The Effectiveness of Instructional Video in the Acquisition of Cognitive, Affective and Psychomotor Skills in Practical Sports Therapy Rehabilitation.

COOPER, DARREN,JAMES

The Effectiveness of Instructional Video in the Acquisition of Cognitive, Affective and Psychomotor Skills in Practical Sports Therapy Rehabilitation.

Abstract

The use of instructional multimedia, particularly video, within education is steadily increasing although the evidence-base regarding its usage typically only indicates that it is equivalent to or as effective as live demonstration or traditional teaching methods. The current study undertook a longitudinal quasi-experimental crossover study, over three consecutive academic years to evaluate the efficacy of instructional video to teach cognitive, affective and psychomotor skills to level 5 undergraduate sports therapy students. Through the use of a crossover design students undertook both the video and control conditions, they were assessed formatively on a weekly basis to provide a consistent measure of performance throughout the eighteen weeks of data collection within each year. The instructional videos used within the study were based upon (as far as possible) the multimedia principles proposed by Mayer to reduce extraneous cognitive load and maximise essential intrinsic and germane cognitive load. The results from the study were analysed with the use of effect size statistics and interpreted though the use of magnitude based inferences, an emerging alternative to the traditional use of null hypothesis testing. The findings of the study indicate that the use of the instructional videos was beneficial to the vast majority of the students, which builds upon the current evidence-base as it demonstrates that they can be used to enhance academic practice rather than be used as an equivalent resource.

DURHAM UNIVERSITY

The Effectiveness of Instructional Video in the Acquisition of Cognitive Knowledge, Affective and Psychomotor Skills in Practical Sports Therapy Rehabilitation.

Darren James Cooper

Thesis submitted for partial fulfillment of the

Degree of Doctor of Education

School of Education

University of Durham

2015

Contents

Table of Illustrations

Table of Tables

DECLARATION

No part of the material offered in this thesis has previously been submitted by me for a degree in this or in any other University.

STATEMENT OF COPYRIGHT

ACKNOWLEDGEMENT

I would like to take this opportunity to thank Professor Steve Higgins and Dr Nadin Beckmann for their guidance, support, invaluable advice and feedback throughout the duration of this thesis that has enabled it to be completed together with publications submitted along the way.

I would also like to thank my previous colleagues and the students of Teesside University for their assistance and participation in the study, without them the study could never be undertaken.

Finally, I would like to thank my wife for her support and patience during the write-up of the thesis.

Chapter 1: Introduction

1.1 Background of the study

Sports Therapy is a relatively young allied health profession (established in 1990) that encompasses musculoskeletal examination & assessment and a variety of treatment modalities ranging from spinal mobilisations to electrotherapies and rehabilitation. Sports Therapy differentiates itself from other professions such as physiotherapy, chiropractors and osteopathy by placing a significantly greater emphasis upon rehabilitation. This ensures that the patient is able to return to work as quickly and safely as possible but more importantly to return them to their chosen sport at the previous level (or higher) they were competing at, at the time of injury. As a result, a range of protocols has been devised over the years for specific injuries, arguably the most famous of these is Shelbourne's ACL rehabilitation program. The program commences after surgery and aims to return the professional sports participants in a 6-9 month time frame (roughly half to a third of the time most sedentary ACL patients would need to return to full function (Shelbourne, 1990). Whilst the specific protocols are of great use to the Sports Therapist there has to be an over-arching continuum that can be applied to any injury that the sports participant may encounter from complete bone breaks to minor ligamentous sprains. Many authors have proposed specific rehabilitation criteria and whilst there are subtle discernable differences they are all fairly similar. As such the protocol from Smith, (1998) was implemented as the protocol for the Rehabilitation and Remedial Therapy module where the current study undertook data collection. This was due to the fact that the author was the Chairman of the Society of Sports Therapists (the validating body for the course) so any queries about the protocol could be addressed with ease, ensuring an evidence-based practice approach to the module. The rehabilitation continuum from Smith, (1998) is divided into distinct stages relating to the time frame of the sports participants progression, these are early, intermediate, late and pre-discharge. Each of the first three can also be further sub-divided into early, intermediate and late i.e. intermediate early, which is useful when goal setting with the patient to ensure patient adherence. Smith, (1998) details strict criteria that must be met before patients can progress from one stage to another that are easy to apply and for the patient to understand. They are also useful as a pre-test before sessions to ensure they are able to progress safely, if not then the patient will have to regress to ensure their safety and minimise the risk of reinjury.

The aim of the Rehabilitation and Remedial Therapy module is to ensure that students understand the theoretical implications and application of the continuum, but more importantly, how it should be applied in a practical setting. This aspect however presents a grand task as the continuum can be applied to any injury, at any muscle, tendon, ligament or joint in the human body, in the case of the latter there are 340-360 depending upon the sources used and the inclusion or exclusion of pseudo-joints. As a result, the module refined the application to the main peripheral joints and grouped them together for simplicity (Ankle, Knee, Hip, Shoulder, Elbow and Wrist). Whilst it maybe a simplification it is necessary as the Elbow encompasses the humeroulnar joint, humeroradial joint and the radioulnar joint, which to provide a session for each individually would make the module entirely unfeasible as the elbow is the simplest of the peripheral joints.

The practical aspect of the module is core to the competencies of a Sports Therapist upon graduation, and it is the application of appropriate exercises that are safe and effective for the patient, dependent upon their injury status, which is of utmost importance. Therefore, one of the main learning outcomes of the module is for students to demonstrate their practical competence. As a result, the students are informed in advance, as to which joint and which stage they will have to devise a session for. During the practical sessions the students will perform these sessions in groups, with a group of mock patients and receive formative feedback from their peers and their lecturer. The practical sessions are in preparation for their main summative assessment item (70%), where the students are required to perform a 20-minute rehabilitation session on any of the peripheral joints or rehabilitation stages.

1.2 The use of videos in education

The use of computers, computer technology and the internet is becoming an increasingly important educational resource that is radically changing the teaching paradigm, this has resulted in electronic learning (e-learning) being seen as one of the most popular learning environments in the information age (Alonso, López, Manrique, & Viñes, 2005; Liaw, Huang, & Chen, 2007;

McNulty et al., 2009). E-learning is essentially a different way of teaching and learning and is the ‘umbrella term’ for the various methods that can be used by instructors (Liaw et al., 2007). Regardless of the method employed, e-learning has been cited as a promising alternative to the traditional classroom setting, and it is generally concluded that e-learning will improve, is equivalent to, or in some cases is superior to traditional teaching methods (Chen, 2011; McNulty et al., 2009; Zhang, Zhou, Briggs, & Nunamakerjr, 2006). The vast majority of current University students, have grown up surrounded by multimedia instructional content and are completely comfortable, familiar and able to use the learning resources with no additional training (Lawlor & Donnelly, 2010).

The use of videos has also been cited as being beneficial from a teaching perspective, as their application means that the instructor does not have to continuously repeat lectures, or skilled demonstrations resulting in greater consistency for the learners (Chan, 2010).

One particular area that is starting to adopt the methodologies available within the field of e-learning are the medical professions (Duijn, Swanick, & Donald, 2014a; Ford, Mazzone, & Taylor, 2005).This has been achieved via the rapid expansion of web-based video hosting platforms, although there is a very limited evidence-base regarding the use of instructional videos to teach psychomotor skills, the vast majority of pedagogic research in allied health education has concentrated on the acquisition of cognitive rather than psychomotor skills (Duijn, Swanick, & Donald, 2014b; Ford et al., 2005; Gallagher, Gilligan, & Mcgrath, 2014). Typically professional skills were acquired during placements, this has now changed to the University setting were they are taught in a face-to-face laboratory setting, where they observe a demonstration of the clinical skill and then practice on each other (Duijn et al., 2014b; Gallagher et al., 2014). This does however present a unique set of problems, as the complexity of teaching a hands-on profession in a traditional academic setting, rather than a purely theoretical or philosophical subject have been documented in previous literature (Perry, 1999) and it would be a fair assumption, that these would carry over into the realm of e-learning.

As a result of the challenges posed, educators have responded by adopting various e-learning methods in an attempt to standardise, and enhance the efficiency of teaching (Veneri, 2011). Amongst these methods video has been highlighted as one of the most consistently used methods, however video is an accepted medium that has been used in academic health for a long time (Attstro et al., 2005; Gallagher et al., 2014). Due to the modern technologies currently available, the distribution of electronic videos is easy than ever (Chen, 2011; Pieter Wouters, Tabbers, & Paas, 2007). The videos can illustrate abstract cognitive processes or concepts, and it is this reason that is responsible for bringing video back to the forefront of learning and teaching within medical education (Chen, 2011; McNulty, Sonntag, & Sinacore, 2009; Wouters, Tabbers, & Paas, 2007). Although videos are still being utilised within their typical settings to present theoretical content, there is an emerging area pertaining to the implementation of instructional videos to teach psychomotor skills (McNulty et al., 2009; Smith, Cavanaugh, & Moore, 2011).

Typically the traditional strategies used to teach psychomotor skills include the use of lectures, textbooks, self-instruction and live demonstration (Duijn et al., 2014; Smith et al., 2011). Learning motor skills by watching a live demonstration has been long been recognised as a successful and a well-researched instructional method for over 30 years (Wouters et al., 2007). Previous instructional videos of demonstrations where either broadcast 'off-line' through TV or stored on CD-ROMS, these were not generally made available to students and were included only when the teacher felt it was appropriate for the course (Attstro et al., 2005; Green, 2003; Zhang et al., 2006). However, even though modern technology now makes the dissemination possible, introducing e-learning into the learning process does not necessarily ensure the promotion of learning effectiveness and efficiency (Chen, 2011; Kala, Isaramalai, & Pohthong, 2010). More significantly there is limited conclusive evidence demonstrating its effectiveness for teaching clinical skills (Bloomfield, Roberts, & While, 2010). The acquisition of psychomotor skills among students and the relationship to teaching methodologies is a relatively new area of research in health care education (Veneri, 2011). In general multimedia educational research has been characterised by inconsistencies that have led to inconclusive findings as to the effect that multimedia resources can have upon the learning process (Samaras, Giouvanakis, Bousiou, & Tarabanis, 2006). Tallent-Runnels et al., (2005) expand on this further by stating that there is a serious dysfunction between the technological features that are emerging, and the pedagogical teaching principles relating to e-learning. Pedagogical principles provide a theoretical structure that guides strategies, and activities that form the foundation of good educational practice, within e-learning that are informed by instructional design (Alonso, Lopez, Manrique, & Vines, 2008; Alonso et al., 2005; Kala et al., 2010). Instructional design has developed in combination with three basic learning theories; behaviourism, cognitivism and constructivism (Alonso et al., 2008, 2005). Each of the three theories focus on different perspectives of learning (Kala et al., 2010). Behaviourist theory centres around observable behaviour (objectivity), cognitivism on unobservable behaviour (subjectivity) and constructivism identifies the importance of new knowledge construction and learner-centred experiences (Kala et al., 2010). Whilst it is generally acknowledged that the cognitive and constructivist approaches have dominated the evolution of learning theories, given the nature of the multimedia to be examined within this study, all three basic components will be discussed and their contribution evaluated (Samaras et al., 2006). This is because each makes different assumptions about the way a person learns and remembers knowledge (Samaras et al., 2006). Therefore to maximise the instructional effectiveness of multimedia, it is useful to be guided by the research-based theory of how people learn, and to connect applications with evidence and theory (Mayer, 2002).

The adoption of e-learning technologies, for the education of health professionals has increased progressively throughout the past five decades (Triola, Huwendiek, Levinson, & Cook, 2012). Principally this is due to the literature regarding e-learning, citing the method as a promising alternative to the traditional classroom setting (Chen, 2011; Zhang et al., 2006). It is generally concluded that e-learning will improve learning, is equivalent to or in some cases is superior to traditional teaching methods (Chen, 2011; McNulty et al., 2009; Zhang et al., 2006). Since identifying literature to support the rationale for the pilot study (see Cooper & Higgins, 2014) the use of instructional videos has gained significant momentum. Examples now include Music instruction (Kruse & Veblen, 2012), being used as a pre-teaching tool for English language courses (Seilstad, 2012), as aids for DIY hobbyists, who are moving away from the traditional manual to instructional videos (Swarts, 2012) and by students to visualise elements from chemistry manuals (Benedict & Pence, 2012). However there is a new area that is emerging, and these are what can only be perceived from an academic perspective as negative instructional videos that share practices about how to cheat on academic work (See Seitz, Orsini, & Gringle, 2011).

1.3 Rationale of the study

The video examples above where a model demonstrates how to perform a task, are increasingly being used in an educational setting, as they have become simple to create and disseminate via e-learning environments (Van Gog, Verveer, & Verveer, 2014). A predominant problem associated with multimedia learning, is whether it is possible to promote constructivist learning from passive media (Mayer, 2002). An aspect that is still relevant today as Meij & Meij, (2013) state video is fast becoming the primary medium for instructing users about procedures, however questions arise as to its design and effectiveness. Mayer's Multimedia Learning Theory (MMLT) provides a well-evidenced instructional design approach (Moreno & Mayer, 1999a) that encompasses the main pedagogical theories. It is also used as a foundation for further instructional design principles that are emerging such as those by Meij & Meij, (2013). They cite multiple sources and state that their guidelines summarise key notions of accepted thinking, rather than advancing new theory for instructional design (Meij & Meij, 2013). As a result the efficacy of the e-learning resource is the main area that can be advanced, and that is the prime focus of the thesis, to advance the findings from the pilot study which has been published (Cooper & Higgins, 2014).

1.4 Definitions and terminology

Throughout this thesis instructional video and its use will be referred to in various ways, this is due to the different terminologies used throughout the world. It is referred to as Computer Aided Instruction (CAI), online video; CD or DVD based video, YouTube or iTunes U videos to provide but a few examples. Nevertheless, the videos described and discussed in the literature review will not differ significantly from those used in this thesis, and will have only been included due to their ability to be compared with the current study. There will be variations in structure, editing, timing and many other variables but these are all factors that will be addressed throughout the thesis.

1.5 Summary

As a result of the overview of the context above, it is clear that the use of instructional video is becoming widespread in educational settings; this is due to the development of technologies that can be aligned with teaching methodologies. There are many ways that instructional videos can be deployed when being used to teach cognitive, affective or psychomotor content. However the current study will investigate the role of instructional videos during student's preparation before practical rehabilitation sessions. The study will assess their potential with the use of a formative scale throughout the taught practical sessions of a course, designed to teach rehabilitation skills.

1.6 Aims and Objectives

The aims and objectives of the current study are to evaluate the effectiveness of instructional videos to supplement the teaching of cognitive, affective and psychomotor skills. The instructional videos will be designed in accordance with instructional design principles (MMLT) wherever possible and will be evaluated against evidenced based pedagogical approaches. The current evidence base around psychomotor-based instructional videos will be investigated and utilised to inform and evidence based approach and recommendations.

1.7 Overview of the thesis

The structure of this thesis is as follows: A two-part literature review is used to set the context, and then an overview of the general methodology is presented. After the first two years of data collection a methodological change was implemented therefore a more detailed account of the initial and second phases is presented to introduce each of these sections for the first two years of data collection and then the third year. The results of the study are presented in chronological order and the discussion follows the same order. The discussion is book ended with an overall discussion that focuses on the calculated meta-analysis from all years of data. The discussion encompasses the findings relative to the underlying literature (specifically Mayer and Bloom) and highlights prominent aspects. A final conclusion and implications rounds off the study.

Chapter 2.1 Literature Review

2.1.1 Introduction

The use of instructional videos to teach cognitive, affective and psychomotor skills encompasses a large number of theories, principles and models. The purpose of the first part of the literature review is to provide an overview of theoretical perspectives and the current evidence-base regarding the design and deployment of instructional video within a practical setting.

2.1.2 Observational Learning

It is widely acknowledged that learning a task from scratch, that is with no prior knowledge, is a daunting undertaking and it is an approach that is rarely used by humans (Schaal, 1997). Learning by observing and imitating others, has long been recognised as one of the most basic, yet powerful learning strategies that is extensively employed by primates (Gog, Paas, Marcus, Ayres, & Sweller, 2008; Hodges, Ong, Larssen, & Lim, 2011; Iacoboni & Dapretto, 2006; Rizzolatti & Craighero, 2004). The prominence of its application is also being applied in other settings as Schaal (1997) states that learning from demonstration (also known as programming by demonstration, imitation learning or teaching by showing) has been employed as a way to program robots in an assembly line, to replace the time-consuming manual programming process.

The mechanisms that underpin the learning strategy (imitation) were discovered by accident when researchers were studying cortical activation in monkeys whilst grasping objects from a box (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992). The researchers then observed the same neural activation when the researchers also grasped the objects from the box leading to the formation of the mechanism for imitation (di Pellegrino et al., 1992). As a result it was assumed that the monkeys were able to understand the goal and could copy it, but they could not encode the details of the movement to enable an accurate replication, an aspect that humans can undertake, enabling imitation (Gog et al., 2008).

