigger the display of digital content. However, it is not possible to paste QR codes all over the world!

New software by Aurasma allows for pictures and objects (someday it is possible that something as complex as a human face can be used to trigger AR) to be used as a way to display digital content. Obviously, AR that is displayed because of a marker (QR code, picture, or object) is generally much more accurate in terms of spatial location because the marker will be in a very reliable, fixed location that is right there in front of you!

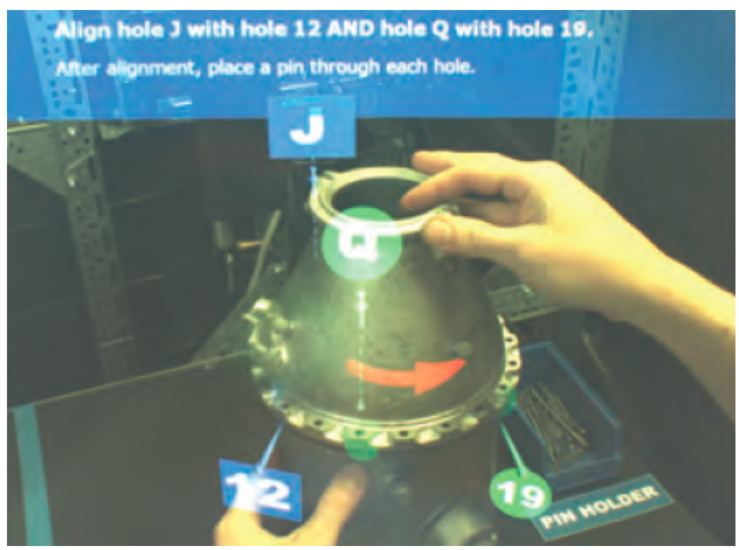
real-time:

This is an essential part of the definition. AR must relate to what is happening in the present. Think of a movie in which digital special effects have been used. Despite the addition of digital content, this example would not count as AR because the events are not taking place in real time.

direct or indirect view of the real world:

The movie example would also fall short here because it is not a direct or indirect view of the real world. The flat tire scenario described in the introduction would count as a direct view of the real world and the yellow line that indicates a first down in American football broadcasts would also count- as an indirect view (you are not seeing the field directly with your own eyes but rather through your TV).

Let's review the parts of the definition by looking at an example from a recent study. Here, we see AR being used to train users to assemble an aircraft engine. The learner is experiencing the AR content via a wearable display. Note that digital information (in the form of both words and pictures) is being overlaid in a relevant location on a real-time, direct view of the real world resulting in an example of AR that meets all of the requirements specified in the definition.



Henderson, S. J., & Feiner, S. K. (2011a). Augmented reality in the psychomotor phase of a procedural task. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on* (pp. 191-200). IEEE.

Learning Aided by AR

Most of the writing and research related to AR involves teaching users to perform/practice a skill or task. This type of learning is called procedural learning. As you saw in the tire scenario, the ability to place relevant digital information directly in the field of vision can minimize the mental and physical effort of performing a procedural task (Henderson & Feiner, 2011a, 2011b; Neumann & Majoros, 1998).

For example, have you ever looked at some printed instructions when assembling a piece of furniture and been confused about how to complete a certain step? Have you experienced repeatedly looking back and forth between the paper and the piece of furniture in an attempt to figure out where to put a specific screw or nail? Have you ever made a mistake when assembling a piece of furniture only to find out about the mistake 10 steps later? Issues such as these are exactly the ones that AR can help with.

By using digital arrows to direct your attention or by showing you a 3D model of how an object should look after a step is completed, AR can help you with the mental aspects of procedural learning. Essentially, AR makes it easier for you to orient yourself to a task and to figure out what needs to be done. Once you have figured out what to do, digital AR content can help you do it correctly by presenting animations that show how a specific part should be moved or using color coding to help you properly align real-world objects.

Researchers have used AR to study how it can help with a large variety of procedural tasks (De Cresenzio et al., 2011; Haritos & Macchiarella, 2004; Henderson & Feiner, 2009; Henderson & Feiner, 2011a, Henderson & Feiner, 2011b, Neumann & Majoros, 1998, Tang et al., 2001). For example, AR has been used to train mechanics to maintain airplanes, tanks, and other machinery. AR has been used to help physicians learn how to perform intricate injections and other surgical procedures. AR has been used to help workers assemble objects and select specific objects from a large warehouse and it has even been used to help people learn to play the bass guitar and to allow them to practice acting out a movie scene.

Since AR has proven useful for a wide variety of procedural learning goals, such goals will be the focus of the guidelines presented in this document. Ideally we will be able to identify some procedural learning goals that apply to the average person (things like tying a shoe or writing words in cursive) and use AR to help teach them. Such use will take the documented benefits of using AR to help with procedural learning and apply them to everyday life.

Also, as you are about to hear, the pending release of wearable displays means that the average learner will now have full use of their hands while experiencing AR. This development matches up quite well with using AR to teach procedural tasks.

AR Past, Present, and Future

The term AR dates back to the early 90s (Caudell & Mizell, 1992). Early uses of AR involved clunky head mounted displays that made it hard to move around and could only display very basic digital content.

Fast-forward to the present where AR can be accessed through applications and browsers loaded on smart phones or tablets. Typically, users hold their device in front of their eyes and see digital content overlaid on their view.

In the very near future, wearable displays such as Google's Project Glass will be available for purchase by the general public. Wearable displays hold the potential to radically change the manner in which people experience AR and could lead to the development of some amazing applications that take advantage of the new displays. In fact, Google has even started a Glass Collective related to their Glasses to help them figure out how people are going to use the devices.