Learning by imitation is thought to be due to the role of mirror neurons, which are a class of visuomotor neurons that were initially discovered in a monkeys premotor cortex (Rizzolatti & Craighero, 2004). Whilst there existence can be confirmed in monkeys, most papers can only present evidence (neurophysiological and brain-imaging experiments) of their existence in humans, as it is not possible or indeed ethical to identify individual neurons in a humans brain (Lacoboni & Dapretto, 2006; Oztop, Kawato, & Arbib, 2006; Rizzolatti & Craighero, 2004). Examples of mirror neurons can be found when the role of motor imagery is considered, and the link between the visual and motor cortex is established. A particularly pertinent example, is given by Jeannerod (1994) about internal motor imagery that most of us can fully appreciate. When a sports addict watches a game/match/race on television they mentally perform the appropriate action that is required at a necessary time to score/win a point/win, which can induce changes in heart and respiration rate due to the vivid nature of the imagery (Jeannerod, 1994). The first evidence of this was provided in the 1950's by Gastaut & Bert, (1954) who observed that an EEG (Electroencephalography) rhythm changed not only when the participants performed movements themselves, but also when they purely observed movements. The work by Jeannerod (1994) gave rise to the hypothesis that mirror-neuron activity mediates imitation, which complements another hypothesis formed by Rizzolatti, Fogassi, & Gallese (2001) that states that mirror neurons are the basis of action understanding, that is the ability to make sense of another person's action.

The processes discovered in monkeys underpin the theory of behaviourism, a learning theory that views the mind as a 'black box' where something is observed quantitatively (objectively) and then copied but without thought processes occurring in the mind resulting in limitations to learning (Alonso et al., 2005). The primary principle of behaviourism is that there is a predictable and reliable link between a stimulus and the response it produces (Deubel, 2003). Skill acquisition takes place when the behaviour, strategies, or thoughts of observers are moulded after viewing experts who perform flawless and error free movements, or from peer models that make and correct errors (Wouters et al., 2007). Learners learn by doing, experiencing and participating in trial and error, the outcome of the learning will be observable and measurable (Deubel, 2003). Buccino et al., (2004) conducted a study in an attempt to form a neural basis for the learning of a new motor pattern. The authors used fMRI to analyse naïve participants brains whilst they observed guitar chords being played by an expert, and concluded that the mirror neuron system is at the core of imitation-based learning (Buccino et al., 2004). However the existence of the mirror neuron or mirror system in humans is still un proven as documented by Turella, Pierno, Tubaldi, & Castiello (2009) in their review who state that the evidence is not compelling.

The application of imitation-based learning is integral to the aim of this thesis; its relevance is contextualised further in the sections below that will identify the key challenges and problems presented with the use of instructional videos. The most pertinent of these is how imitation (through visual observation) can be transferred into a motor output, that can then be stored as meaningful knowledge, that can then be applied and adapted in various situations. This area has seen more than a centuries worth of research, yet it stills poses as a crucial functional problem (Brass & Heyes, 2005). However, some research on the mirror-neuron system suggests that video, may be at its most effective when the task to be learned involves human movement, as it will automatically trigger an effortless process of embodied simulation (Van Gog, Paas, Marcus, Ayres, Sweller, 2009). This is similar to the example of internal motor imagery which then primes the learner for the execution of similar actions (Van Gog et al., 2009). There is also additional research that states that observers in their experience acquire more explicit/strategic knowledge as a result of watching, rather than undertaking themselves (Hodges et al., 2011). Despite this research the evidence is not yet conclusive, and the challenge of transforming an instructional video, which is deemed to be a passive media into constructivist learning has been a significant limitation for a prolonged duration (Mayer, 2002). Mayer's

Multimedia Learning Theory (Mayer, 1999) provides the most definitive and well-evidenced instructional design approach (that will be discussed in 2.1.7) and is founded on all three of the main pedagogical theories and the extensively tested and substantiated Cognitive Load Theory (CLT) (Paas, Gog, & Sweller, 2010). Therefore to maximise the instructional effectiveness of multimedia, it is useful to be guided by the research-based theory of how people learn and to connect the applications with evidence and theory (Mayer, 2002).

2.1.3 Learning Theories

There have been a multitude of learning theories proposed since the first by Ivan Pavlov (1849-1936) who became the "Father of Classical Conditioning" (Burton, Moore, & Magliaro, 2003). The most well-known learning theories that have emerged and stood the test of time are behaviourism, cognitivism and constructivism. Before each is discussed and applied to the current study, the mechanisms as to how learners process information shall be addressed. These mechanisms have a profound impact on each learning theory, when applied in an instructional video context. The first of these is the dual-coding theory; this theory paved the way for others including the Cognitive Load Theory and the most relevant to this study, Mayer's Multimedia Learning Theory.

2.1.4 Dual coding theory

Dual-coding theory emerged from experiments that Allan Paivio undertook in the early to mid 1960's (Paivio, 1991). It theorises that people obtain information through the visual and auditory channels and process them in different cognitive subsystems (Clark & Paivio, 1991; Paivio, 1991). If a learner receives a similar amount and scope of information that combines, intertwines and interacts together through both visual and auditory channels then their learning can become considerably more effective (Clark & Paivio, 1991; Ouyang & Stanley, 2014; Paivio, 1991; Schnotz, 2002). When both the visual and textual content complement each other and overlap, they can then serve as mutual retrieval cues, which should then, as a result enhance recall performance (Mayer & Moreno, 1998; Merkt, Weigand, Heier, & Schwan, 2011). Furthermore studies have demonstrated that when the visual and auditory channels are used simultaneously also known as bimodal presentation, the speed of processing and recall can be enhanced (Montali & Lewandowski, 1996).

It was posited that the dual coding theory could provide a useful foundation for a general psychological model of education, additionally it could assist with the efforts at the time to explain certain educational phenomena related to cognitive mechanisms (Clark & Paivio, 1991). An example, that also confirms the theory, was identified by Marcus, Cooper, & Sweller (1996) who found that a picture, can assist in understanding complex text based instructions as both cognitive channels can be used, rather than relying on one.

Despite being one of the earliest cognitive theories, it is still being used as a foundation for many research studies today, Snelson & Elison-Bowers (2009) state that the basis of a good design for educational videos is founded on the dual coding theory. A recent example of this is provided by Chi, Pickrell, & Riedy, (2014) who designed their study to examine the difference between student learning outcomes, with either a video case, or a paper case for dental students.. The theory is also still being examined by many as Berk (2009) identifies an extensive array of theories of learning that have tested dual coding. Each of them have tested verbal information and conclude that the majority of them, support the use of additional media to increase memory, comprehension, understanding and deepen learning (Berk, 2009). Each of these attributes has been extensively tested in a vast number of research studies throughout the past five decades, (far to many to cite, however a large proportion will be cited throughout this work) that have either used dual coding, or a theory that has been based upon dual coding. Dual coding has tremendous advantages due to each of the qualities cited above, however the theory does have limitations and these generally, relate to the capacity of the system and our 'working memory'.

2.1.5 Working Memory

After many years of research into short term memory and its cognitive functioning, Baddeley & Hitch (1974) proposed a three-component model of working memory (Baddeley, 2000). They proposed that it could be divided into three subsystems: the central executive, phonological loop and the visuospatial sketchpad (Baddeley, 1979; Reed, 2006). The central executive is an attentionally-limited control system, it selects and operates the various control processes in conjunction with the phonological loop, which is responsible for verbal and acoustic information and the visuospatial sketchpad which is its visual equivalent (Baddeley, 2003; Baddeley, 1979; Reed, 2006). It was also stated that as the central executive has multiple roles, both the phonological loop and the visuospatial sketchpad can be used to help offload some of the storage demands from the central executive as it is surprisingly limited, in both capacity (7+\- 2 units) and duration (retention for 15-30 seconds) (Baddeley, 1979; Pociask & Morrison, 2004; Young, Van Merrienboer, Durning, & Ten Cate, 2014) which can constrain learning by creating a “bottleneck” in the process (Young et al., 2014). However, Reed (2006) states that a limitation of the initial model by (Baddeley & Hitch, 1974) was that it did not provide a distinct mechanism for the integration of visual and verbal codes, the model in conjunction with the dual coding theory (Clark & Paivio, 1991) were more useful for studying independent contributions of each, rather than the integration of the information. Despite the initial model having a central executive that was intended to store and integrate the visual and verbal information, it was more concerned with attention and omitted this key factor that underlined both the working memory model and the dual coding theory (Reed, 2006). As a result of this a fourth subsystem has been proposed, the episodic buffer (Baddeley, 2001). This is a limited-capacity short-term storage system that works in conjunction with the central executive and stores chunks, or blocks of information which can then be accessed consciously (Baddeley, 2000, 2001, 2003).

Working memory is part of the process for memory storage, one of the first models was proposed by Atkinson & Shiffrin (1965) who derived a model with three memory states. This included a very short-lived memory system that they referred to as the sensory buffer; a temporary memory state called the rehearsal buffer and a long-term storage state from which items can be stored and recalled (Atkinson, Brelsford, & Shiffrin, 1967). The first aspect is addressed in the dual coding theory and working memory model, as the sensory information is obtained through the visuospatial and phonological systems. The central executive and the episodic buffer, in the working memory model address the temporary memory state before items are codified, and stored in the long-term memory system. The process of integrating and codifying the sensory information has received a large amount of interest, as the working memory has distinct limitations. The most notable of which is capacity, this is particularly relevant to dynamic visualisations such as instructional videos. When viewing instructional videos students are required to process the information in front of them, whilst remembering previously presented information, relate and integrate the information together into chunks in the episodic buffer, so that they can comprehend the visualisation (Ayres & Sweller, 2005). This is also referred to as the split-attention effect, which will be revisited in 2.1.8. The activities can create a high working memory load, which could compromise the comprehension of the subject matter due to the continuous flow of information, which is being inputted. A limited capacity or limited duration working memory, results in there being very little time or capacity for the information to be integrated and understood (Gog et al., 2008). There is also the possibility that the system could become overloaded when the learners working memory is exceeded, this also undermines the students capacity to process information effectively and therefore learn (Iserbyt, Liesbet, Jan, & Daniel, 2012; Moreno & Mayer, 1999a; Young et al., 2014). Lusk et al., (2009) discovered that students with a high working memory capacity where able to generate more valid interpretations, and recall a greater amount than those with a low working memory. The authors also identified that those with low working memory, would benefit from segmentation of the dynamic visualisation for recall and transfer of the information being addressed (Lusk et al., 2009).

As a result working memory is an integral step in the learning process, it helped to foster the development of theories and models that encompassed or have attempted to encompass the entire process. One such model is the social cognitive model of sequential skill acquisition summarised by Wouters et al., (2007). The theory describes how learners initially start with observing a model, but then start practicing and gradually learning how to self-regulate their own performance, this is accompanied by different cognitive processes such as problem solving and cognitive behaviour modification (Wouters et al., 2007). Whilst this is an encouraging approach, the literature and evidence-base for this theory is significantly lacking. Another promising instructional design model has emerged however from the cognitivist approach, the Cognitive Load Theory (CLT) (Ayres, Marcus, Chan, & Qian, 2009).

2.1.6 Cognitive Load Theory

Cognitive load theory (CLT) was first described by John Sweller in the 1980's (Sweller, 1988) and represents a general theory of cognitive learning. CLT is receiving increasing recognition when used in a medical education setting (Ayres, 2015; Young et al., 2014) which underlies the use of the theory within the context of this study. CLT was designed to provide guidelines that could assist in the presentation of information to a learner, it is designed to encourage mental activity that optimises intellectual performance (Sweller, Merrienboer, & Paas, 1998). CLT argues that because of the way human cognitive architecture is organised, learning by observing and/or imitating what other people do, say, or write, is a much more effective and efficient way of acquiring knowledge than trying to devise this knowledge by ourselves (Merriënboer & Sweller, 2005; Paas, Renkl, & Sweller, 2003; Sweller et al., 1998; Sweller, 1988). This fundamental premise is congruent with previous literature already cited (see sections 2.1.2 and 2.1.3) regarding the fundamental methods of learning.

The CLT has become an extremely successful experimental and instructional theory throughout the last 30 years, despite coming under scrutiny by many researchers (Ayres & Paas, 2012). CLT utilises the knowledge of the working memory and the long-term memory, to generate instructional procedures that can seem counter intuitive to some (Mayer, 2011). The fundamental principle of CLT is that instructional messages, must be designed to keep the cognitive load of learners to a minimum during the learning process, to be effective (Samaras et al., 2006). CLT builds upon Paivio's dual coding theory, it acknowledges that information can be acquired in differing forms via the auditory and visual channels. Some media may recruit only one channel but (Mayer, 2002) theorises that information is conveyed and understood further, when both channels are used synchronously. The units of information are then processed by the working memory, that as outlined previously (see section 2.1.5) is limited in terms of the amount of information it can simultaneously integrate, and the duration in which it can retain these (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). The working memory aims to form a mental schema in combination with retrieved information from the long-term memory, this can then be transferred into the unlimited long-term memory within the episodic buffer (Baddeley, 2000; Merriënboer & Sweller, 2005). The limitations of the human working memory have been acknowledged as a critical factor that needs to be considered when instructional materials are designed (Ayres & Paas, 2012; Samaras et al., 2006). This is to ensure that too much cognitive load is not placed upon the student which could compromise learning (Ayres & Paas, 2012; Samaras et al., 2006).

There are three distinct forms of cognitive load that form within the working memory; these are intrinsic, extraneous and germane (Schnotz & Kürschner, 2007).

Intrinsic Cognitive Load

Intrinsic cognitive load is deemed to be essential to the task, it relates to the intrinsic intellectual complexity of the knowledge that is being acquired, but it is not concerned with the methods of acquisition (Jong, 2009; Mayer, 2011; Vogel-Walcutt, Gebrim, Bowers, Carper, & Nicholson, 2011; Young et al., 2014). It has been extensively documented that intrinsic cognitive load cannot be modified by the instructional strategy, it is a function of the nature of the materials being learned, the expertise and motivation of the learner (Jong, 2009; Merriënboer & Sweller, 2005; Sweller et al., 1998; Verhoeven, Schnotz, & Paas, 2009). However, Pollock, Chandler, & Sweller, (2002) suggested a mechanism that was beneficial to learners; presenting information in isolated elements, that could be processed iteratively and more effectively when all of the elements where presented together. Hollender, Hofmann, Deneke, & Schmitz (2010) present this information as a principle to adjust intrinsic load, but do not cite any further studies or information to supplement this principle. This is despite a study by van Merriënboer, Kester, & Paas (2006) citing a similar approach to Pollock et al., (2002) where they referred to the part-whole approach, that initially simplifies the task and then adds more and more elements sequentially. Another study by Kalyuga, Ayres, Chandler, & Sweller (2003) cites a method of scaffolding the information or adapting the level of instructional guidance to each learners appropriate expertise level. Therefore, whilst there maybe processes whereby the intrinsic load could be reduced, they require extensive planning and a greater volume of materials and resources, which may make the process unfeasible in most situations. Furthermore, more recent studies have explicitly stated that intrinsic load cannot be manipulated (Hasler, Kersten, & Sweller, 2007; Wouters, Paas, & van Merriënboer, 2008). As a result the role of intrinsic load is one that requires further research, nonetheless intrinsic cognitive load is an important component to consider, as a very high intrinsic cognitive load from difficult material, could exceed the capacity of the working memory itself because it is high in element interactivity (Paas et al., 2010; Pollock et al., 2002). However a reduction in intrinsic load decreases total cognitive load, allowing learners to use the freed capacity to form advanced schemas (Paas et al., 2003).

Extraneous Cognitive Load

Extraneous cognitive load is inflicted by information that is deemed irrelevant to the task and places additional, unnecessary load on the learners working memory (Austin, 2009; Ayres, 2015; Jong, 2009; Paas, Tuovinen, et al., 2003; Sorden, 2005; Sweller & Chandler, 1991; Sweller, 1994, 2011; Wong, Leahy, Marcus, & Sweller, 2012). Extraneous load is often thought to be the consequence of poorly designed instructional resources, which is within the control of the instructional designer (Schnotz & Kürschner, 2007; Vogel-Walcutt et al., 2011; Wong et al., 2012). As a result the vast majority of research into cognitive load theory, until the second half of the 1990's, has been into techniques of reducing extraneous cognitive load from instructional resources (Schnotz & Kürschner, 2007; Wong et al., 2012). The classical view of CLT, states that learning materials should keep the learners extraneous cognitive load to a minimum during the learning process, and a range of methodologies are available to the instructional designer to accomplish this (Verhoeven et al., 2009).

One of the main criticisms of CLT is that it generally only focuses on the acquisition of procedural knowledge, in a well-structured and organised environment, which may not always be possible (van Merriënboer & Sweller 2005). Van Gog, Paas, & van Merriënboer (2006) undertook a challenging study examining the use of worked examples for complex problem solving. This inherently applies a large volume of intrinsic cognitive load on the learner due to the unorganised, ill-structured and unpredictable nature of the task, which when combined with extraneous load could impair the learners ability to solve the problem (van Gog et al., 2006). The load is also known to be considerably larger in novices than experts, as the latter have previous mental schemas and solutions from experience that they can draw upon to assist in the problem solving process. The authors found that using worked examples as guides and worked-out solutions largely confirmed their hypothesis. They were able to prove that they would lead to higher near and far transfer test performance, less mental effort and a decreased time requirement from the learners (van Gog et al., 2006) thus addressing the main criticism raised. Paas, Camp, & Rikers (2001) present an alternate method for reducing extraneous cognitive load and detail the use of goal-free problems on maze learning between age groups. The author's report that instruction based on CLT, can reduce extraneous load within this context and is able to compensate for age related cognitive declines. Despite the above examples Kirschner, Ayres, & Chandler (2011) state that the goal of research on cognitive load should not necessarily be about reducing cognitive load during learning, but finding methods to optimise learning. As a result the use of the methods above can be integrated into the three aims that Kirschner et al., (2011) and Vogel-Walcutt et al., (2011) outline for cognitive load research. These are for instructional design to keep extraneous load to a minimum, for cognitive load and learning to be correlated optimally and for any load that is incurred to by an instructional design is germane in nature (Kirschner et al., 2011 and Vogel-Walcutt et al., 2011).