Over the next few years, we will work together (programmers, instructional designers, entrepreneurs, educators) to figure out how to maximize the benefits AR delivered via wearable devices. Based on the work of researchers who have identified the benefits of using AR for procedural learning, it is not hard to imagine that such benefits will be applied to the commercial/public domains. The guidelines discussed in this document should provide some assistance to those who want to design instructional applications that attempt to help the general public learn procedural tasks via AR enhanced instruction.

Learning Theory

Cognitive Information Processing

All of the AR instructional message design guidelines that are going to be covered in the next section of this document have been adapted from some principles of multimedia learning. Multimedia learning involves learning from words and pictures (Mayer, 2005). Since the type of AR learning goals we are talking about will involve words and pictures, it is imperative that you learn how to use them effectively when designing instruction. Much of the research related to multimedia learning is built upon a theory called cognitive information processing (CIP) (Baddeley, 1999; Paivio, 2007). A theory is a system or set of ideas that is intended to explain something. CIP attempts to explain how the human mind works. According to CIP, there are three fundamental parts of the human processing system. Below, they are explained as simply as possible:

(1) *sensory memory*-

Think of the room you are sitting in right now and imagine that you were forced to pay attention to every possible minute detail at once: the humming of the lights, the smell of food coming from the next room, the noise of the wind outside, the feel of your chair, the brightness of the lights, the pressure your elbows are placing on your desk, and so on. Sounds like a nightmare, right? How would you ever be able to accomplish anything if you had to attend to every possible detail at the same time? This is where sensory memory helps us out. Thankfully, our body has an amazing ability to block out some things (usually things that are not important to us) while actively attending to others (like this document). This process of ignoring some things and attending to others allows select information to enter into our working memories.

(2) *working memory*-

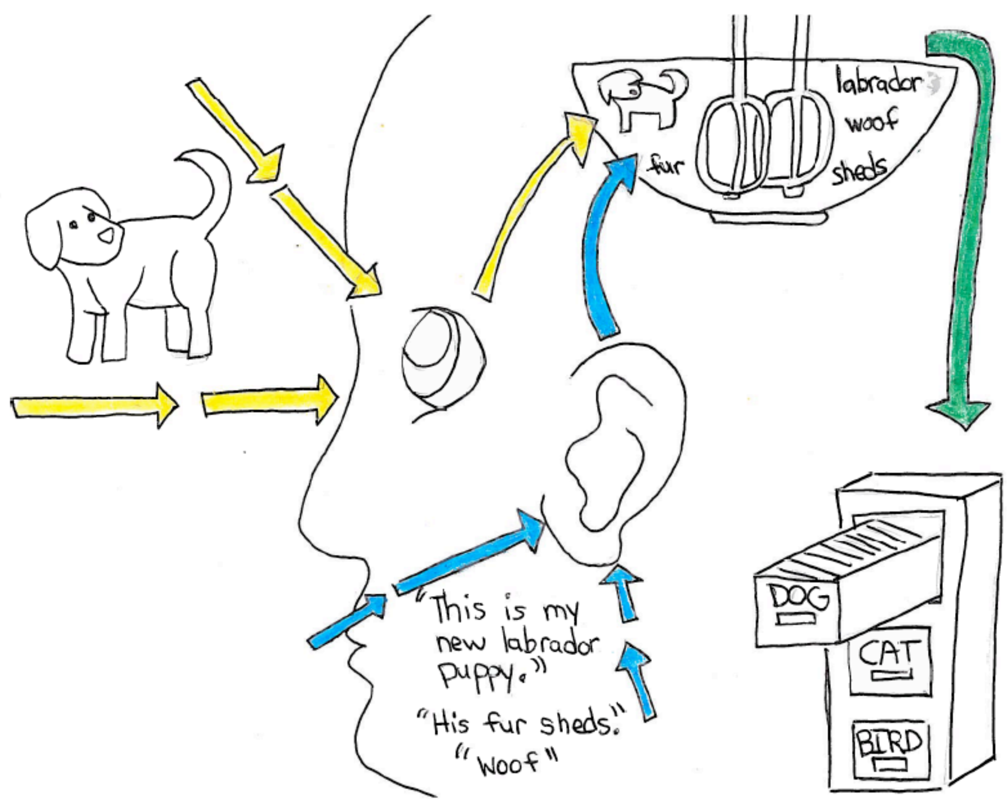
This part of the CIP model has been given a perfect name because this is where much of the “work” of learning is done. In working memory, we manipulate information that entered through our senses and attempt to process the information in a way that will help us remember it. Just like you can only jump so high and do so many pushups, working memory also has limits. As you have likely heard at some point, working memory can usually process approximately seven “chunks” of information at the same time. Also, if we do not use it (the information held in working memory), we will lose it! For example, think back to what you listened to on the radio while driving to your office this morning. You may be able to remember a song or a part of a news story because something about that information stood out while you held it in your working memory. However, the information that did not stand out is now gone for good!

This metaphor may help you better understand working memory: Imagine a pizza chef making dough. The chef takes flour and water (some pieces of information that entered through sensory memory) and combines the two in a pile. The chef starts mashing the two together, pulling and stretching and tossing the dough in the air and then stretching and pulling some more. This is similar to the kind of "work" that is done in working memory. Eventually, after enough manipulation, the two ingredients have been successfully changed into a piece of dough that is ready to be used for something else.

(3) *long-term memory*-

Generally, once we have meaningfully worked with information in working memory, we are able to store it in our long-term memory for later use. The more we have worked with some information or the more meaningful it is, the easier it will be to access the information when we want to "retrieve" it from our long-term memory.

Hopefully, this simple graphic will help you remember the model of the mind that is suggested by CIP.



Here is a final example of the CIP model in action. Try to recall the name of your fourth grade teacher. Possibly you were able to say his/her name right away, perhaps you had to think for a few moments before you came up with the name, or perhaps you were unable to recall the information. At some point when you were in fourth grade you heard the teachers name with your ears and read it using your eyes (sensory memory) and you found a way to remember the name (working memory) until you knew it (long-term memory). Now, years later, the name is obviously long gone from your working memory. Perhaps you were able to quickly bring it back to your working memory because the teacher did something really nice for you (making him or her very meaningful). Maybe the teacher was boring and you had to think about the look of the classroom, a field trip you took, or the names of some classmates to help you remember. Perhaps you could not remember the name because the teacher did not stand out enough or you have since learned the names of so many teachers that they have all blended together. Either way, hopefully you now have a basic understanding of how CIP works and we can go ahead and move on to the next section.

Multimedia Learning

All of the multimedia principles that have been adapted in this set of guidelines come from a body of research by scholars such as Mayer, Sweller, Chandler, Low, Moreno, and Clark.

Dr. Mayer (2005) defines multimedia learning as learning from *words and pictures*. The words can come in the form of printed text or spoken text and the pictures can be static (like a photo or diagram) or dynamic (like a video or animation). Dr. Mayer's definition is really useful because it does not limit us to one medium. For example, a teacher giving a lecture while referring to a diagram, an educational video, or the last miserable PowerPoint lesson you sat through can all be considered multimedia learning according to Dr. Mayer's definition.

Dr. Mayer (2005) developed a theory called Cognitive Theory of Multimedia Learning (CTML) that attempts to explain how people best learn from words and pictures. CTML borrows concepts from other researchers and it is the theory that directly underlies the five guidelines that will be discussed in the next section. CTML sounds complex but it is pretty easy to understand! CTML is based on three fundamental assumptions that are described below. The good news is that the three assumptions are very much related to what you learned about CIP.

(1) *people have separate channels for processing words and pictures*

This simply means that CTML supposes that you process verbal (printed text or spoken text) and pictorial (static or dynamic pictures) information in different ways. Researchers have conducted experiments where they have shown how channel interference (asking people to look at a picture and respond to a printed text question- this caused interference because people had to use their pictorial channel for both tasks) led to bad results on a performance test. Additionally, researchers have attempted to prove that we have separate channels by presenting a rapid sequence of pictures and a rapid sequence of words. People consistently did better at recalling the pictures in any order but were better at recalling the words in sequential order, suggesting that people processed the words and the pictures in different ways.

It is widely believed that pictures are remembered better than words. When people are shown a picture, they automatically assign it a "word label" in their working memory. So when a person is shown a picture of a dog, they will have a mental image of that picture and they will assign it a label like "Chuck's terrier". If a person is shown a new word like, "faineant" and they do not know what they word means, they will be unable to form a mental image (of me!) and will have to rely on only one channel to remember the word. See how, in certain cases, pictures have the ability to outnumber words 2 to 1?

(2) *each channel is limited in capacity*-

If you remember, we have already discussed some of the limits of working memory in the last section. For CTML, this assumption means that each channel (the verbal channel and the pictorial channel) can only handle so much information at one time. The amount of information that can be handled may vary from person to person but generally, the channels are quite limited. Imagine trying to take notes on an audio recording while your teacher is talking aloud, giving you instructions about what to do with the notes. This double demand placed on your verbal channel would have very negative consequences. Try to remember that working memory can only manipulate about 2-4 elements at one time and that without use, the contents of your working memory will be lost in about 20 seconds.

Have you ever used an old computer that had too many applications running? Did you notice how this made the computer strain and how it slowed down every piece of software you tried to use? This metaphor is not unlike what happens when channel limits are tested. As with an overloaded computer, the performance of working memory can be compromised due to information overload.

(3) *learning is an active process*-

Think back to our pizza maker in the last section. Learning involves active mental effort. How easy would it be if our minds were simply hard drives that could have information uploaded to them with no effort required on our part? Obviously that is not the case. CTML suggests that as learners, we are required to select relevant words and pictures, organize the words and pictures into something meaningful and find a way to integrate this new information with what we already know so that we can access it when we need it. Of course, all of this requires work on your part. Anyone who has ever taught a room full of tired middle or high school students knows how challenging it can be to get them to jump start their inner pizza maker!

That's it. We are ready to take a look at the five guidelines in depth. Remember that each of the guidelines are designed with CIP in mind and each of the guidelines can be used by designers to help people learn efficiently and effectively from instruction that contains pictures and words.

5 Guidelines

Preface

As you know, each of the five guidelines are grounded in CIP and have been taken from the multimedia learning research base and the research of Dr. Mayer and his associates. The guidelines have been adapted to apply to using AR to help teach procedural tasks (the type of learning that most of the AR research has addressed).

For each guideline, you are going to be presented with a definition, a short description and a pictorial example to get an idea of how the guideline looks when it is applied. Each pictorial example will be based on the same scenario: A person wearing AR glasses is standing in front of a mirror and is being presented with instructions about how to tie a tie using a Windsor knot. Note that this scenario would also be possible if a person was using the front facing camera on a standing tablet device.

Let’s get started!