Germane Cognitive Load

Germane cognitive load (also known as effective cognitive load) enhances learning, it results in task resources being devoted to schema acquisition and automation when resources are available in the working memory (Schnotz & Kürschner, 2007; Sorden, 2005; Sweller et al., 1998). Whilst this increases cognitive load it also contributes to, rather than interferes with learning (Artino, 2008; Schnotz & Kürschner, 2007; Sorden, 2005; Sweller et al., 1998).

There has been some debate around whether Germane load is sufficiently distinct and thus should really be referred to as a constituent component of intrinsic load (Kalyuga, 2011; Young et al., 2014). Germane cognitive load was added after the CLT was initially derived (Sweller et al., 1998) as it became apparent that cognitive load did not always compromise learning. It was an established aspect that both intrinsic and extraneous load needed to be reduced, in order to facilitate effective learning and the theory therefore did not fully account for how learning occurred (Kalyuga, 2011). Therefore whilst intrinsic and germane cognitive load are responsible for the generation of mental schemas enabling humans to learn, remember and process information (Samaras et al., 2006) extraneous load is non-conducive to the generation of the mental schema’s and results from poor instructional design (Paas et al., 2003).

As a result, it was postulated that the goal of learning is to balance total cognitive load by maximising the amount of germane and intrinsic load, whilst reducing or eradicating extraneous load (Paas et al., 2004). All three types of cognitive load are additive (Galy, Cariou, & Mélan, 2011; Jong, 2009; Moreno, 2010; Schnotz & Kürschner, 2007; Merriënboer & Sweller, 2010) so there is in theory a direct positive relationship between decreasing extraneous load, increasing intrinsic and germane cognitive load, within the subject area of e-learning and specifically the utilisation of videos as a teaching medium (Chandler, 2004; Dijksterhuis et al., 2009; Hollender, Hofmann, Deneke, & Schmitz, 2010a; Merriënboer & Ayres, 2005; Sorden, 2005; Triola et al., 2012; Wong et al., 2012).

As a result of this Richard E. Mayer (with colleagues) proposed the Multimedia Learning Theory (Mayer, 1999) that provides the most definitive and well-evidenced instructional design approach to enhance cognitive processing, based upon the principles that have been outlined up until this point.

2.1.7 Mayer's Cognitive Theory of Multimedia Learning

Multimedia learning refers to two or more sources of information; this has been extensively used for many decades in classrooms. For example when a teacher draws an item onto a blackboard or overhead projector whilst simultaneously talking about the item, they are employing a multimedia approach. Multimedia as a result purely refers to multiple sources of information, presenting information in this way is seen to be the most effective form of aiding learning (Mayer, 2014). Typically when you state the term multimedia to anyone they will depict a high-tech methodology in their minds. This however is merely a modern application of the teacher with a blackboard; the principle between that and an instructional video remains the same. Therefore the process by which learners build mental representations from words and pictures, is the focus of Mayer's cognitive theory of multimedia learning that seeks to foster meaningful rather than rote learning (Mayer, 2009; Mayer, 2014).

Mayer provides a succinct example of the rationale for multimedia learning from earlier studies using a bicycle pump (Mayer, 2014). For the first part of the experiment, the learners listen to a description of how the bicycle pump works, they are then asked to write down an explanation to test retention and to answer problem-solving questions designed to test transfer. It was identified that most students retained some of the words, but had difficulty applying the material to the problem-solving questions (Mayer & Anderson, 1992). In the second part learners are shown an animation of the internal workings of a bicycle pump, the same tests are conducted, however this time (Mayer & Anderson, 1992) found that most learners did not perform well on either test. In the final part of their example, learners are shown the animation with verbal descriptions as the example works through the different steps of how the bicycle pump works. The authors found that students performed well on both tests, particularly problem-solving (Mayer & Anderson, 1992). The fundamentals of the study (termed the multimedia principle) were replicated in nine out of nine studies that Mayer and colleagues have conducted, and when analysed yielded an effect size of 1.50 (Mayer, 2009).

Mayer’s cognitive theory of multimedia learning is outlined pictorially below (Figure 1) and illustrates how the information presented, is obtained via two channels (visually and audibly) in line with the dual-coding theory (Clark & Paivio, 1991). The information is then coded and transferred into the working memory via the visuospatial and phonological systems. The central executive and episodic buffer, in conjunction with previous knowledge from the long-term memory form new schema’s that can then be stored for retention and transfer tests (Baddeley, 2001; 2003).

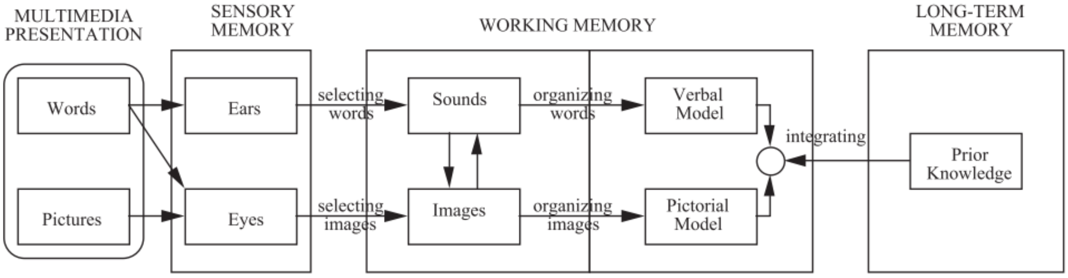

Figure 1 Mayer's Cognitive Theory of Multimedia Learning, reproduced from Mayer (2008)

In order for the desired information to be presented effectively, appropriate guidance regarding the design of instructional videos and the science of instruction should be followed. As a result Mayer proposed a range of guidelines that he termed the *"Basic Principles of Multimedia Learning"* which have varied in number throughout the years, but have been extensively tested with experimental studies by himself and colleagues (for example: Mayer, 1999; 2005; 2008; 2009; 2011 and 2014). The most recent collection of basic principles can be found in *The Cambridge Handbook of Multimedia Learning* (Mayer, 2014) alongside a set of advanced principles of multimedia learning. The evidence-based and theoretically grounded principles of multimedia instructional design, are divided into three subsections; those that reduce extraneous processing, those that manage essential processing and those that foster generative processing.

A fundamental principle of the science of instruction is that the use of instructional practice is grounded on empirical evidence, to that end (Mayer, 2008, 2011 and 2014) derived effect size statistics from controlled experiments based upon and interpreted by the work of Cohen (1988). Effect sizes can be interrpted at different levels however Cohen (1988) states that an effect size of 0.2 is considered small, 0.5 is moderate and anything greater than 0.8 is considered a large effect (see effect size section 3.11 for further details).

2.1.8 Instructional Principles to Reduce Extraneous Processing

Coherence Principle

The coherence principle states that multimedia materials are more effective and understood better, when they include a lesser number of extraneous words and sounds (Kirschner, 2002; Mayer & Fiorella, 2014; Reed, 2006). This is proven to have a large median effect of 0.86 (23 out of 23 studies) in controlled experiments and by concise materials leading to better learning rather than embellished materials (Ely, Pullen, Kennedy, Hirsch, & Williams, 2014; Kirschner, 2002; Lehman, Schraw, McCrudden, & Hartley, 2007; Mayer & Moreno, 2002; Mayer, 1999; Mayer & Fiorella, 2014; Mayer, 2008; Reed, 2006). The coherence principle is also known as the seductive detail effect that implies an unusual connotation to the use of embellished materials. Rey, (2012) conducted a meta-analysis of the seductive detail effect that encompassed various forms of embellishment (written passages added to instructional resources, illustrations to multimedia messages, spoken materials during lectures, video clips, sounds and background music during multimedia messages) that may well be deemed interesting but are irrelevant (Mayer, 2005). The initial results of the meta-analysis identify 39 studies encompassing four main seductive areas with contrasting results as they contain Cohen's *d* effect sizes ranging from 2.29 to -0.93 (Rey, 2012). In conclusion the findings of the meta-analysis support the existence of a seductive detail effect for retention (small to medium effect size) and transfer (medium effect size). The theoretical causes of the seductive effect remain unclear and can be attributed either to overloading the working memory, distracting a learners attention, interfering with schema formation or by disrupting the coherence of the learning material (Rey, 2012). The conclusions of the meta-analysis are further supported in the chapter by Mayer & Fiorella (2014) however they only identify 23 studies, as opposed to the 39 studies of the meta-analysis. This may be due to the fact that they focused on multimedia presentations only, but the methodology is not completely clear and whilst Rey, (2012) identifies multiple studies with negative effect sizes, only one (with a small effect size of -0.17) is identified by Mayer & Fiorella (2014). Therefore if it were not for the congruence between findings and recommendations from both of these sources, the efficacy and relevance of the coherence principle could be questioned. It is also worth noting that the vast majority of studies examining the principle use college or undergraduate students as participants, therefore the role of the principle is of direct relevance to this study. Whether the effect can be extrapolated to other populations remains unclear, as the demographics of undergraduate students can be quite broad. However, this is only an aspect that can be analysed if differing age groups of undergraduate students are compared, an aspect that is beyond the scope of this study.

Signaling Principle

The signaling principle (also known as the attention-guiding principle) aims to reduce extraneous load by providing cues for the learners, identifying and directing the learners attention to the importance and organisation of the pertinent information (Castro-Alonso, Ayres, & Paas, 2014; Croft, 2010; Mayer, 2011). The principle has a median effect size of 0.41 (supported in 24 out of 28 studies) which, whilst positive is the lowest of all the fourteen principles (Mayer, 2014). Cues can either be provided with additional elements such as arrows, frames, symbols and the like which may prove to increase extraneous cognitive load and defeat the purpose of the cueing (Castro-Alonso et al., 2014). They can also be provided without additional elements, such as zooming, panning, colours or patterns and the like which would be the favoured mechanism (Castro-Alonso et al., 2014).

One of the most logical methods to quantify the principle is through the use of eye tracking. In a commentary for a special issue relating to eye tracking and the signaling principle amongst others Mayer, (2010) identifies that prior research has sometimes failed to empirically evidence the use of visual cues such as arrows in animation. The special issue does however contain two studies that identify differing versions of visual cueing, these have proven to be effective in improving learning outcomes or cognitive processing whilst learning (Boucheix & Lowe, 2010; de Koning, Tabbers, Rikers, & Paas, 2010). The first study utilised a technique called “spreading-colour cues” in which ribbons of colour overlaid the most relevant aspects of the graphic, this was synchronised with the propagation of the content being depicted. The second adjusted the luminance of the surrounding aspects of the graphic, so that the key information is, in essence highlighted to the learner whilst being able to be contextualised as the remaining content is still visible. Whilst both of these techniques proved to have a beneficial effect on learners, the practicality of either of their implementation, or any of the other techniques for that matter would significantly enhance the production time of any instructional videos. Although they can be seen to be effective, they would need to be precisely controlled and could only be employed within an animation or video whereby the content is of sufficient volume so that it can be signaled, cued or highlighted effectively. As such the use of the signaling principle for general practical skills that involve movements at multiple joints needs to be questioned, as it would be very easy for the designers of instructional materials to violate the coherence principle if any form of signaling is over utilised as this could distract or confuse the learner. As Mayer, (2014) identifies it only has a small effect size, the value of this principle will require consideration as there appears to be a limited evidence base that is applicable to the context of the current study.

Redundancy Principle

The redundancy principle demonstrates that when pictorial animation and narration are used together, higher rates of retention and transfer performance are achieved (Mayer, 2014b). This is supported with a large median effect size of 0.86 from 16 out of 16 studies when compared to information that is presented as pictorial animation, narration and text (Austin, 2009; Kirschner, 2002; Mayer, 2014; Samaras et al., 2006). When pictorial animation, narration and text are used to present information the visual pathway becomes overloaded, but this is only when they are presented simultaneously, successive presentation did not result in an overload or affect learners performance (Kalyuga, Chandler, & Sweller, 1999; Mayer & Moreno, 2002). Not all research is in agreement with this viewpoint as Montali & Lewandowski (1996) found that when information was presented with on-screen text and accompanying narration, benefits where identified for less skilled readers. The study had numerous limitations; for example the small sample size for less skilled readers (eighteen), the fact that the sample had internal variability between the factors, this affected the reading of the less skilled readers and that only three reading sessions were used. Despite these limitations the findings should be considered by an instructional designer, as violating the redundancy principle could (in practice) be beneficial, depending on the abilities of the learners for whom the resources are intended.

If the learning resource is predominantly text based, Kalyuga et al., (2003) drawing upon the findings of multiple authors support a similar advantage for low level readers as Montali & Lewandowski, (1996). Kalyuga et al., (2003) state that with the use of certain enhancements, in addition to written text they will perform better. The authors go further by stating that higher-level readers were disadvantaged when using the enhanced texts, as the additional material required additional cognitive processing even though it was deemed redundant by them.

Therefore the benefits of the redundancy principle are evident, it is recommended when the information is predominantly pictorially represented with accompanying narration. There may be a case for adding additional text for low level learners if it does not disadvantage the higher level learners, but this is a judgment call for the instructional designer and it should be informed by the target populations capabilities. When the resource is text based similar complications arise that again should be settled by considering the spectrum of abilities of the learners. In both cases if the population is mixed it would appear that omitting the additional material would provide learning opportunities for all levels, given the stronger evidence base relating to its omission rather than inclusion. Whether it is the case or not however is purely speculative and in a longitudinal study, such as the current one where higher level learning is targeted, compliance to the redundancy principle would appear to be advantageous if not necessary once some prior knowledge has been achieved to ensure continued success. If the resource violates the redundancy principle and uses pictorial animation, narration and on-screen text it may limit the achievement of all of the learners.

Split-Attention Principle

The Split-attention effect identified in the late 1980's is a robust, easily demonstrated effect that leads to the split-attention principle (Ayres & Sweller, 2014). This occurs when a learner has to split their attention between multiple sources of mutually referring information, before they can cognitively integrate them (Ayres & Sweller, 2014; Mayer, 2014; Sweller, 2003). Interestingly this is the only principle that Mayer does not provide an effect size for, however they cite Ginns, (2006) who conducted a meta-analysis and established that the split-attention effect had a effect size of 0.85, which whilst large should be interpreted with caution as it does not include a body of evidence from recent years. The instructional material can only be understood when both the text and picture are presented, when only one or the other is provided the material is unintelligible (Sweller, 2003).

However this is not always the case, exceptions are found when the text merely describes the picture and does not contribute any essential additional information, in this case the text is redundant (see redundancy principle) (Ayres & Sweller, 2014). Mayer (1989) applied the principle to a computational setting and found that when pictures were presented without labels, and when labels were presented without pictures, was inferior to providing both to learners. Moreno & Mayer (1999b) undertook further work to demonstrate the subtly of the split attention effect, they found that placing the text either in close proximity to, or underneath the picture resulted in a significant difference between learners retention, transfer and matching performance. As a result it is advised that materials are appropriately constructed, this is so that disparate sources of information are physically and temporally integrated, which avoids the necessity for the learner to mentally integrate the information sources (Ayres & Sweller, 2014). The research undertaken for the split-attention effect led to the development of the spatial and temporal contiguity principles. Whilst there is an overlap between the three principles they are presented independently in Mayer (2014). Ginns, (2006) however states that the split-attention effect is also known as the Spatial Contiguity effect, which if not sufficiently differentiated by authors could create confusion. The split-attention principle shall be used within the context of this study as a general overview of learners dividing their attention, whereas the spatial contiguity principle has specific boundaries as outlined below.

Spatial Contiguity Principle

The spatial contiguity principle indicates that text should be positioned next to the respective part of the pictorial animation, this is to reduce visual scanning and therefore reduce extraneous overload (Doolittle, 2001; Issa et al., 2013; Mayer, 2011; Moreno & Mayer, 1999b; Reed, 2006; Schnotz & Kürschner, 2007; van Bruggen, 2005). The theoretical rationale is that the corresponding words and pictures, should be in the working memory at the same time so that constructive links between the materials can be made (Eitel, Scheiter, Schüler, Nyström, & Holmqvist, 2013; Ginns, 2006; Scheiter & van Gog, 2009). This theoretical standpoint has been subsequently proven to have a large effect size of 1.10 from 22 out of 22 experimental studies, in particular when applied to complex learning materials were substantial learning gains can be found with the principles implementation (Ginns, 2006; Mayer, 2014). Despite the large effect size the finding is, unfortunately not directly transferable to the context of the current study as Iserbyt et al., (2012) undertook a study investigating the Spatial Contiguity and Multimedia Principles on participants ability to learn Basic Life Support (BLS). BLS is a psychomotor skill which was taught from task cards, this differs considerably from those previously investigated that focus on how lightning or bicycle pumps work, those that informed the effect size statistic. BLS skills were assessed practically via an instrumented resuscitation simulator, the scores from dual observers were analysed with exceedingly high levels of inter-observer reliability (0.97 and 0.98). The authors found that the spatial contiguity group (integrated text) did not achieve greater BLS skills than those who received the text and picture or text only group. They did however, identify that when the learners had to verbalise responses as opposed to practically demonstrate them, those who received the integrated text condition outperformed the other groups supporting the spatial contiguity principle. The findings from Peter et al., (2012) question the validity of the spatial contiguity and multimedia principles within this context. Whilst it is not directly transferable to instructional videos, task cards when sequenced could provide a resource not dissimilar to an instructional video. However, it is worth noting the scarcity of research in this area and as such the findings of a single study do not warrant the dismissal of the principle, merely that the principle is applied with deliberate forethought in a psychomotor context.

Temporal Contiguity Principle

The Temporal contiguity principle states that individuals learn more effectively, when pictorial animation and narration are presented simultaneously and preferably synchronised, as opposed to following each other (Mayer, 1999; Mayer & Anderson, 1992; Reed, 2006). The relevance and context of the principle for the current study was identified by Baggett (1984). They conducted a study where they provided students with a video which included a voiceover that was either synchronous with the visual component, 7s, 14s or 21s before or after the visual component creating seven distinct conditions. The results of their study concluded that those who used the resource with the voiceover being synchronous or 7s after the visual component performed at a superior level to the other five conditions. This was confirmed in the study via recall tests exemplifying the importance of timing between visual and audible materials. Mayer, (2014) identifies nine experimental studies that yield a median effect size of 1.22 demonstrating that it is one of the most practical principles to reduce extraneous processing. The finding is in line with the theoretical foundation regarding constructive links between materials within the working memory (Eitel et al., 2013; Ginns, 2006; Mayer, 2008; Scheiter & van Gog, 2009).

2.1.9 Instructional Principles that Manage Essential Processing

Segmenting Principle

The segmenting principle states that learning is more effective (effect size of 0.79 from 10 out of 10 studies) when a multimedia lesson is presented in bite-size segments, that learners can view at their own pace rather than as a continuous item which could overload their working memory (Castro-Alonso et al., 2014; Croft, 2010; Issa et al., 2011; Mayer, 2011, 2014). The principle aims to give the learners time and capacity to organise and integrate the selected information, before moving on to the next segment (Mayer & Moreno, 2003). The application of the principle is not universal however; as it appears that it may only be of use to learners with low prior knowledge as established by Spanjers, Wouters, Van Gog, & Van Merriënboer (2011) who undertook a segmenting study with animations. The authors used continuous, animated worked-out examples that were approximately 2 minutes in duration and were presented one after another. These were compared to segmented, worked-out animated examples which were divided (by three content related experts) into appropriate segments, with two-second pauses were the screen darkened. The results of their study demonstrated that the segmented videos, for learners with low levels of prior knowledge attained an equal level of performance with less cognitive load, an aspect that disappeared for those learners with higher levels of prior knowledge (Spanjers et al., 2011). The role of segmenting has been built upon further by Cheon, Crooks, & Chung (2014) who manipulated the role of the pause. In previous studies the pause has been passive i.e. a darkened screen, which acted as the control versus an active pause whereby learners were presented with embedded questions during the pauses. In order to counter the effect of the questions, those in the passive pause group were presented with the questions at the end of the animation. The results from their study showed that those in the active pause group, significantly outperformed those in the passive pause group on recall and transfer tests (Cheon et al., 2014). The authors do identify a key limitation of their paper as the pause (within both conditions it appears) was not controlled, therefore there is the potential for this to be the main factor for the differences identified in line with the segmenting principle, or a combination of the two. Either way, the role of an active pause appears to be substantially beneficial regardless of the mechanism by which it is achieved. The predominant limitation of their implementation is the extra production time, however some recent software available online, is addressing that aspect and bringing the use of active pauses to fruition (Educanon, 2014).

Pretraining Principle

The pretraining principle states that students can learn better when they are familiar with the names, locations and characteristics of the key concepts presented in the instructional resource (Mayer, 2005; 2008; 2011; 2014). The aim of the pretraining principle is to provide students with relevant prior knowledge. This is so that generative processing can be accomplished, as it indicates which pieces of prior knowledge are important to integrate with the new information, this can also reduce the complexity of the task (Ibrahim, 2012; Moreno & Mayer, 2007; Wouters, Paas, & van Merrienboer, 2008). The pretraining principle has been identified to have a large effect size of 0.75 from 13 out of 16 studies reviewed (Mayer, 2014). There is also limited evidence that the principle is most effective when learners have to conceptualise a model or system, for low-knowledge learners or for complex presentations that are delivered at a fast pace (Mayer, 2011; 2014).

Modality Principle

The modality principle postulates that combing pictorial animation and spoken narration, provides better retention and transfer performance than a combination of written text and pictorial animation (Alonso et al., 2008; Austin, 2009; Kirschner, 2002; Mayer, Heiser, & Lonn, 2001; Mayer & Moreno, 1998; Mayer, 2008; Samaras et al., 2006; Schüler, Scheiter, Rummer, & Gerjets, 2011; Wouters, Paas, & van Merrienboer, 2008). This principle has been shown to have a large effect of 0.76 from 53 out of 61 studies identified (Mayer, 2014). This principle has its foundations in the limited capacity component of the working memory model, and identifies that using pictorial animation and spoken narration, is a more effective use of the visuo-spatial sketchpad and the phonological loop (Austin, 2009; Schmidt-Weigand, Kohnert, & Glowalla, 2010; Schüler et al., 2011). This theoretical underpinning is supported by Kalyuga, Chandler, & Sweller, (2011) who found that the working memory becomes overloaded, when the visual system is used to process text and images and can be termed the split attention effect. The efficacy of the principle is predominantly found when the learner does not have any control over the pace of the material, there are minimal/no opportunities to take notes and there are no pauses (active or passive). Jong, (2009) identified a range of recent studies that demonstrated a lack of support for the principle when the learners could dictate the pacing. The pacing provided opportunities for them to form schema's before the next segment of information is presented (Jong, 2009). This viewpoint is also supported by Cheon et al., (2014), who stated that whilst the modality and segmenting principles both aim to utilise the learners cognitive resources as effectively as possible, the segmenting principle may in fact neutralise the modality principle due to the reasoning cited above.

2.1.10 Instructional Principles to Foster Generative Processing

It has been stated by Mayer, Fennell, Farmer, & Campbell, (2004) that there are two mechanisms to foster generative processing from multimedia resources. Those that reduce cognitive load, as outlined above and those that increase the learner's motivational commitment for active cognitive processing (Mayer et al., 2004). Whilst the vast majority of research has been undertaken on the former of the two, the role of social considerations is an area that warrants further research (Mayer, 2014). This is an area where advances can be made with effective planning and implementation to engage the learner in a constructive manner.

Multimedia Principle

The multimedia principle states that students learn better from words and graphics or pictures as opposed to words alone (Iserbyt et al., 2012; Issa et al., 2011; Mayer, 2005; Samaras et al., 2006; van Bruggen, 2005). Iserbyt et al., (2012) state that this approach enables learners to build two mental representations (verbal and visual) and therefore build connections between them to enhance learning. The enhancement of the learning has been identified by Mayer, (2008) as having an effect size of 1.39. However Butcher, (2014) does not provide an overall effect size when discussing the principle, instead they cite multiple positive effect sizes to support the use of the multimedia principle. This is due to the various forms multimedia could take, as the principle is not associated with a single visual mechanism as the instructor could provide static pictures, dynamic pictures, animation, video or a combination of the above.

Issa et al., (2011 and 2013) conducted studies to test the multimedia principle with medical students when being taught theory. The outcome of both studies revealed a statistically significant effective when compared to a control or traditional group for theoretical knowledge. A study by Smith et al., (2011) contributes to the work of (Issa et al., 2011, 2013) by undertaking a crossover study comparing the differences between a live demonstration and instructional multimedia, when teaching theory to physical therapy students. The study concluded that there was no difference between the two methodologies, identifying that multimedia resources are an effective teaching mechanism. The authors elaborate by stating that the instructional multimedia may also improve efficiency and in theory, promote higher level processing during supervised practice sessions (Smith et al., 2011). Whilst the study identifies some key attributes and benefits the method of analysis for every relevant aspect assessed was subjective via questionnaires. This study identifies theoretical and cognitive benefits for the use of the multimedia principle within a psychomotor domain, but it fails to apply them in a practical setting, as opposed to the aim of the current study.

Upon further reading of the cited sources in this section, it appears that the efficacy of the principle is inextricably linked to a large proportion of the other principles listed here. As a result it seems as though there is little need for this principle to be referred to as an isolated principle, 'multimedia' has been used in an educational context for a considerable number of decades, if not centuries up until now with appropriate success. Therefore this principle is automatically complied with anytime visual and verbal materials are presented. As a result instructional designers should focus further attention on those that reduce extraneous processing.

Personalisation Principle

The personalisation principle states that the design effects are stronger when learners hear narration that is formed in a conversational style, rather than a formal style, that is including ‘I’ and ‘you’ for example or embedding the content within an overarching narrative like a story (Evans, 2008; Gog & Rummel, 2010; Mayer, 2003; Mayer et al., 2004; Mayer & Moreno, 2002; Moreno & Mayer, 2007; Paas & Sweller, 2011; Sorden, 2005). The effect of this principle is large as (Mayer, 2014) identified an effect size of 0.79 in 14 out of 17 studies. There is a distinct paucity of studies that have assessed the role of the personalisation principle in a psychomotor context; this is due to the inherent limitations that a study of this nature would present. The personalisation principle aims to engage the learner by encouraging them to relate the material they are processing with pre-existing content. Although this may appear straight forward enough to apply, Mayer et al., (2004) advocate caution when doing so, as it could create seductive distractions that are irrelevant, therefore the personalisation should be subtle and focused.

Voice Principle

The voice principle enables the learner to create a feeling of social partnership with the computer when they hear a human voice (especially an accepted human voice), this encourages them to put in a greater amount of effort to comprehend the message i.e. they participate in generative processing (Castro-alonso, Ayres, & Paas, 2014; Gog & Rummel, 2010; Mayer, 2009; Mayer, 2011; Paas & Sweller, 2011). The voice principle is extended further by Castro-alonso et al., (2014) who state that female speakers are not only deemed to be more appealing than males, but the feminine voice elicits greater problem-solving performance. However there is no distinct cognitive explanation, as it should take the same amount of processing for a male or female voice. Mayer, (2011) also identifies a potential boundary condition that may impact upon the learner, the narrator's accent, as the learner may not like the accent, which could negatively affect their cognitive processing. Despite these potential pitfalls the voice principle has a large effect size of 0.74 calculated from 5 out of 6 studies (Mayer, 2014).

Embodiment Principle

The embodiment principle (also none as the video modeling example) suggests that learning is more effective, (effect size of 0.36 from 11 out of 11 studies) when an on-screen agent, (typically animated, although it could be a puppet or similar) displays humanlike gestures and movements (Clark & Feldon, 2005; Mayer, 2014; Mayer, 2009; Van Gog et al., 2014; Wang & Crooks, 2015; Wouters, Paas, & van Merrienboer, 2008). The principle could be undone if the voice principle is violated, as there is a tendency for animated agents to have a machine or synthetic voice (Mayer, 2009). The embodiment principle exemplifies the aspect mentioned earlier as there is a distinct lack of specific research into the principle. Studies that have been undertaken do not generally refer to it as the embodiment principle; therefore any researchers in this area have to draw comparisons to studies that use similar, but different methodologies. One recent study by Wang & Crooks,

(2015) examined the effect of combining the personalisation and embodiment principles upon learner delayed retention. Their study demonstrated that learners, who received both, outperformed those who only received the personalisation design. However learners from the latter group preferred the instructional design. Therefore, there is a distinct divide, despite the calculated effect size from Mayer, (2014) as to the efficacy of the principle. For some it may be generative, but for others the embodiment may prove to be a distraction and violate the coherence and split-attention principles.

Image Principle

The image principle states that the visual presence of the narrator is not strictly necessary for learning to be effective (Gog & Rummel, 2010; Mayer, 2009, 2011 and 2014). This principle as a result may well be of great relief to many instructional designers. The image principle is stated to have an effect size of 0.20 (from 7 out of 14 studies) which according to (Mayer, 2011) is too small to be of any educational significance (Mayer, 2014). Gog & Rummel, (2010) stated that the presence of the narrator in the video may well have an adverse effect on the learner's attention, as it may have a similar negative impact as the embodiment principle above.

2.1.11 Advanced Principles of Multimedia Learning

The set of basic principles listed above is expanded upon further by Mayer, (2014) with the list of Advanced Principles of Multimedia Learning below. At the time that both sets of videos were created the basic principles were the most prominent, therefore they were utilised during the production of the media. The advanced principles will be defined and referred to in the discussion, should the area that each principle occupies be an area for attention after analysis of the study's findings. The advanced principles are:

Guided discovery learning: Learning is increased when guidance is incorporated into multimedia sources that are discovery-based.

Worked examples principle: During initial skill learning learners attain more success when they receive worked examples.

Self-explanation principle: When learners are encouraged to generate explanations of their own it increases there learning.

Drawing principle: Learners learning is increased when they create drawings when reading text.

Feedback principle: Learning from multimedia lessons is enhanced when learners receive feedback regarding their performance.

Multiple representation principle: At particular times learning will be enhanced when aspects are represented in multiple ways/times.

Learner control principle: Providing a learner with control of the selection and organisation of the material will not necessarily result in more effective learning.

Animation principle: Animation may not always be superior to static pictures.

Collaboration principle: Collaborative online learning activities can promote further learning.

Prior knowledge principle: Aspects of the instructional design that are integrated to assist novice learners may hinder experienced learners.

Working memory principle: The capacity of the learners working memory is integral to the efficacy of the instructional material.

(Mayer, 2014)

Many of the principles have been discussed, or in part referred to in the basic principles above and therefore do not warrant a more extensive discussion at this stage. There is an evidence-base for each, however there also needs to be an appropriate constraint on the number of principles that are used. Having an overly extensive list, may in fact restrict the instructional designer too much, limit creativity, or be a barrier preventing the development of multimedia instructional material that could be used to enhance students learning.

As a result Meij & Meij, (2013) attempt to address the extensive array of principles that can be applied to instructional videos and they present eight concise guidelines (Table 1). The eight guidelines are to aid the instructional designer based upon the work of various authors who have researched instructional design and human cognition. Meij & Meij, (2013) also state that their guidelines summarise key notions of accepted thinking and can be thought of as a way of highlighting “best practice”.

Guideline 1	Provide easy access
	1.1 Craft the title carefully
Guideline 2	Use animation with narration
	2.1: Be faithful to the actual interface in the animation
	2.2: Use a spoken human voice for the narration
	2.3: Action and voice must be in synch
Guideline 3	Enable functional interactivity
	3.1: Pace the video carefully
	3.2: Enable user control
Guideline 4	Preview the task
	4.1: Promote the goal
	4.2: Use a conversational style to enhance perceptions of task relevance
	4.3: Introduce new concepts by showing their use in context
Guideline 5	Provide procedural information
Guideline 6	Make tasks clear and simple
	6.1: Follow the user's mental plan in describing an action sequence
	6.2: Draw attention to the interconnection of user actions and system reactions
	6.3: Use highlighting to guide attention
Guideline 7	Keep videos short
Guideline 8	Strengthen demonstration with practice

Table 1 Eight Guidelines for the Design of Instructional Videos for Software Training (Meij & Meij, 2013).

The guidelines above are closely aligned to the basic multimedia principles that are provided by Mayer, (2014) and whilst the table is a concise summary, the remainder of the paper elaborates on each of them further. Although the authors of the paper state that they have merely summarised the current position, it provides the instructional designer with a clear set of simple descriptors when devising their materials. As a result the guidelines from Meij & Meij, (2013) and those established by Mayer, (2014) are the most well evidenced, applicable guidelines currently available for instructional materials. In order to enhance their application, the review will return to the three established learning theories and contextualise each within the setting of instructional videos and their design principles. It will then detail how each can be monitored and assessed within the use of functional hierarchies.

One aspect that is unclear when reviewing the literature is 'what is the effect size going to be when instructional designers apply all (or most) of the principles?'. The vast majority of the individual studies cited up until now by Mayer, (2014) when calculating the net effect size for each principle tend to be focused on one, two or three principles at most for obvious reasons. Therefore it is unclear if the effects are additive or averaged for each principle when they are incorporated into the instructional video design and production. Is it possible to have too many of the principles within each video which could then result in a negligible or negative effect? This is a question that could be difficult to address due to the vast number of potential variables that would be incorporated into any study. Additionally, the sheer number of videos that would have to be produced, to encompass all of the permutations possible to investigate the most effective combination would prove to be unfeasible.

Chapter 3 Literature Review

3.1 Introduction

The second part of the literature review aims to outline and contextualise the underlying pedagogical theory against which the efficacy of the videos is to be evaluated against. The literature review will have a broad approach encompassing the basic tenants of learning through to specific taxonomies that aim to provide an evaluative tool for the quantification of learning.

3.2 Learning Theories

3.2.1 Behaviourism

Behaviourism has had a long and extensive history with many authors, predominantly Thorndike, (1913), Pavlov, (1927) and Skinner, (1974) contributing to its development over more than a century of investigation. Aspects of behaviourism are well known by a large proportion of the population thanks to the classical conditioning work undertaken with dogs by Pavlov, (1927). Behaviourism as a theory is relatively simple, Skinner, (1974) proposes that learning occurs when there is an observable change in behaviour that has been brought about by an external environmental stimulus. A common example is the increase in saliva produced by Pavlov's dogs when a bell is heard, due to the associative nature of the bell with food. Behaviorists state that this change in observable behaviour is the indicator that learning has occurred, and what is occurring in the learners head is not an indicator or indeed as some state important to the process, which is a controversial standpoint with many educators (Alonso et al., 2008; Deubel, 2003; Eccles & Feltovich, 2008; Ouyang & Stanley, 2014; Royer, 1979; van Bruggen, 2005). The viewpoint has been rationalised by Skinner, (1974) by stating that it is not possible to objectively prove any of the inner processes of learning, however it is possible to quantify observations and as such researchers should focus on 'cause and effect relationships' (Hung, 2001). The simplicity of behaviourism has been its strength, but also its downfall in an educational context. This is due to its absence of attention to the processes or understanding occurring in the learners head, a prime example of this was that behaviourism is unable to explain some social behaviours (Alonso et al., 2005). Other authors also state that behaviourism has few applications in adult knowledge-based education, which may have particular consequences for the current study given the population (Lau, 2014).