INSTRUCTIONAL DESIGN GUIDELINES FOR AR

1. use both words and pictures and use them appropriately

Guideline 1 is adapted from what Dr. Mayer refers to as the "multimedia principle" (Mayer, 2008; Mayer & Anderson, 1991; Mayer, Bove, Bryman, Mars, & Tapangco, 1996; Mayer & Gallini, 1990). This principle simply states that people learn better from words and pictures than they do from words alone. In the previous section, we have already talked about how using words and pictures takes advantage of both channels and gives your working memory more information to use.

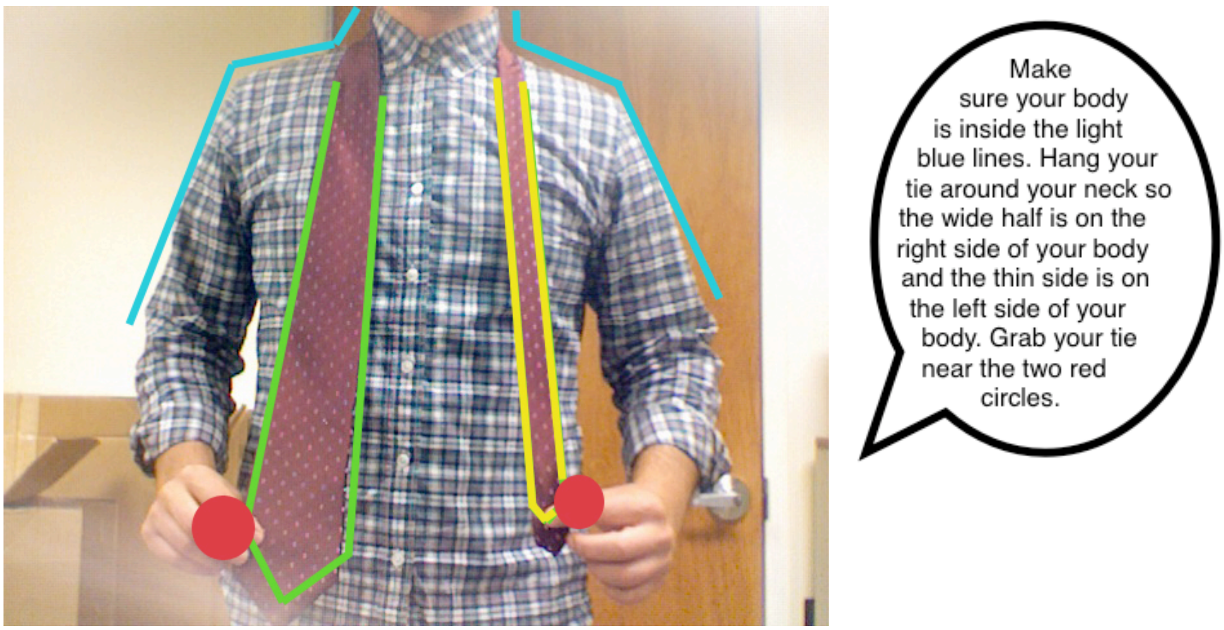
The multimedia principle is most useful when it is applied to long chunks of text or long lectures that do not use any pictures to help facilitate learning. Think of how a well-designed picture can help you gain a better understanding of a challenging concept. Hopefully you have already seen evidence of such use in this document! The most important thing to remember when adding pictures to words is that the picture is related to what is being talked about in the text. Adding a visual that is unnecessary or unrelated will potentially confuse your learners and interfere with their learning.

So, what does this have to do with learning a procedural task via AR? The cool thing about AR is that, as we saw in the definition, visuals play a huge role. Because AR involves looking a view of the real world, there will always be a relevant, high definition visual in place. Instructional designers just need to add a little digital information to it in order to help our learners understand what we want them to do. Think of the flat tire example that was presented at the beginning of this document. The real world provided the majority of the visual and we just needed to add a bit of digital information (something like an arrow or a simple animation) to help learners understand exactly what we want them to do. As long as the digital information is placed in the right location, we will be on our way to using words and pictures in concert with each other.

Going back to the theories we talked about, using both words and pictures that are **related** to each other and that provide relevant information takes advantage of both channels and gives our learners better chances to make connections in their working memories.

Here is an example of Guideline 1 in action. Remember that in this scenario we will see a person standing in front of a mirror wearing AR glasses and learning how to tie a Windsor knot. Notice how there is a digital outline of a body that helps the user stand in the correct spot. This helps make sure that the digital content is presented in the correct location and it related back to the spatial registration term we learned earlier.

Notice how a small amount of visual information is added to the view of the real world in order to help our learner get started while, at the same time, some spoken text is played aloud to better explain what is happening in the picture.



What would happen if we took away our AR instructions and gave our learners a different set of directions that contained only printed text without pictures? Learners could easily misinterpret the text and make a mistake, which would complicate the learning process. Printed text directions with some pictures would be a bit more helpful but would still require our learner to expend effort making sense of the words and pictures and looking back and forth between the directions and a mirror. Our AR instructions seem to minimize both the opportunities for mistakes and the physical and mental effort required of our learners.

2. generally, use spoken text instead of printed text

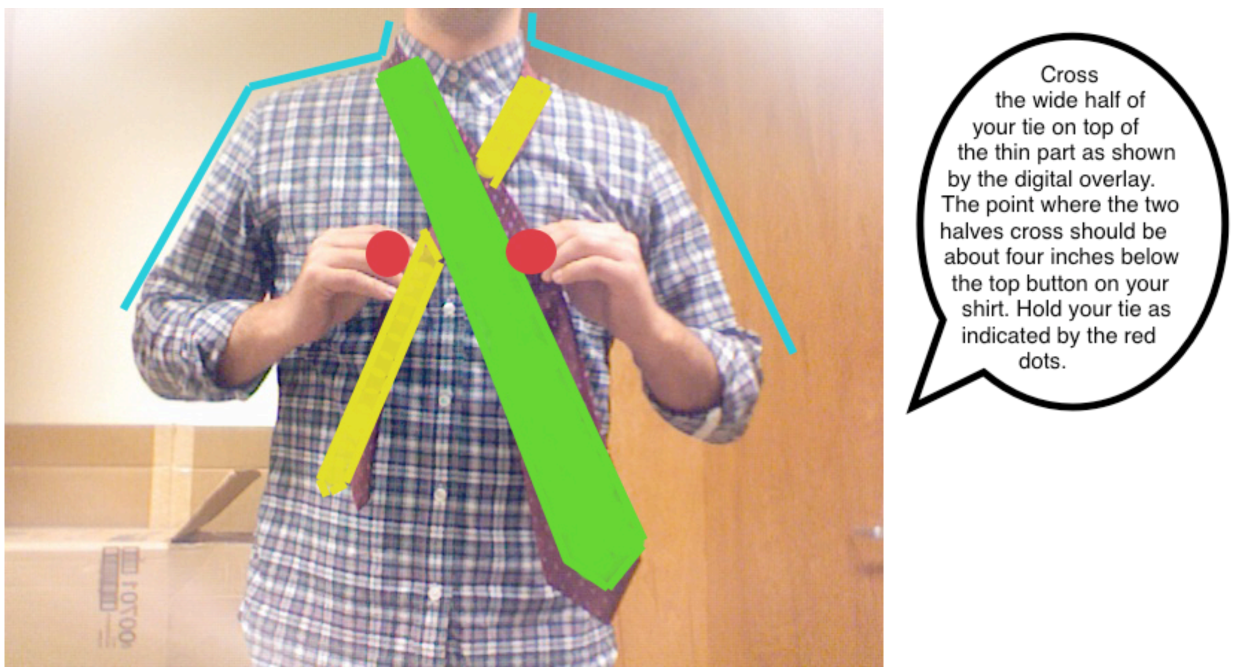
Guideline 2 is adapted from what Dr. Mayer calls the "modality principle" (Mayer, 2008; Mayer, Dow, & Mayer, 2003; Mayer & Moreno, 2002). The modality principle seeks prevent learners from being asked to split their attention between two sources of information. Think words and words or pictures and pictures! Refer back to the scenario in which a person was listening to an audio recording while a teacher spoke instructions aloud. This double requirement of the verbal channel is problematic because it puts a learner in the uncomfortable position of having to listen to two things at once.

For the purposes of AR instruction, we know that the visual channel is always going to be asked to process some relatively complex information. Adding to the demand on the visual channel by presenting large chunks of printed text seems like a really bad idea. Such a design decision will require learners to split their visual attention between the printed text and the corresponding parts of the visual.

Think back to when we learned about how humans are believed to have a verbal and a pictorial channel and that both are limited in capacity. When we double up in one channel (in this case, it is the visual channel that is under assault), we create a traffic jam of sorts. By simply switching the printed text to spoken text, we can avoid the traffic jam and make use of both channels.

Please do not interpret this guideline as, "never overlay printed text on a view of the world when designing AR instruction". When a person is being asked to look at a complex visual with many relevant parts (under the hood of a car, for example), some small printed text labels placed close to the part they are referring to can actually help your learner get oriented. What we are looking to avoid here is presenting large chunks of printed text directions on a view of the world. Such a decision would unnecessarily ask your learners to split their attention.

Look at this example of Guideline 2. Our learner has now moved on to step 2 of learning how to tie a Windsor knot. Notice that the directions are presented as spoken text (via the speaker on the AR glasses) instead of printed text. This design decision avoids requiring our learners to split their attention between two sources of visual information. Also, thinking back to our first principle, notice how the spoken directions correspond to what is happening in the picture.



3. make sure related words and pictures are presented at the same time

If Guideline 3 seems related to Guideline 1, it is. Guideline 3 is based on that Dr. Mayer calls the "contiguity principle" (Mayer, 2008; Mayer & Anderson, 1992; Mayer & Sims, 1994; Moreno & Mayer, 2000). The contiguity principle states that people learn better when corresponding words and pictures are presented simultaneously.

To better help you understand Guideline 3, let's imagine that instead of using a still picture (the simple green and yellow digital version of the tie that we have been seen in all of our examples), we will be overlaying a moving animation that indicates how the tie should be moved in order to complete each step. Now imagine if we played some complex spoken text that stated, " While holding the wide end of your tie in your left hand and the thin end of your tie in your right hand".....and so on, AND THEN the animation showing the corresponding movement played. What would be problematic about not showing playing the spoken text and showing the animation at the same time?

The answer goes back to the fact that we have a working memory that is limited in capacity. By separating the spoken text from the animation, we are forcing learners to expend mental effort by holding the words in their working memory until the animation is played. By playing the spoken narration and showing the animation at the same time, we can help learners make more significant connections between the two pieces of information.

We will not have an example for this guideline since it is not possible to play an animation in this document. Hopefully you get the idea that presenting related words and dynamic visuals such as animations at the same time is a simple design decision that can be used to help learners comprehend what is being taught. Guideline 3 is another example of what it means to use words and pictures effectively when designing instruction.