Regardless of its limitations behaviourism is credited with the legacy of educational technology (Deubel, 2003). It is the foundation for innovations such as computer assisted instruction, mastery learning, minimal competency testing, educational accountability, situated cognition and social constructivism and other educational methodologies (Deubel, 2003; Eccles & Feltovich, 2008) that use the concept of instructional design (Chen, 2007). The role of behaviourism within multimedia design has predominantly been associated with lower-level learning tasks, such as the rapid acquisition of basic concepts and skills, it is yet to be proven with higher-order learning tasks (Deubel, 2003; Samaras et al., 2006). The rationale behind this argument is that behaviourist designs, are passive and do not foster cognitive activity that is generally associated with higher-order tasks (Rovai, 2004; Samaras et al., 2006). As a result of this, behaviourism has been declared redundant many times over, it has been viewed as an outmoded approach that is inferior to cognitively orientated studies to such an extent that, in principle it can be ignored without reprisals (Keijzer, 2005). Nevertheless, there are signs that a variation of behaviourism might emerge again as cognitive multimedia design approaches appear to have a severe quandary; principally that there is no easy way to quantify their effectiveness, as it is extremely difficult to determine what is occurring in the learner's mind as initially argued by Skinner (Keijzer, 2005; Samaras et al., 2006). This is in distinct contrast to the inherent simplicity of the behaviourist approach. It is increasingly accepted that many cognitive processes rely on external processes and the dynamic interplay between internal, cognitive processes and bodily and situational characteristics (Keijzer, 2005). As a result of the basic tenets and the direct link to psychomotor skill acquisition, any investigative approach into the use of video and practical skills will need to incorporate an element of behaviourism, fundamentally to provide a measure that can be used to determine the effectiveness of the learning process. How this is applied however is a distinct area of contention. The vast majority of practical skills will be composed of many performance levels, various specific intricacies and subtleties that may well determine whether the skill performed is correct or incorrect. Whilst this does relate back to the central premise of behaviourism, i.e. the change in behavior will be correct or incorrect, there is the requirement for a sliding scale to objectively quantify this, as the skill could be correct with errors that will require further learning by the participant.

3.2.2 Cognitivism

In the late 1950's learning theories began to move away from the use of behavioural models and toward models of the cognitive sciences which views learning as the acquisition or reorganisation of cognitive structures to process and store information and acknowledges the importance of prior knowledge during learning (Alonso et al., 2008; Eccles & Feltovich, 2008; Lau, 2014). Cognitive functioning was originally devised as a process that occurs predominantly within the head, and could in theory be performed by a computer (Keijzer, 2005). Cognitivists believe that humans are able to receive, process and use information via various means (most of which have already been outlined) and integrate them to form appropriate cognitive development depending upon the learners specific stage (Ouyang & Stanley, 2014). The process is thought to have many stages and can involve memory, thinking, reflection, abstraction, motivation and meta-cognition (van Bruggen, 2005). Cognitivism has had a profound effect on instructional design and this is evidenced by the organisation, chunking and the use of metaphors in a simple to complex way in instructional materials (Alonso et al., 2005). Cognitivism also relates to Cognitive Load Theory, which has been extensively discussed previously and is the foundation of the predominant design principles for instructional multimedia. Cognitivism however is not an over-arching all-encompassing learning theory but one which emerged through the contributions of leading theorist such as Piaget, (1959) and Vygotsky, (1962). Their detailed work and contributions along with others led to the development of contemporary constructivism.

3.2.3 Constructivism

Constructivism has many varying interpretations but each of them has a common philosophy. Learning is understood, stored and applied most effectively when learners develop their own mental representations from presented information (Burton et al., 2003; Eccles & Feltovich, 2008; Fee & Fee, 2003; Mayer, Moreno, Boire, & Vagge, 1999; Vogel-Walcutt et al., 2011). Constructivism builds upon behaviourism and cognitivism as it accepts multiple perspectives and maintains that learning is a learners personal interpretation of the world (Alonso et al., 2005; van Bruggen, 2005). At present, educationalists are emphasising the need for a student-centred learning approach, an aspect that constructivism emphasises as the fundamental theory which supports the transition from a teacher-centred to a student-centred learning approach (Applefield & Huber, 2000; Chen, 2007; Kala et al., 2010). Although it has been noted that constructivism is complex and full of controversy, a constructivist approach in education can be reduced to the concept that learners actively construct their own knowledge rather than passively receive it (Kim & Reeves, 2007; Vogel-Walcutt et al., 2011). As a result, the contrasts with behaviourism are clearly evident and the theoretical link with cognitivism is implicit during the initial phases of the constructivist approach, with constructivism continuing the learning process by applying situational and personal interpretations to the subject matter that is being presented. Therefore a significant challenge within multimedia learning concerns the notion as to whether it is possible to promote constructivist learning from passive media (Mayer, 2002). This is particularly beneficial during psychomotor skill acquisition when we consider behaviourism, which could then facilitate the cognitivist and constructivist approaches to potentially develop higher-order learning. Constructivism-based instructional goals often include providing the learner with skills or support and encouraging the learner to actively construct his or her own personal learning experience (Vogel-Walcutt et al., 2011). Mayer, (2002) elaborates further by stating that although the media is passive, constructivism can be fostered when cognitive activity is present within the learner. As a result (Mayer, 2002) concludes that it is not necessary to have a hands-on activity or social collaboration but in essence a well-designed multimedia explanation within an appropriately interactive learning environment should provide conditions that will promote constructivism. The role of constructivism within the learning process from instructional videos is confirmed, as deeper learning occurs when students can transfer and apply the concepts to a novel situation (Austin, 2009). However, the MMLT is typically related with information acquisition from a cognitivist approach (Mayer, 2002) and therefore the application of the theory is warranted to maximise the efficiency of the instructional media from a learning theory perspective. This theoretical concept has been empirically tested by Vogel-Walcutt et al., (2011) who compared CLT to constructivism when teaching complex military command and control task. Whilst the authors identified no distinct differences between the groups for the acquisition of procedural, declarative, and conceptual knowledge and decision-making skills, there was a difference between the recall of integrated knowledge. This finding is contrary to their expectations as the CLT approach was superior to the constructivist approach, which aligns itself with the integration of knowledge by placing the learner at the centre of the learning experience. Therefore, whilst constructivism may have its distinct advantages according to those authors cited here, the use of cognitivism via the CLT and the MMLT principles in conjunction with constructivism may yield the most effective approach. This is an area which has begun to emerge and some authors Kala et al., (2010) for example are integrating constructivism with e-learning approaches.

3.3 Interim Summary

In summary, up until now both parts of the literature review have detailed how instructional designers should design, construct and edit multimedia materials that can present and deliver information effectively. Each of them should be formulated so that learners can acquire cognitive and psychomotor knowledge through the visual and auditory pathways. This directly relates to how the learners will theoretically process the information within their working memory, form mental schemas that can then be stored in the long-term memory ready for retrieval when required. As a result, the review has focused on all of the most fundamental aspects of an educational study bar one; assessment. Whilst it has been touched upon at various points throughout the review thus far, the assessment of any learning needs to be grounded in an accepted methodology. There are a multitude of potential models available to educational practitioners. However a dominant theoretical classification of learning objectives that incorporates a psychomotor element,, that is also ecologically valid given it is derived to provide a taxonomy for the analysis of University level education, and has construct validity given its foundations in behaviourism is Bloom's Taxonomy (Ferris & Aziz, 2005).

3.4 Bloom's Taxonomy

Bloom's Taxonomy was originally published in 1956 with the aim of providing a method of classifying educational system goals, this was to aid the discussion of curricula and evaluation challenges with a level of precision (Amer, 2006; Ferris & Aziz, 2005). The taxonomy addressed one of the major problems within education, rote learning and this is a contributory factor to its success. As a result of its use after only twenty years Seddon, (1978) stated that there was no doubt that it has had a considerable impact on education all over the world, this is further supported by Kennedy, Hyland, & Ryan, (2007) who state that is still the most widely quoted in literature.

Bloom's taxonomy was proposed with three overlapping domains; cognitive, affective and psychomotor (also known as the knowledge, skills and attitudes (KSA) taxonomy). Each of the domains has its own internal hierarchy. Achievement at a particular level was required for progression up the pyramid in which they were visualised (Alonso et al., 2008; Croft, 2010; Dawson, 1998; Pickard, 2007). Although basic, the three-way divide appears to be the fundamental and most enduring structure of a large number of educational taxonomies (Lynch, Russell, Evans, & Sutterer, 2009). Despite being composed of three domains the cognitive domain is the most well-known and was first published in 1956 (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956; Thomas, 2005) with the Affective Domain following just under twenty years later (Krathwohl, Bloom, & Masia, 1973). The psychomotor domain was omitted from the original publication and the team led by Bloom never published a hierarchical taxonomy. This was because the original psychomotor committee members were unable to find psychomotor objectives in the literature. They deemed that it was not important to teach manual skills to college students (Anderson et al., 2001; Ferris & Aziz, 2005; Hanna, 2007) and whilst it is acknowledged that no single theoretical foundation exists for instructional design, a blending of behaviourist and cognitive approaches seems inevitable (Deubel, 2003).

3.4.1 The Cognitive Domain

The original taxonomy divided the cognitive domain into six major categories (Figure 2) all except for application, were further sub-divided and ranked in a hierarchy from simple to complex (Bloom, 1956; Krathwohl, 2002). The initial aspects of the cognitive domain are as simple as recalling previously learned material. This is directly comparable to the basic tenets of cognitivism, through to creatively combining an assortment of ideas to synthesise something original, a significant concept in constructivism (Croft, 2010). Whilst the original taxonomy was of great use within an educational context, it received a large amount of criticism throughout the years. Most notably the criticism centred on the linear, hierarchical nature of the domain from simple to complex and how it is not representative of how learners acquire knowledge (Amer, 2006). Furthermore it does not have a direct practical applicability to curriculum planning (Cannon & Feinstein, 2005), the taxonomy has an over-reliance on behaviourism (Amer, 2006; Ferris & Aziz, 2005) and the taxonomy did not include a section for 'understanding' which is a commonly used term within a teachers vocabulary (Krathwohl, 2002). As a result of these criticisms the original taxonomy was revised by a group of cognitive psychologists, curriculum and instructional researchers, and testing and assessment specialists in 2001. This differs from origins of the original taxonomies which were developed by an American educational committee (Amer, 2006; Anderson et al., 2001; Cannon & Feinstein, 2005; Krathwohl, 2002; Seddon, 1978).

The revised taxonomy includes several significant changes related to certain assumptions, structure and terminology. Most notably these see the taxonomy change from one, to two dimensions by separating the use of nouns (Knowledge Dimension) and verbs (Cognitive Process Dimension) (Amer, 2006; Cannon & Feinstein, 2005; Krathwohl, 2002; Pickard, 2007). The revised taxonomy as a result is now structured as a matrix (see Figure 3) rather than a hierarchy, which visualises the potential for interaction and is seen as "a tool to help educators clarify and communicate what they intended students to learn as a result of instruction". This can be seen as another way to help align learning objectives, curriculum and assessment that complements the complexity of learning, in a subject matter-specific instruction (Anderson et al., 2001; Cannon & Feinstein, 2005; Hanna, 2007).

The Knowledge Dimension relates to three categories of cognitive psychology. These arose since the conception of the original taxonomy and the rise of metacognition as a pedagogically significant approach, that was not widely recognised at the time that includes: Factual, Conceptual, Procedural and Metacognitive Knowledge (Amer, 2006; Krathwohl, 2002). Metacognitive knowledge has two distinct types of knowledge; explicit/conscious, factual knowledge and implicit/unconscious knowledge (Hanna, 2007). Whilst metacognitive strategies can enhance learning they can also hinder learning if used in an environment that is to challenging i.e. particular problem-solving scenarios (Merriënboer & Sweller, 2005b; Wittrock, 1989). The metacognitive demands are greater for learners in a loosely structured learning environment; therefore adherence to instructional guidelines in the formulation of videos (such as those used within the current study) will decrease demand and result in greater learning in theory (Deubel, 2003).

The Cognitive Process Dimension relates closely to the original taxonomy, although three of the six categories have been renamed, the order has been changed to that below (inverted to the schematic representation). The descriptors all utilise verbs to ensure their congruency with instructional objectives:

- Knowledge became Remember
- Comprehension became Understand
- Synthesis became Create
- Application became Apply
- Analysis became Analyse
- Evaluation became Evaluate

Whilst they are listed in an order there is no defined hierarchy as they are allowed to overlap which gives ‘Understand’ in particular far greater scope (Amer, 2006; Anderson et al., 2001; Krathwohl, 2002). The Cognitive Process Dimension reflects learners cognitive and metacognitive activity that is fundamental to the learner-centred constructivist approach, enabling them to make sense of the information in conjunction with prior knowledge (Amer, 2006; Pickard, 2007).

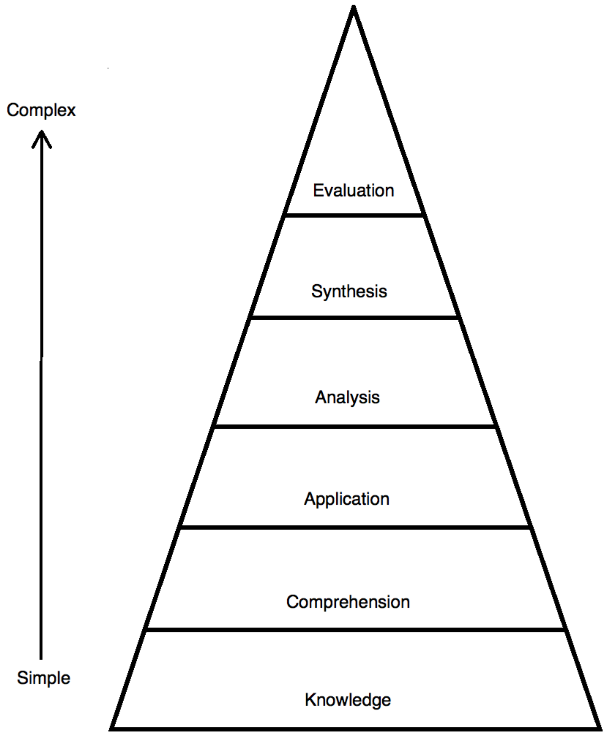

Figure 2 Cognitive Domain from the Original Taxonomy sourced from Bloom et al., (1956)

The Knowledge Dimension	The Cognitive Process Dimension				
	Remember	Understand	Apply	Analyse	Evaluate
Factual Knowledge					
Conceptual Knowledge					
Procedural Knowledge					
Meta-cognitive Knowledge					

Figure 3 Bloom's revised taxonomy table sourced from Anderson et al., (2001)

The differences between the taxonomies are clearly apparent, the transition from behaviourism to constructivism, demonstrates the modernisation of the cognitive domain (Amer, 2006) however, there are still criticisms of the revised taxonomy. Fuller et al., (2007) summarise some researchers thinking that Bloom's Taxonomy does not give sufficient emphasis to critical thinking (something that the revised taxonomy endeavours to capture), a core component of University level education. Critical thinking is far greater than the cognitive domain as it encompasses reflection, another particularly important aspect of University level education (acutely relevant to the context of this study). Whilst there are taxonomies that focus on reflective judgment and critical thinking, both are predominantly just for this aspect and do not share the same kind of evidence-base that Bloom's taxonomy enjoys.

Although the current study encompasses all three domains, it is biased towards the psychomotor domain, due to the overly practical nature of the content being conveyed in the instructional videos. The cognitive domain is nevertheless important, and there is an inherent difficulty assessing cognitive aspects in a practical demonstration/setting. Unfortunately, the ability to be critical and reflect on the practical performance, during the session are integral skills that require attention and is therefore an area were Bloom's taxonomy may not be as valid as it could be. Furthermore the cognitive aspects can generally be assessed in a one-dimensional way i.e. they understand the exercise or they do not. Whilst this is a fairly basic analogy however, it does demonstrate that the application of the revised taxonomy maybe over-complicating the assessment of the cognitive domain. This is due to the current study being skewed towards behavioural aspects, despite its criticisms; it seems that utilising the original taxonomy may in fact be more valid and efficient. In particular, it may also seem that an amalgamation of the two and by using the cognitive process dimension as a linear hierarchy, would provide a suitable method for assessment. That said however the terminology of the original taxonomy is well known and has been employed in a vast number of subject areas (Hanna, 2007). As a result this presents a distinct quandary from which there are numerous advantages and disadvantages to each of the cognitive taxonomies. Given that the fundamental aim of the current study focuses on the psychomotor aspects due to the nature of the task and the vocation, the original taxonomy shall be adopted and employed. This will be used with reference to the knowledge and cognitive dimensions as they have been clearly mapped from one taxonomy to the other. It is anticipated that this will provide a suitable methodology for assessment and enable discussion that aligns itself with cognitivist and constructivist learning theories and not just behaviourism.