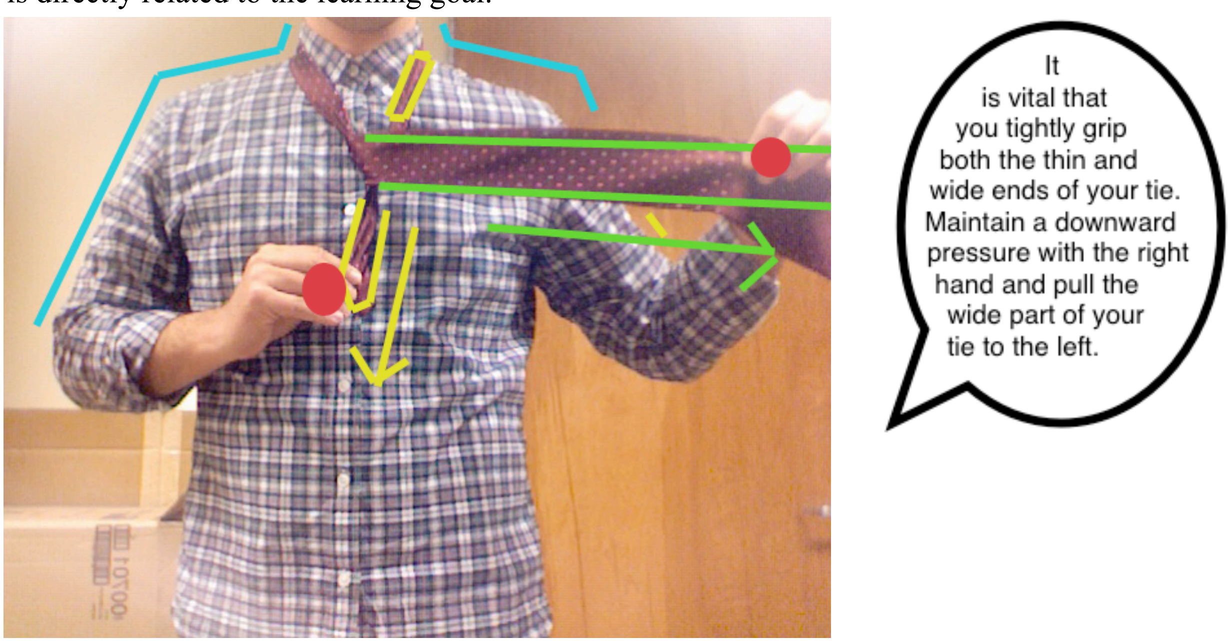
4. do not add any unnecessary information to your instruction

Designing instruction to help a person learn a procedural task from AR is all about making their life easier. Adding extra information that is not related to the task at hand is a really bad idea. Guideline 4 is adapted from what Dr. Mayer calls the "coherence" and the "redundancy" principles (Mayer, 2008; Mayer, Heiser, & Lonn, 2001; Mayer, Bove, Bryman, Mars, & Tapangco, 1996). Both principles address taking steps to make sure that only essential material is presented to learners.

What do we mean by unnecessary information? Things like sound effects that do not help the learner better understand the subject, extra visuals that are not related to the subject, and unnecessary stories or anecdotes related to the learning goal. The idea is to make sure that the learner is presented only with material that is essential to the learning goal.

Let's think back to how this relates to the theories we learned about earlier. We know that people have limits on how much information they can process, right? By presenting learners with unnecessary or unrelated content, you are asking them to waste valuable time and mental effort that could otherwise be spent on making connections between useful pieces of information.

Guideline 4 is especially important for AR learning because we know how busy and distracting the real world can be. Generally, trying to learn in a real-world environment can be a working memory nightmare. This makes it essential that we avoid contributing to the busyness of the learning environment by adding things like sound effects or decorative animations. Notice how in this example (and all of the others) the presentation is clean and concise. All words and visuals are directly related to the learning goal. Learners are presented only with information that is directly related to the learning goal.