3.4.2 The Affective Domain

The affective domain was originally started by Bloom and colleagues however it was completed by Krathwohl, Bloom, & Masia in 1964. Whilst the cognitive domain has received the vast proportion of attention, it is acknowledged that the cognitve and affective domains overlap through the metacognitive area (Lynch et al., 2009). The affective domain (an area often inadequately addressed in education) includes objectives that describe changes in interest, attitudes and values (Bloom et al., 1956; Krathwohl et al., 1964). These relate to values ranging from the act of simply giving attention, to the profound internal qualities such as character and conscience (Bloom et al., 1956; Cleveland-Innes & Ally, 2013; Croft, 2010; Hanna, 2007; Krathwohl, Bloom, & Masia, 1964; Rovai, Wighting, Baker, & Grooms, 2009; Sahasrabudhe & Kanungo, 2014). Some researchers extend this by stating that the emotions should be positive or about professional values. Others have also referred to the affective domain in learning outcomes for workplace training; this can be used as a mechanism to develop Soft Skills Training. This is due to the fact that it is recognised as the external expression of internalised emotion, through the stages shown in Figure 4 (Cleveland-Innes & Ally, 2013; Krathwohl et al., 1964; Reid, 2012; Rovai et al., 2009). As learners progress from receiving, to characterising by value or value concept this denotes an increasing amount of internalisation of interest, attitudes and/or values in the subject or material being taught (Lynch et al., 2009). Snelson & Elison-Bowers, (2009) provides a clear link between the use of YouTube videos and the role of emotions from the affective domain, highlighting the importance of emotion. The perspective is also supported by Shen, Wang, & Shen, (2009) who stated that the influence of emotions on learning is still under-emphasised. As such the role of emotions in any learning endeavour and outcomes, especially online learning is only now beginning to be understood. Shen et al., (2009) undertook a study evaluating the difference between an emotion-aware system and a non-emotion-aware system for e-learning. Despite a number of limitations (mainly in the quantitative assessment of emotions and the small sample size) they were able to show a significant improvement in learning with the emotion-aware system (Shen et al., 2009). Snelson & Elison-Bowers, (2009) also provide examples emphasising the importance of emotion, which relates to a case study describing an emotionally charged health care event, with multiple instructional objectives that map to the affective domain. The objective of doing so by the authors is that each learner will internalise the information as part of their core value system, and then respond appropriately in the professional setting. The affective domain is engaged through the emotional story, evocative imagery, and mood-provoking music much in the same way that theatrical movies evoke emotions to move audiences to laughter or tears (Snelson & Elison-Bowers, 2009). Multiple authors have highlighted the requirement for further research into the mechanisms behind e-learning/instructional videos to understand how they can be used effectively to meet all the intended learning objectives. This is especially the case for those in the affective domain, which are paramount to subject areas that heavily integrate values, empathy and emotion (Cleveland-Innes & Ally, 2013; Rovai et al., 2009; Shen et al., 2009; Snelson & Elison-Bowers, 2009).

The affective domain has received more attention than the psychomotor domain but significantly less than the cognitive domain. As a result, there have been no dramatic revisions proposed to its structure, namely because it is apparent that the limitations of the affective domain centre on how it is assessed. The use of Electroencephalography, heart rate, blood pressure and skin conductance sensors by Shen et al., (2009) is at the most detailed end of the hypothetical assessment continuum, quantifying as many factors as possible, however emotions are never easily quantified. Further aspects relating to the assessment of the affective domain will be discussed further in the literature review in section 3.4.2.

The examples provided within this section demonstrate two differing approaches to the application of the affective domain; the current study will attempt to apply them in a suitable way. Whilst it will not be feasible to create an emotionally charged instructional video, similar to the case studies referred to by Snelson & Elison-Bowers, (2009) (an aspect to which there is no distinct guidance within the multimedia design principles from (Mayer, 2014) all attempts to engage empathy from the learner within the videos will be adopted where possible. This will be achieved through the use of affective learning outcomes, similar to those used by Shen et al., (2009) and recommended by other authors (Cleveland-Innes & Ally, 2013; Lynch et al., 2009; Rovai et al., 2009). The use of the affective domain and its overlap with the cognitive domain will hopefully create generative learning through the use of positive emotions, whilst the learners view the instructional videos. However affective aspects, although they have a long tradition in research on learning instruction have been neglected in a technology-based environment (Leutner, 2014).

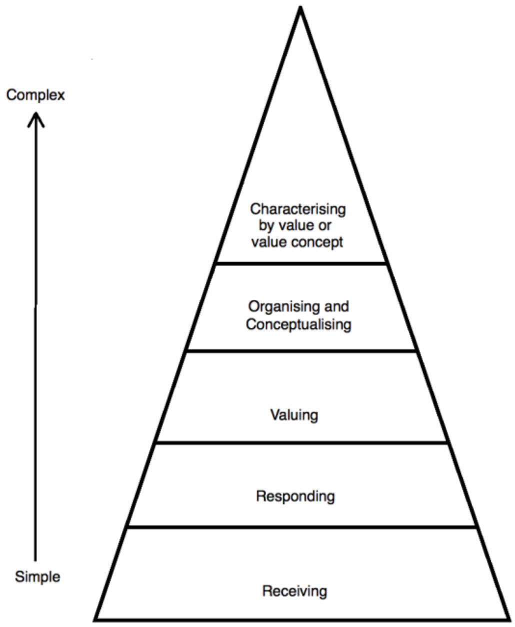

Figure 4 Affective Domain

3.4.3 The Psychomotor Domain

Bloom's (1956) taxonomy of educational objectives has been a cornerstone for analysing and thinking about the goals of educational activities in the cognitive and affective domains. Thoirs & Coffee, (2012) state that the original taxonomy did not develop the psychomotor domain however, Simpson, (1966) provides a quote from Bloom and colleagues that justifies their decision:

> "Although we recognise the existence of this domain, we find so little done about it in secondary schools or colleges, that we do not believe the development of a classification of these objectives would be very useful at present"

Subsequent statements by the authors of the original taxonomies reaffirmed their positional stance and lack of intention to develop a third domain

(Simpson, 1966). However the psychomotor domain addresses the obvious issue, that conscious knowledge or values and attitudes are not viable mechanisms to explain; the effective performance of learned physical tasks (Rovai et al., 2009). As a result many authors since this quote have sought to extend or complete Bloom’s Taxonomy so that it fully encompasses a psychomotor domain. There has however, not been a unanimously agreed consensus of opinion regarding the classification of this domain. Simpson, (1966) attributes this to the involvement of the other two domains within the psychomotor domain, that creates a ‘special problem of complexity’ when developing a classification system for it. Whilst many authors have proposed and reviewed (within their own papers) the variety of psychomotor domains currently available, only those that have been cited by other authors will be included within this review. A plethora of psychomotor taxonomies have been proposed for a wide variety of subject areas and skills, some without obvious foundation or supporting evidence. The psychomotor domains reviewed will be presented in chronological order and critiqued at the end of this section (see section 3.5).

Simpson’s Psychomotor Domain (1967)

The first psychomotor domain to be developed was undertaken by Simpson and published in 1967; the methodology is documented in a report published in 1966. In the report Simpson, (1966) details the step wise process that was undertaken to form the tentative system over the period of approximately two years (Appendix 2). As the authors sought to make the domain applicable to multiple subject areas (Industrial Arts, Agriculture, Home Economics, Music, Physical Education and Art) the domain is quite elaborate and complex. Particularly when viewed in comparison to the relatively simple hierarchies and matrix of the two original and the revised taxonomies (Urbach, Singer, Simpson, Greer, & Fleishman, 1972). When the proposed domain is viewed in a linear hierarchy (Figure 5) representing the progression through the stages, the title headers for the domain are not entirely understandable, without further explanation.

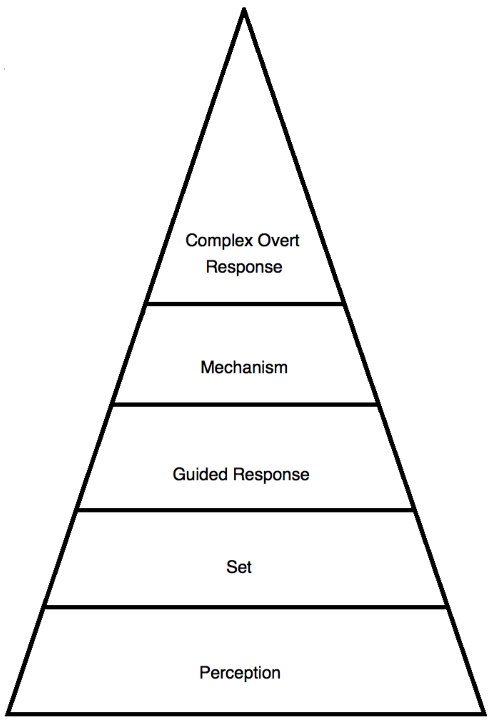

Figure 5 Simpson's Psychomotor Domain

The domain was adapted in 1972 with the inclusion of Organisation and Adaptation above Complex Overt Response, this is to encapsulate and extend the applicability of the domain to problem solving (Simpson, 1972). Simpson's domain has been adopted by some due to the inclusion of two preparatory levels as prerequisites to the development of a physical task. This has been deemed to be useful when mapping and aligning psychomotor learning with learning outcomes (Olson, 2008). Kovacs (1997) provides one of the best contextualisations of Simpson's domain, in a subject area related to the current study by defining the levels as:

Perception - is the initial step that requires awareness of performance. This could include the tactile sensation of a suture needle penetrating the skin, selecting appropriate cues to respond to during a procedure such as bag-mask ventilation or altering (translation) the cue that caused the initial problem.

Set - is the state of perception or readiness that requires mental knowledge to perform the procedure. Physical knowledge i.e. how to hold and use the tools for the procedure and emotional set, which can be seen as the desire to correctly demonstrate the procedure.

Guided Response - is the motor act under supervision from an instructor (imitation) or as a result of self-evaluation from trial and error.

Complex Overt Response - requires the learner to proficiently, efficiently and effectively performs the previous steps with an optimal outcome from the procedure.

Adaptation - requires the modification of an established skill to a novel area i.e. reduction of a fracture from a long bone to an irregular bone.

Originating - is where new procedural skills are devised.

Kovacs, (1997) states that Simpson's psychomotor domain has been adopted by the Advanced Trauma Life Support (ATLS) Instructors Course. This emphasises the importance of this and/or a psychomotor domain, as this is a high-pressure skillset that doctors in Accident and Emergency are required to obtain prior to attempting to save a patient's life. Simpson's domain is extensively cited in text that refers to psychomotor domains, and is seen as one of the original models from which others are based. It is also used to provide the inspiration for others, which exemplifies the importance of the work undertaken and published.

Dave's Psychomotor Domain (1970)

Dave, (1970) proposed a taxonomy that has five levels and bases its foundations clearly upon behaviourist theory. The first level is *imitation* and the ascending pathways lead to *naturalisation* of the activity and skill (Figure 6). Dave's domain has been referred to extensively (Krathwohl, 2002; Radin Salim, Puteh, & Mohd Daud, 2011; Rupani, 2011; Thoirs & Coffee, 2012) within literature as an alternate psychomotor domain. There are however, very few empirical studies that actually employ the domain in a practical setting which is a distinct concern relating to the validity/applicability of the domain. A web-based critique cited by Radin Salim et al., (2011) analysed the psychomotor domain from Dave. They stated that the levels within the domain were suitable for adults, which was in contrast to the domain from Simpson, (1966) which has been identified as being better suited to the development of children and young people. It is clear as to how this conclusion has been arrived at as the titles of Dave's domain and the explanations are as follows:

Imitation - Observing and copying the behaviour of another person.

Manipulation - Performing actions governed by instructions or self-directed practice.

Precision – The learner is able to perform the task with minimal error and with practice will become more precise without the presence of the original example.

Articulation – The learner is able to co-ordinate more than one action by amalgamating skills that can be modified to fulfill requirements or solve problems.

Naturalisation – Skills are performed with a high level of performance and accuracy autonomously. Skills are also combined, sequenced and performed consistently with minimal effort.

(Kennedy et al., 2007)

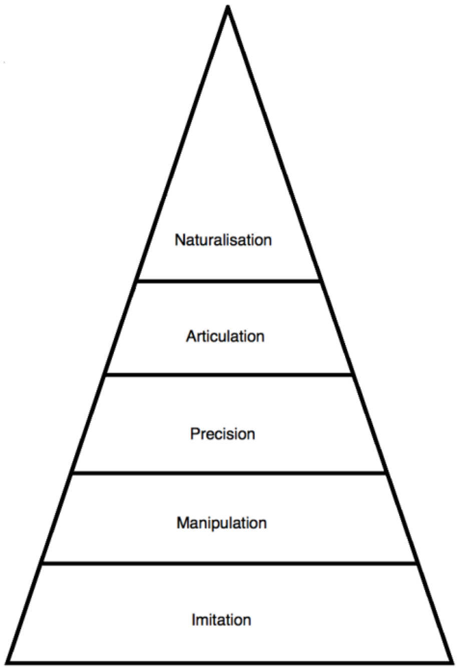

Figure 6 Dave's Psychomotor Domain

Harrow’s Psychomotor Domain (1972)

Soon after Dave’s taxonomy another was put forward by Harrow, (1972) with the emphasis on physical ability for children with physical needs. The taxonomy is designed to assess ability in a recreational or sporting environment, rather than skills or tasks in a workplace (Thomas, 2005). Harrow’s psychomotor domain is cited extensively within articles (Kennedy et al., 2007; Kovacs, 1997; Krathwohl, 2002; Lynch et al., 2009; Thoirs & Coffee, 2012; White, 2010) to name but a few, but very few of them discuss the domain in any detail. It has not been mapped to the outcomes of their study or referred to in the discussions, as such the domain is particularly difficult to evaluate given the scarcity of information about it. The levels are defined below and are represented in Figure 7.

Reflex – generally deemed too low a level for objectives to be written as they are elicited without learning.

Fundamental movements – predominantly applicable to young children (run, jump, reach etc.) that can be formed from a group of reflex actions.

Perceptual abilities – these can be defined as the ability to catch, write, balance and manipulate objects, i.e. stimulus has to be interpreted for an appropriate response to be triggered.

Physical abilities – these can be defined as the ability to stop, increase pace, change direction or react.

Skilled movements – these are classified as being able to hit an object, play, swim and dive, generally these are as a result of acquiring a complex skill/task.

Non-discursive communication – is defined as being able to express, create, mime design and interpret, generally these are from facial expressions but can include postures and gestures.

(Harrow, 1972; Isaacs, 1996; Thomas, 2005)

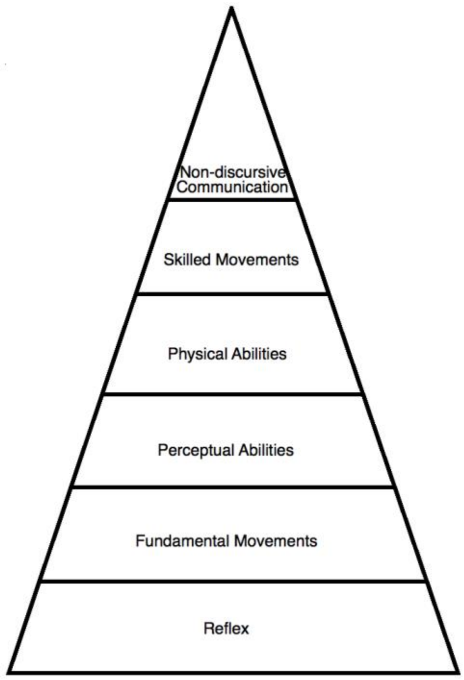

Figure 7 Harrow's Psychomotor Domain

Dawson’s Psychomotor Domain (1998)

Dawson’s taxonomy is similar in nature to that of (Dave, 1970) and provides a hierarchy from observation through to mastery (Figure 8). Dawson, (1998) aligns the pertinence of the psychomotor domain to sport, craft and trade training. In addition they also align it to the objectives of the performing arts and those professions that require manual dexterity, such as laboratory practice and surgery. The domain builds on the behaviourist approach, as there is a clear pathway for the development of higher-order learning, given the mastery component of the hierarchy. However this is purely theoretical as the paper by (Deubel, 2003) stated that it is still yet to be proven.

The application of Dawson’s domain has been limited predominantly due to the fact that the book "Extensions to Bloom's taxonomy of educational objectives” was only published in Australia. The textbook had a limited run, meaning that there are only a few copies outside of Australia. Nevertheless Dawson, (1998) has been used to develop a competency testing instrument for assessing musculoskeletal sonography, when using a multimedia DVD. Thoirs & Coffee, (2012) accepted that the domain could be seen as an extension of Bloom's taxonomy, they state that it can be used to assist students and teachers in the stages of learning. Dawson, (1998) defines each level as the following:

Observation – the learner will observe a demonstration form an expert as to how the task should be performed.

Trial – an initial attempt is made to perform the task, if it is not successful then a verbal correction from an instructor might be appropriate or re-observation depending on the nature of the task.

Repetition – when a trial is performed correctly it is followed by a number of repetitions to improve their performances of the task. They are not classified as trials but seen as practice, intervention is only required if an incorrect element of the task is adopted through repetition.

Refinement – is composed of further attempts by the learner to perform the task brining about improvements in their performance.

Consolidation – is seen when the tasks are firmly instilled in the learner and they are able to achieve them as proficiently after a specified interval.

Mastery – every time the task is performed it is of the highest possible standard.

(Dawson, 1998)

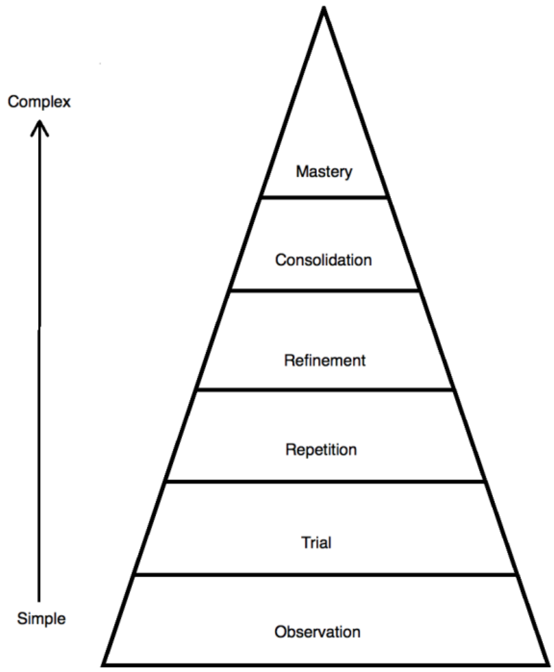

Figure 8 Dawson's Psychomotor Domain

Ferris and Aziz's Psychomotor Domain (2005)

Ferris & Aziz, (2005) sought to create a psychomotor domain, as they could not see how any of the previous taxonomies would be applicable for use in higher education, principally for use within engineering education. Upon creation of their own domain, other authors have sought to validate it or map the domain to other electronic laboratory skills with a moderate level of success (Radin Salim et al., 2011).