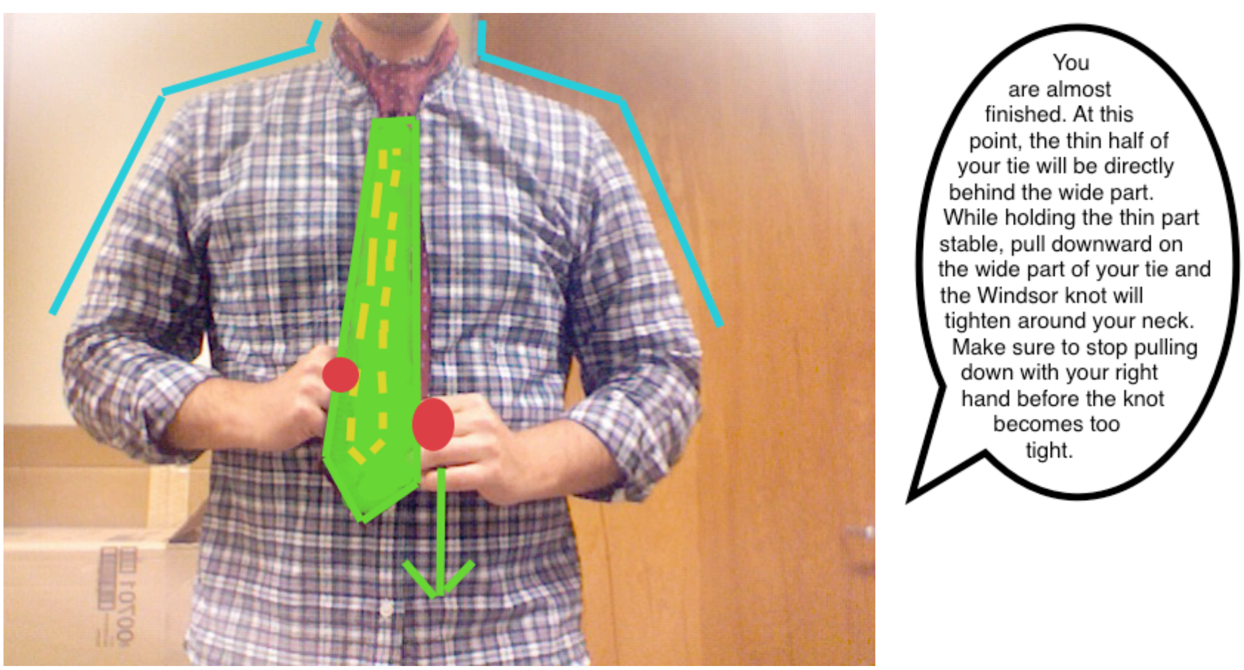


5. use a conversational tone

Guideline 5 is adapted from what Dr. Mayer calls the “personalization principle” (Mayer, 2008, Mayer, Fennell, Farmer, & Campbell, 2004; Moreno & Mayer, 2004). The personalization principle suggests that people learn better from a multimedia lesson when words are presented in a conversational style rather than a formal style. Sounds pretty simple, right? Dr. Mayer conducted an experiment where sentences like, “During inhaling, the diaphragm moves down, creating more space for the lungs, air enters the nose or mouth, moves down through the throat and bronchial tubes to tiny air sacs in the lungs” were compared to the same sentence except the word “the” was changed to “your” so the new sentence read, “During inhaling, your diaphragm moves down...”. Dr. Mayer found that people who read the conversational sentences performed better on tests than people who received the formal text. This simple change actually had an impact on learners’ ability transfer their knowledge to a new task.

Thinking back to what we know about learning being an active process, perhaps Guideline 5 facilitates that active process by establishing a sense of social partnership with learners, thus making the content just a bit more meaningful to them.

When you are drafting the spoken text for your instruction, make sure that you attempt to maintain a conversational tone and personalize the text wherever possible. In this final example, as with all the examples, we see that the spoken text is presented in a conversational style rather than a formal style.



Summary

So that is it. Here are the five guidelines once more:

1. use both words and pictures and use them appropriately

2. generally, use spoken text instead of printed text

3. make sure related words and pictures are presented at the same time

4. do not add any unnecessary information to your instruction

5. use a conversational tone

Remember that these Guidelines are adapted from multimedia learning principles that can be applied to ANY type of learning that involves words and pictures...a lecture, an educational video, or a PowerPoint presentation.

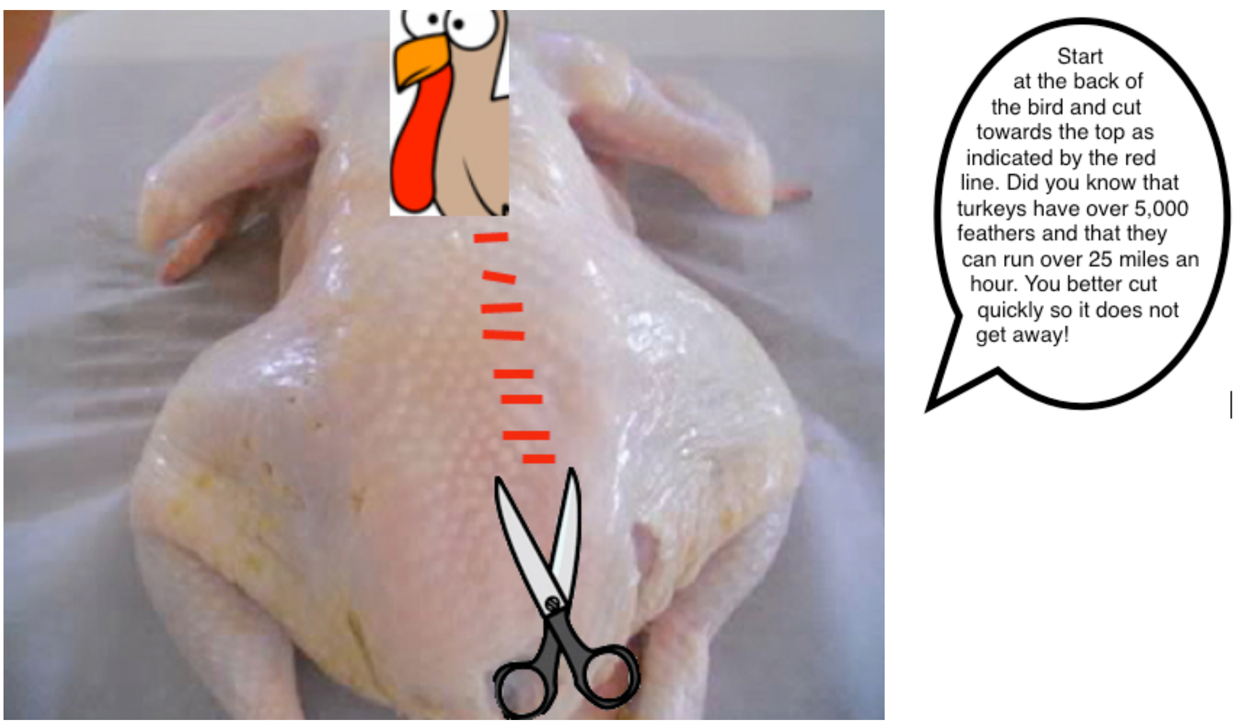
In this case they have simply been applied to AR which is an emerging instructional delivery method. We have focused on procedural learning since the overwhelming amount of AR research has involved AR being used to assist with the learning/performance of specific procedures. And with an affordable (relatively) wearable AR display ready to his the market in the near future (devices that will allow users to use both of their hands to complete procedural tasks), everyone is going to be scrambling to figure out how AR will be used.

If you are an entrepreneur who has an idea of how AR can be used to help the average person learn a procedural task or if you are a developer who has been hired to make someone else's vision come to life, these principles should help you effectively manage words and pictures thereby developing a more effective AR application.

Assessment

Finally, here are some scenarios (again a person is using a wearable display to learn a procedural task) where the AR instructional design principles have been misapplied. See if you can determine what is wrong in each example.

In this scenario, a person is learning the 8 steps involved in carving a turkey. What would you change about the instruction?

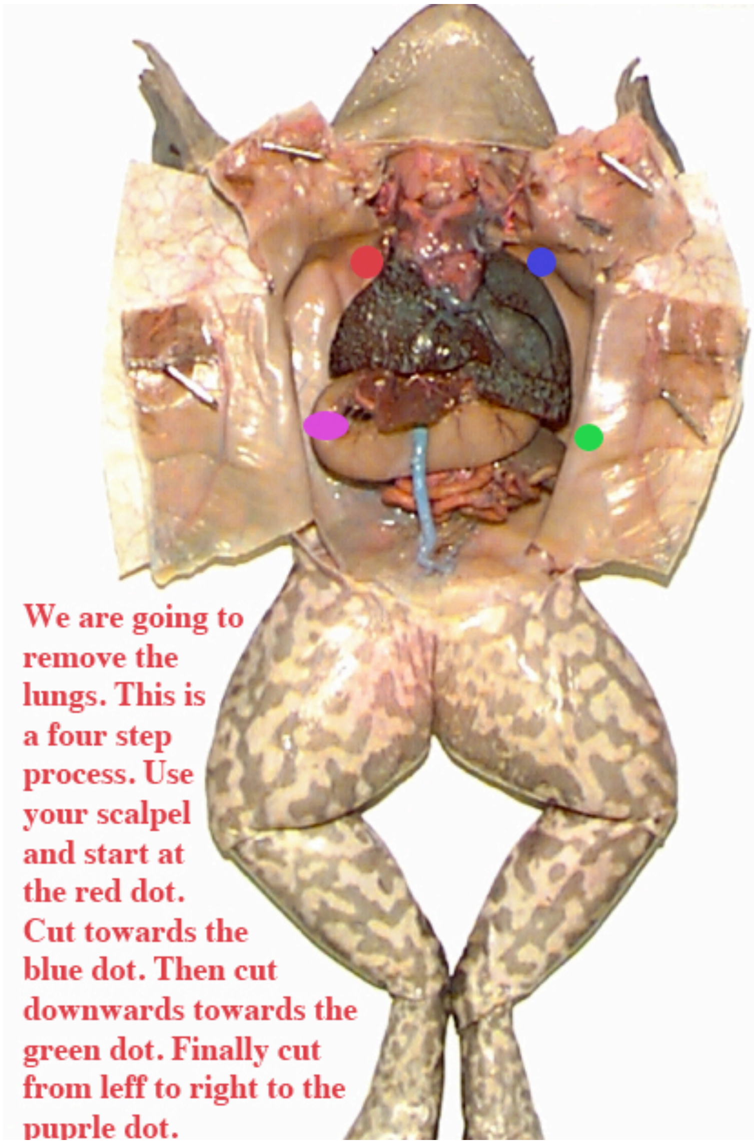


Surely you noticed all of the nonessential information (visual and spoken text) that was included in this scenario. The extra information has nothing to do with learning how to carve a turkey and seems to violate Guideline 4.

Here we see some instruction that attempts to help teach how a car engine works. What would you change in this scenario?

Surely you noted how the text could easily be changed to make it a bit more conversational in style as was discussed in Guideline 5. Again, it may seem simple but changing “the” to “your” has been shown to make a difference in terms of learning transfer.

In this scenario, a person is using an AR display to dissect a frog. What changes would you make to the instruction?



Surely you noticed the printed text on the screen. This visual is complex enough without adding a large chunk of printed text. Adding the printed text places an unnecessary burden on the visual channel. As discussed in Guideline 2, where applicable, try to present text in spoken rather than printed form.

References

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, *6*(4), 355.

Baddeley, A. D. (1999). *Essentials of human memory*. Psychology Press/Taylor & Francis (UK).

Carmingniani, J., & Furht, B. (2011). Augmented reality: An overview. In B. Furht (Ed.), *Handbook of augmented reality* (pp. 3-45). New York, NY: Springer.

Caudell, T.P., & Mizell, D.W. (1992). Augmented reality: An application of heads-up display technology to manual manufacturing processes. *Proceedings of the Hawaii International Conference on Systems Sciences*

De Crescenzio, F., Fantini, M., Persiani, F., Stefano, L. D., Azzari, P., & Salti, S. (2011). Augmented reality for aircraft maintenance training and operations Support. *IEEE Computer Graphics & Applications*, *31*(1), 96-101.

Haritos, T. & Macchiarella, N. D. (2005). A mobile application of augmented reality for aerospace maintenance training. In *The 24th Digital Avionics Systems Conference, DASC 2005*, pp. 5.B.3–5.1-9.

Henderson, S. J., & Feiner, S. (2009). Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In *Mixed and Augmented Reality, 2009. 8th IEEE International Symposium on* (pp. 135-144). IEEE.

Henderson, S. J., & Feiner, S. K. (2011a). Augmented reality in the psychomotor phase of a procedural task. In *Mixed and Augmented Reality, 2011. 10th IEEE International Symposium on* (pp. 191-200). IEEE.