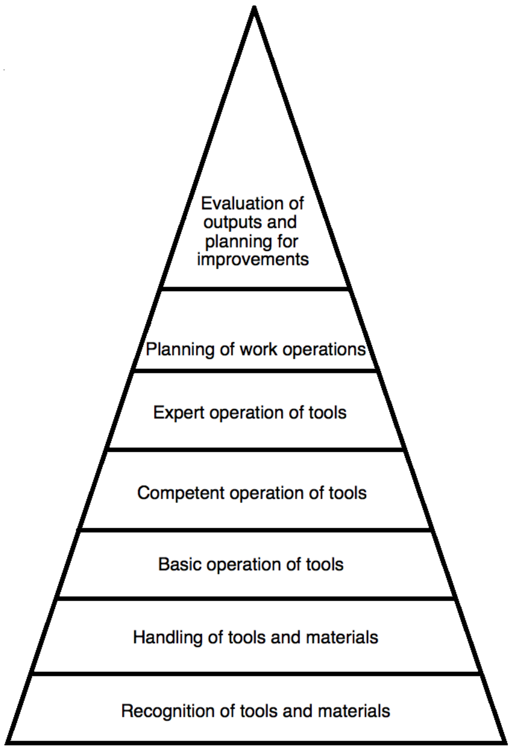

Figure 9 Ferris and Aziz's Psychomotor Domain

The levels of the taxonomy have a fairly clear meaning from their titles until the final stages of the domain:

Recognition of tools and materials – Ability to recognise the appropriate tools and materials required.

Handling of tools and materials – Ability to appropriately handle the previously identified tools and materials.

Basic operation of the tools – Ability to use the tool to complete basic tasks.

Competent operation of the tools – Ability to use the tools fluently when performing a task.

Expert operation of tools – Ability to use the tools extremely effectively and efficiently in a safe way.

Planning of work operations – Ability to transform a piece of work subject to a specification.

Evaluation of outputs and planning for improvement - Ability to review the final output and indemnify deficiencies and discrepancies.

(Ferris & Aziz, 2005; Radin Salim et al., 2011)

The authors of the taxonomy state that theirs has returned to the primary target of the taxonomy, higher education. By referring to the development of the manual skills, that correspond to professional competencies rather than just gross physical motor function, can be seen as the foundation of other psychomotor domains (Ferris & Aziz, 2005).

3.5 Critique of psychomotor taxonomies

After an extensive search through the literature, the five taxonomies listed above represent the most frequently cited and used hierarchies for the psychomotor domain. The domains cover a large amount of time between each of them and that is certainly reflected in the terminology and descriptions used within each. This however is insufficient a reason to exclude one over another due to the age of the cognitive and affective domains from Bloom and colleagues, despite the updates to the former. One of the first critique's is cited by Ferris & Aziz, (2005) who state that the taxonomy extensions (up until theirs) have been designed to cover the lower levels of education (children's physical and skill co-ordination). As a result, it is stated that the domains are not applicable to higher education, which poses a unique problem given the context of the current study. Whilst this would seem to be a fair assessment of Simpson's and Harrow's taxonomies, due to the rationale behind their formulation (even though Harrow's has been cited as being used in course design for undergraduate students (Reeves, 2006)) it is unclear as to how this can be applied to the other taxonomies. Those presented by Dave and Dawson, whilst general in their description could easily encompass higher education specifically due to this reason. As a result there is a broad scope for interpretation and application, as it could be argued for those from Simpson and Harrow. It could be stated that the taxonomy from Ferris and Aziz is overly specific and starts at a level, which is rather simplistic for higher education i.e. recognition and handling of tools. As a result this appears to be a particularly harsh assessment of the other taxonomies and an unconvincing rationale particularly as Simpson's is used within the training of doctors for emergency life support. This is certainly not something that could be classified as a lower level skill which strongly contradicts the claim by Ferris & Aziz, (2005).

It appears that the vague nature of each of the domains would be the rationale behind other criticisms. The taxonomies from Simpson, Harrow, Dave and Dawson are seen as vague and Ferris and Aziz propose a taxonomy that is very specific, the argument as to which is better is the second main criticism. As four out of the five taxonomies listed above are fairly vague in their descriptors, it would be prudent to assume that this is the most applicable way to write a taxonomy (it has certainly been employed effectively for the cognitive and affective domains). This could be so that the taxonomy is transferable to a large number of subject areas, however is the vagueness of the taxonomies, also their downfall. As a result of the vagueness, interpretation by the instructors could result in certain inaccuracies being written into the objectives, due to their understanding or indeed misunderstanding of the headers and the descriptors. Some authors have explicitly sought to apply the taxonomies to particular subject areas, however the vast majority of these are not critiqued or assessed by anyone independent to the paper. As a result, the validity of their interpretation is potentially biased and subjective. Although the taxonomy by Ferris and Aziz is very specific, this does not make it a useful tool for someone outside of their subject area i.e. any practical discipline that does not use tools. As a result, it has clear advantages within the specific context, but is it really feasible to write a specific taxonomy for each subject area, most definitely not. Therefore the adoption of the taxonomy by Ferris and Aziz would not be recommended, as it is overly subject specific. Whilst it does have other literature to support its use within a related field, the subject area is so dissimilar to that of the current study that it would not be a feasible hierarchy for use within the practical rehabilitation setting.

When assessing the four remaining taxonomies it can be seen that each has a distinct hierarchy. Whilst it may be overlapping for some of the taxonomies, all of the various taxonomies in the psychomotor domain describe a progression from simple observation to mastery of physical skills (Kennedy et al., 2007). The mechanism by which this is achieved with language and the level descriptors is the key distinguishing feature between the taxonomies.

Whilst prevalent in the literature, Simpson's appears to encompass too many aspects of the cognitive domain in the lower levels of their taxonomy. Both Simpson and Harrow also do not address one of the key aspects of skill acquisition, which is cited as one of the most powerful learning strategies, namely observational learning (Gog et al., 2008; Hodges et al., 2011; Iacoboni & Dapretto, 2006; Rizzolatti & Craighero, 2004). This final point is the predominant weakness of the domains from Simpson and Harrow within the context of this study. As the current study will utilise a range of practical skills and neither of the taxonomies have a specific way of taking account of this crucial aspect of the learning, this leads to the justification for not employing each of these taxonomies in the current study.

As a result despite the range of available taxonomies highlighted, those from Dave, (1970) and Dawson, (1998) appear to be the most appropriate classifications of the psychomotor domain for the current study. This is due to their acknowledgement of the initial Imitation or Observational level, respectively, from which the rest of the hierarchy is based upon. When comparing and contrasting between these two taxonomies there is very little difference between them. Dave, (1970) does have the distinct advantage of being cited in a far greater number of studies than Dawson, (1998). However very few are empirical studies applied within a context that is similar or indeed relevant to the context of this study. Whilst Dawson's exposure has been limited due to the reasons cited previously, it has been used and adopted for studies that have specific relevance to the current study. Dawson's taxonomy also enables greater differentiation between the levels of learning due to the additional level that they include within their hierarchy. As a result, from the five taxonomies identified, the current study will utilise the psychomotor taxonomy from Dawson, (1998) to complete the learning frame work initially proposed by Bloom and colleagues.

3.6 Progression of Mayer's Multimedia Learning Theory

Following the development of Mayer's multimedia learning theory, Moreno & Mayer (2007) revisited a proportion of their work. This related to interaction in multimedia learning environments and aimed to tackle the challenge of promoting meaningful learning, whilst reducing cognitive load for media such as virtual reality, agent-based and case-based learning environments (Mayer, 2014a). Following the evaluation of previous work, the authors proposed the Cognitive Affective Theory of Learning with Media (CATLM). CATLM is based upon numerous assumptions from cognitive and motivational research such as the dual-coding theory, the theory of working memory, cognitive load theory, motivational factors that may increase of decrease cognitive engagement, metacognitive factors and prior knowledge all of which have been presented up until now (Moreno & Mayer, 2007). When compared to Figure 1, CATLM, depicted in Figure 10 is clearly an expansion of the multimedia learning theory with additional sensory memory components. One in particular, tactile could be seen to have a distinct relevance to the current study and additional interaction with the long-term memory and the role of metacognition illustrated via motivation and affect.

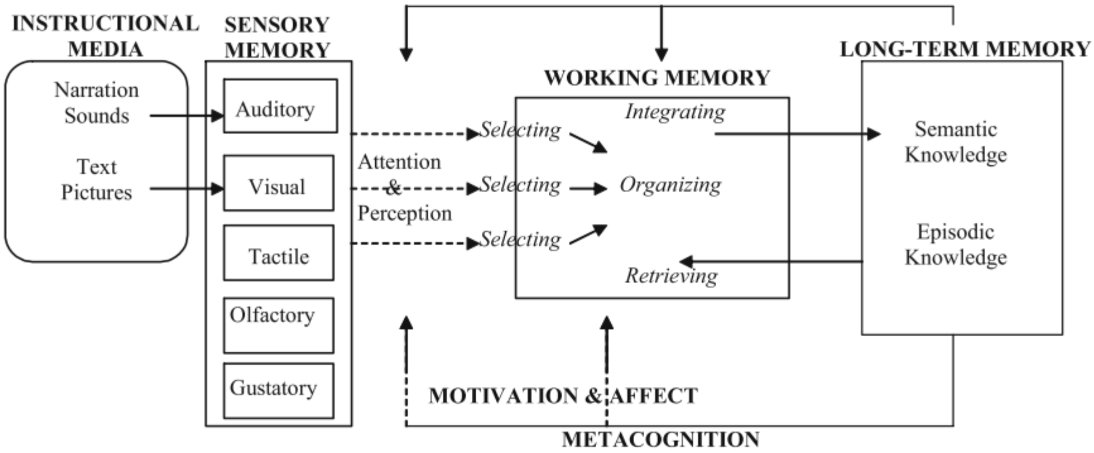

Figure 10 Cognitive affective model of learning from media from Moreno & Mayer (2007).

The authors of CATLM then proposed a series of principles for instructional design; guided activity, reflection, feedback, pacing and pretraining which overlap/replicate the already presented multimedia principles. The concept of emotional design has been well established, however it has not been until recent times that the concept has been researched in an educational setting (Mayer & Estrella, 2014). In a commentary about the role of incorporating motivation into multimedia learning Mayer, (2014a) describes different ways it can be achieved. They cite examples of emotional design (with the use of exciting colours and round face-like shapes), increasing interest (through increasing the appeal of graphics) or through the confusion hypothesis (where posing contrary correct opinions that forces the learner to resolve the contradiction could improve learning). The latter appears to be somewhat counter intuitive and the author states that the results are complicated, but there is evidence that students report that they are less bored, more confused but more engaged when resolving the problem (Mayer, 2014a).

Throughout the commentary it is acknowledged that the use of additive or overly complex instructional material could increase extraneous load and some of the papers have proved that this is the case. It is concluded however, that the motivational features can improve learning by fostering generative learning when they are not distracted or overloaded.

Mayer & Estrella, (2014) present a study were they used enhanced graphics (incorporating two of the three methods detailed in the commentary) to evaluate the effect of emotional design in multimedia instruction. They used a control group with simple black and white graphics, against a group that had graphics that were depicted with human-like features and vibrant colours. Whilst the study is not strictly relevant to the context of this study, it is one of a select few (others are presented within this study relating to case study based learning) that has aimed to address the role of emotions in instructional video. The findings of their study (which have also been partially supported by Plass, Heidig, Hayward, Homer, & Um, (2014)) demonstrated that the enhanced group performed better than the control group. This was seen to be due to the fact that they were more motivated to engage in beneficial cognitive processing, as the graphics were deemed to be more appealing, in agreement with the findings from the commentary. Mayer & Estrella, (2014) explained this with the use of CATLM by stating that the emotional design cues increased motivation, to interpret the information (generative processing) which led to an increase in learning that supported performance when assessed via a comprehension test. The authors identify numerous limitations to their study but the underlying principles are of particular interest for this current study. If it is possible to obtain generative and increased learning from the use of human-like features and vibrant colours (even though this could be seen to be a violation of the coherence principle as the material has increased seductive content), what would the outcome be with the use of humans, to which an emotional attachment (qualified as a normal lecturer-student relationship) and levels of trust that have already been established be? Unfortunately, there appears to be an absence of literature in this specific area. Whilst there has been an extensive amount of research into the role of instructor presence within online learning, the majority of this relates to a live interaction that cannot be facilitated within the context of an instructional video. The results of these studies have shown increased student motivation, satisfaction and cognition which could point towards a potentially beneficial affect relating to the question posed above (Baker, 2010; Ladyshewsky, 2013).

3.7 Instructional Video

The role of the instructional video is growing and whilst there is a plethora of literary papers (ranging from the safe use of wrenches to manipulation of the cervical spine) that investigate the role of an instructional video with a practical task, only a limited proportion quantify the psychomotor component. From the range of studies conducted, a selection relate to the topic of investigation within this proposed study. Salyers, (2007) conducted a quasi-experimental study with nursing students to determine the effects of traditional and web-enhanced instructional approaches that included instructional videos. Whilst their sample size was limited to 36 students the researchers quantified the student's cognitive, psychomotor and satisfactory scores from the two approaches. The study demonstrated that the students in the web-enhanced group achieved significantly higher scores on the final cognitive examination. They also achieved higher scores on the psychomotor skills examination, although not significantly so, but were not as satisfied with the course as their traditional delivery colleagues. The results of the study yield interesting conclusions that support the previous learning presented. This is with the exception of the affective domain that the authors accredit to the technological learning environment, however the fact remains that those who participated in the web-enhanced format did perform better (Salyers, 2007). Another study involving nursing students was conducted by Bloomfield et al., (2010) to assess the efficacy of an instructional hand-washing video when compared to a control group, with a traditional instructional format. The study had a far larger sample size than the previous study with 242 students of mixed gender, age, educational background and first language recruited (Bloomfield et al., 2010). Due to the nature of the task, the study was brief and the follow up measures (knowledge and skill performance) where conducted at 2 and 8 weeks to determine the retention of knowledge. There were no significant differences in the cognitive and psychomotor skill performance scores between the two groups at the 2-week follow-up, with significant differences present at the 8-week follow-up (Bloomfield et al., 2010). Whilst the findings of the study where not as prominent as the study conducted by Salyers, (2007), the findings reveal that the instructional video is a viable alternative for instructing the hand-washing skill. This further supports MMLT effectively, by demonstrating increased retention due to the deeper learning theorised by Mayer & Moreno, (2002). The application of the findings by Bloomfield et al., (2010) are given further educational scope as Ahn, Cho, Shon, Park, & Kang, (2011) conducted a study examining the effect of a 'reminder video' viewed on their mobile phone, on retention of cardiopulmonary resuscitation (CPR) and automated external defibrillation (AED) skills. Seventy-five male students were assigned to the video-reminded or the control group randomly. Those in the experimental group received text messages every two weeks encouraging them to view the 180-second long 'reminder video'. Three months following their initial training, students in the video-reminded group demonstrated significantly more accurate CPR and AED skills than their control counter-parts. As an additional measure, the video-reminder group also demonstrated significantly higher self-assessed confidence and willingness to perform CPR skills than those in the control group (Ahn et al., 2011).

Moore & Smith, (2012) also acknowledge that multimedia is an effective mechanism to teach psychomotor skills in a health care setting. The authors sought to compare the difference between video podcasts and live demonstration for Doctoral Physical Therapy (DPT) students when teaching psychomotor skills. The study used thirty-three DPT students in a 2-week, two topic cross-over design, group A received the lecture and live demonstration in week one, group B received the lecture and video podcasts in week one (transfer skills) with them swapping conditions for week 2 (gait training skills). The students were assessed during week 3 on cognitive and psychomotor knowledge/skills; two experienced practitioners who were blinded to the study assessed them. Although the authors state that eight podcasts were produced for each topic (each between 30-50s in length) there is no specific reference to the use of any instructional guidelines regarding their creation. The findings of their study indicate that the video podcasts were as effective as live demonstration for teaching cognitive and psychomotor skills to DPT students (Moore & Smith, 2012). However the scores obtained by the students were not higher for either topic or test (Moore & Smith, 2012). The authors also undertook a survey of perceptions and an interesting finding was highlighted upon analysis. Students reported that they did not like that they were not able to ask questions during the practice of the skills, when viewing the video podcasts and they were not in favour of the lack of interaction (Moore & Smith, 2012). Duijn et al., (2014) also undertook a study with DPT students (fifty-three students over four cohorts) that were a single blind, two group (online video instruction versus face-to-face teaching) posttest-only, crossover, and experimental design teaching advanced cervical psychomotor skills. The authors sought to teach four skills in total, two skills that involved examination and two that involved treatment of a cervical spine. The author's detail how the videos were formulated under the guidance of faculty experts; however there is no mention of a reference to any instructional design principles regarding their creation. A particular element of the study design, that is noteworthy, is that the same instructor taught all four cohorts and appeared in all of the instructional videos. Student performance of the skills was evaluated using a grading rubric and the same evaluator for each cohort. Whilst this is has potential pitfalls due to the subjectivity of the approach; it appears to be the standardised method of assessment that is accepted. Following analysis of their data Duijn et al., (2014) identified no statistically significant difference between the groups regarding their performance of the four clinical skills at the first assessment point (prior to crossover). However, when the students had received both interventions, each of them increased their levels of performance significantly. No statistically significant difference was identified regarding the order to which the students received the intervention. This confirms that the crossover design is ethically sound and does not disadvantage either group. In conclusion Duijn et al., (2014) demonstrated that the benefits of using both modes of instruction were apparent and supported by the increase in scores obtained following the end of the second intervention. Whilst it is not a direct solution for the negative comments that were identified by Moore & Smith, (2012) it is a fact that could reassure the students, that the use of both approaches will positively impact upon their learning when the subject matter is the same.

Even though it can be seen as a more complex skill, the findings of the study by Stefanidis, Korndorffer, Markley, Sierra, & Scott, (2006) emphasise the need for repetition when learning and maintaining psychomotor skills. Their study investigated the impact of ongoing simulator training on particularly laparoscopic skills. The method the authors employed is exceptionally valid given the reality of the simulator. Whilst it would be unethical for students to practice the psychomotor skills within this study on injured patients for reinforcement, an alternative to simulation is viable. Van Bruwaene, De Win,

& Miserez, (2009) undertook a most interesting study and whilst their sample size is limited, the findings have broad implications. The authors conducted a study aimed to address the difficulty encountered in medical curricula, regarding laproscopic suture (knot-tying) training. In a novel methodology the authors used two balanced groups, with ten senior medical students in each. Each of the students had minimal prior experience with laproscopic suturing before participating in the study. The control group undertook their training by viewing a stepwise video demonstration repeatedly (no specific data is provided regarding the number of times it was played in its entirety) and then practiced hands-on laproscopic suturing, with continuous expert feedback until they achieved the performance objective twice. The performance measure used maybe of great familiarity to those in the medical setting, however for those who are not it adds a level of complexity when evaluating it. The measure was assessed quantitatively (time) and scored via penalty points if the knots slipped; these were increased by a factor of ten and added to the time to form an overall score. A maximal time of 600 second(s) was allowed and students were compared to the mean score obtained by two expert laproscopists over ten trials each. The experimental group began by reading a comprehensive written account of the instructions and watching the same instructional video twice. It is not wholly apparent within the text as to why this group was limited with their initial viewing. The students then practiced in the same way but with access to the video (as many times as they wished) and peer feedback.

The results of their study showed that every student was able to reach the level defined as expert (145s) twice. However those in the experimental group achieved it with fewer trials, than those in the control group, 15.5 (6-25) versus 19 (6-43) respectively. The findings of their study clearly (although not conclusively) demonstrate that video demonstrations and peer feedback can replace expert supervision. This contravenes the previously held premise that skill acquisition is best obtained with expert feedback (Van Bruwaene et al., 2009). This challenges the standard pedagogical approach and provides some evidence, that instructional videos can be a powerful learning tool for highly dexterous skills such as laproscopic suturing with peer feedback. It would have been interesting to see what the combination of expert feedback and continuous viewing of the video could have resulted in. However the implications from a skills training perspective are evident for educationalists, but it is not known if this finding would translate into other subject areas or if the findings can be generalised due to the rather limited sample size that was used within the study.

Another study assessing knot-tying by Nousiainen, Brydges, Backstein, & Dubrowski, (2008) employed a generally more standard methodology. They had 24 students who had no prior knowledge of knot tying and divided them into three groups, after they all viewed an expert-narrated instructional video on instrumented knot tying. Students assigned to group 1 were then able to watch a segmented version of the video once during their practice session. The student's in-group 2 were able to watch an interactive video with identical content, at their own pace between and during practice attempts, the authors define no set limit. The student's in-group 3 had the same access to the video as those in group 2 during the first nine practice attempts. After this they received expert feedback instruction, to overcome a plateau that the authors previously identified when students are learning this skill. By forming these three groups Nousiainen et al., (2008) addressed the combination of variables not covered by Van Bruwaene et al., (2009) and those of segmenting and self-pacing. The findings of their study highlight that the combination of the interactive video and the expert instruction did not improve the development of retention of the knot-tying skill. Those using the segmented video performed as well as those using the interactive self-paced video. The authors identify similar conclusions to Van Bruwaene et al., (2009) and are able to evidence that the combination of instructional video and expert feedback is no better than the use of video alone when either of the multimedia principles of segmenting or self-pacing are employed. The findings from Van Bruwaene et al., (2009) are supported by Attstro et al., (2005) who undertook a similar study on hand washing with dental students. The authors also evaluated the effectiveness between a segmented video and a sequential video by assigning them as the experimental and control groups respectively. Whilst the findings identified no significant difference for the practical outcome, those in the segmented group spent a longer duration viewing each of the videos and achieved a significantly higher score on the written test, further supporting the use of segmenting and self-pacing.

The overall findings of the studies above are also supported by others in a medical setting from Xeroulis et al., (2007), Jowett, LeBlanc, Xeroulis, MacRae, & Dubrowski, (2007) Willis et al., (2012) for knot-tying, by Chenkin, Lee, Huynh, & Bandiera, (2008) for ultrasound guided vascular access training and by Öztürk & Dinç, (2014) for urinary catheterisation knowledge and skills. All of the above studies demonstrate that the use of instructional videos resulted in an equal or superior level, when compared to traditional educational techniques for a range of complex psychomotor skills.

Ford et al., (2005) conducted a study that was closely aligned to the current study however they used a smaller population. They also had a shorter experimental duration and smaller range of content, which has positives and negatives regarding the feasibility of operationalisation, versus a reduction in data obtained. The authors recruited forty-three students over two consecutive cohorts. They randomly assigned them to one of four experimental groups based upon their previous practical examinations results to ensure a balanced level of ability in each group. The students would learn four musculoskeletal tests that they had not previously encountered. This would be through Computer Aided Instruction (CAI), live demonstration with written text and the same practitioner from the video, independent textbook learning or no educational intervention. Students in the CAI group were able to view any of the four videos twice within their session, although no specific information regarding number of live demonstrations is provided. The independent learning group also had access to the same written text and two still photos from two planes of view. All groups had up to 25 minutes to learn the content from their allocated resources. The students were assessed via a 20-question examination evaluating knowledge and by a practical examination, assessed by a grading rubric. Rather expectedly those receiving an educational outcome significantly outperformed those in the control group. CAI was as effective as live demonstration and both were more effective than learning from written text with still photos. Interestingly their findings regarding the CAI are somewhat unexpected when you breakdown their methodology. The CAI intervention was composed of a video on the left side of the screen and bullet point, PowerPoint slides on the right side, which would according to Mayer, (2014) lead to split-attention. As a result, this contradicts the findings of the principle somewhat and it did result in the CAI group obtaining scores that were slightly lower than the live demonstration group. This potentially in theory could be reversed if the CAI was redesigned.

The findings of Ford et al., (2005) were supported by Smith, Jones, Cavanaugh, Venn, & Wilson, (2006) who adopted a similar methodology but with only two groups. This was composed of an experimental group who had access to an instructional CD of clinical orthopaedic techniques, with instructor feedback and a control group receiving live demonstration of clinical orthopaedic techniques. The study was conducted over a two-year period and recruited forty-five participants in total. The context of the study is particularly pertinent to that of the current study. Each group in phase one of the study were taught orthopaedic techniques for the knee, and then they were taught orthopaedic skills for the foot/ankle. Furthermore the amount of videos that they had access to is related as the study by Ford et al., (2005) only used four videos, whereas this study gave the students access to 23 (knee) and 20 (ankle) videos. Each of the videos were approximately 30 seconds in duration, which is more applicable to video platforms such as YouTube or iTunes U were a multitude of videos are presented. Whilst the authors present a detailed breakdown of how the study was implemented, they have made the procedure overly complex. They used a different joint for comparison between live demonstration and the instructional CD's, which has the potential to identify a difference that is not truly attributable to instructional method, but solely attributable to the differing nature of the knee and foot/ankle. The findings of the study will however have a great bearing on the current study, as it proposes to use differing joints every three weeks throughout the duration of each academic year. The students in the study by Smith et al., (2006) were assessed via a written and practical examination and after a complex interpretation of their results, it was revealed that there were no significant differences in cognitive performance between instructional strategies. There was also no significant difference between practical knee scores, but there was a significant difference in practical foot/ankle scores in favour of the instructional CD. Whilst the findings are in line with other literature, the study does have several less than desirable methodological issues. The students from the first year of data collection conducted the experiment in semester one of their program, but those in the second year of data collection conducted the experiment in semester two, after they had already been exposed to some related material in semester one. Prior knowledge has been shown to be a significant factor in improving performance, as demonstrated by the pre-training principle with an effect size of 0.75 (Mayer, 2014). In theory though this should have balanced both of the groups out proportionally. Therefore there is no clear explanation, either provided by the authors or observable from an informed reader as to why this should have occurred. This highlights one potential limitation when comparing differing material as part of the research design that needs to be considered by the current study. Overall however, the study identified no difference or a positive difference in favour of instructional videos for cognitive or psychomotor skills when teaching clinical orthopaedic techniques.

Other studies undertaken identifying the impact of instructional video have addressed other areas such as student perception. One such study by Kelly, Lyng, McGrath, & Cannon, (2009) sought to evaluate student attitudes to instructional videos when being taught clinical skills. The author's findings demonstrate that students are in favour of them due to the flexibility that they provide, but the instructional videos should only be used in conjunction with, rather than as a substitution for live demonstration.

McCutcheon, Lohan, Traynor, & Martin, (2014) undertook a systematic review with the aims of evaluating the impact that online or blended learning, would have versus face-to-face learning, for clinical skills in undergraduate nursing students. The findings of their review for knowledge identified that in seven out of thirteen studies that assessed knowledge, following an online learning modality found highly significant results indicating higher levels of knowledge. In contrast only two reported higher levels of knowledge following face-to-face delivery and five reported no significant difference. The findings of their review for clinical skill performance identified six from thirteen that established a significant improvement following online learning; the remaining studies did not identify a significant difference. Overall ten of the thirteen studies identified either no difference or an improvement in clinical skills and knowledge when online learning was used. Despite their rigorous search strategy, four of the included papers provided an imbalance of educational instruction between the control and intervention groups. This imbalance was predominantly with supplementary information; nevertheless there is the potential for bias as a result. The conclusions of their review is that students obtained a higher or similar level of clinical skill when evaluated against traditional teaching techniques (McCutcheon et al., 2014). However, the conclusions of a systematic review conducted by Veneri, (2011) supported by the work of medical and allied health educators state that, the usage and effectiveness of computer assisted learning, particularly with regard to higher order thinking and psychomotor skills requires further research.

As a result of this review it is clear that, at the very least instructional videos can be as effective as traditional educational methodologies on the whole. Furthermore according to Smith et al., (2011) it is evident that instructional multimedia for cognitive and psychomotor learning, within higher education appears to be most effective when it is delivered in addition to traditional classroom instruction, an aspect reinforced by the findings of Duijn et al., (2014). In order to establish why instructional videos are effective Duijn et al., (2014) stated that it was shown that enabling the student to repeat and review the online activity, as many times as necessary, whenever they chose, at their own pace, enhanced their acquisition of knowledge and skill performance. Whilst it appears that the review is biased towards positive findings, there is a distinct lack of literature showing that instructional videos are not as effective as traditional instructional methodologies when teaching cognitive or psychomotor skills. The statistics employed within each of the studies above may well be the mechanism for this skewing, as all of them analysed their data with the use of null hypothesis testing. This approach may not have been sensitive enough to detect a significant change between instructional approaches. Therefore, a non-significant difference resulted in a positive outcome for instructional videos, which in most other contexts would not prove to be so. In summary it appears that instructional videos in a wide variety of educational contexts, combinations and subject settings are an effective medium to convey cognitive and psychomotor knowledge. It is also apparent that as a result they are being widely employed throughout educational contexts internationally, due to the potential they have for advancing learning, being time efficient and enabling flexible learning. Each of these traits would be beneficial in their own right and whilst the array of topics is wide, each of them has focused on a test or skill that can be assessed as right or wrong, which could directly lead to a particular ontological approach over any other. Many of the studies have also separated the knowledge and practical aspects. This in most settings is applicable, however within the context of the current study there is a distinct overlap/interaction between the cognitive and psychomotor domains. Interestingly none of the studies cited above mentioned the affective domain in any way, which, given its relevance to the medical setting is surprising. As a result, the current study will be unique in its approach by investigating the efficacy of instructional videos across all three domains.

3.8 Cognitive, Affective and Psychomotor Assessment

As detailed above (in section 3.7) there are well-established methods to assess knowledge within learners and the most commonly used from the studies mentioned is a written examination. The advantages of this are well established in an educational context, and this is why this method is the cornerstone for all of the major landmarks within a child's education. The use of grades within University level education has received particular criticism, as their use can have little a relationship with what the students have learnt. They may even not be a reliable way to assess learning particularly when used to test performances, as it has been noted that even the same assessor will not assign grades in a consistent manner (Rovai et al., 2009).

The affective domain in contrast to the cognitive, has received very little attention regarding its assessment. Those that have attempted to address this typically utilise self-reporting measures, this is due to the inherent limitations of the domain being internal in its qualities. Chi et al., (2014) state that they are the first to publish a study evaluating cognitive and affective learning outcomes within a technological study, comparing the use of videos against traditional paper based approaches. Students in their study evaluated themselves on a series of statements, ten for the cognitive domain and two for the affective domain. The findings of their study identified a significant difference across the board for each of the statements emphasising the efficacy of video based case studies. However, when compared to paper based case studies this is not overly surprising. Whilst Chi et al., (2014) may have been the first to assess the affective domain in a technological setting, the issue of assessing the affective domain, in particular attitudes has been a difficult topic for a considerable amount of time. The assessment of the affective domain has always been seen as less important than the assessment of knowledge, cognitive or psychomotor skills (Cate, 2000). The role of attitudes within the medical context has been introduced into medical examinations that typically assess all bar the affective domain. Cate, (2000) published a paper detailing how they summatively assessed attitudes at the University of Amsterdam with the hope of inspiring others. Assessing attitudes summatively poses a range of inherent difficulties due to the various ways in which they are communicated (verbal and non-verbal). As a result, their criteria for assessment includes courtesy, respect, empathy, sense of responsibility and engagement which within defined clinical scenarios may well be able to be assessed, though with some degree of variability. However the applicability of this approach in a research context to anything other than a strictly controlled scenario may not be as effective (Cate, 2000).

The psychomotor domain has typically been assessed with the use of a grading rubric or score card, which experts use to evaluate aspects of a student's performance against and assign a grade to each section, which cumulatively will provide an overall grade (Duijn et al., 2014b; Ford et al., 2005). Jelovsek, Kow, & Diwadkar, (2013) undertook a systematic review of the methodologies used to assess psychomotor skills in medical trainee's.

They stated that the assessment typically involves two components, the knowledge the student is using to learn from and the students' performance (Jelovsek et al., 2013). Interestingly the author's state that if the student successfully 'shows how' or 'does' the required skill, then this demonstrates that the knowledge is in place and there is no need to assess knowledge separately. The authors reviewed over 4000 articles and concluded that whilst there are numerous tools available for the assessment of psychomotor skills, for medical students in the live setting, there is only an evidence-base for a few of them. Each of those identified are very specific to their particular aspect of the medical setting, therefore the applicability of them to any other context would severely impact upon their reliability and validity.

Whilst there are few references to it in literature the variability of this form of assessment is generally accepted as an inherent limitation of the assessment technique, that most educational institutions overcome with the use of double marking and internal or external moderation. However, assessment reliability and validity are of the utmost importance, especially when the assessment instrument is used in important decision making settings such as examination competence (Jelovsek et al., 2013).

Rovai et al., (2009) are the only authors that sought to develop an instrument to assess the cognitive, affective and psychomotor domains. As, in theory an instrument that measures all three domains of learning would be beneficial for researchers and educators. Whilst the study had distinct promise, the authors decided to formulate a self-report perceived learning scale that would only be of use in a formative setting. As a result it is apparent that the summative assessment of the three domains will present a certain amount of problems. It may have inherent limitations that may not be easily overcome, although they could be standardised, for internal validity. This mixture of qualitative and quantitative assessment has been referred to as ‘triangulation’, which is derived from a navigational concept. Fotheringham, (2010) conducted a review of its theory, use and methodology concluding that the utilisation of ‘triangulation’ unfortunately cannot guarantee the validity or reliability of the assessment. This as a result creates as many questions as it resolves, which does not serve to further the use of this type of assessment unfortunately. This therefore is an aspect that will present a great challenge to the current study and one that needs to be considered with due diligence.

3.9 Duration of instructional video

There is one omission from the majority of the instructional video studies above; duration. Whilst it is not directly referred to in the instructional principles from Mayer, (2014b), the guidelines from Meij & Meij, (2013) recommend to keep the video short. The authors cite conference proceedings that are no longer available, from Plaisant & Shneiderman, (2005) detailing that the optimal length to ensure engagement and to facilitate learning from the material is 15 to 60 seconds. This is certainly a duration which may not be feasible to the vast majority of content and therefore the authors recommend the use of segmenting and signaling to facilitate the use of shorter durations (Meij & Meij, 2013). The claim from

Plaisant & Shneiderman, (2005) has some support from McCombs & Liu, (2007) who undertook an extensive podcasting survey which culminated in a series of recommendations for the use of podcasts. The authors concluded that video podcasts should be no longer than 15 minutes in duration. This was founded on the comments from one student that stated that their attention span faded towards the end of long podcasts, in contrast to face-to-face lectures (McCombs & Liu, 2007). The shorter duration is also supported by Chen, (2009) and Chan, (2010) as they interpret the work of Prensky. They state that the current generations of learners (who are deemed to be digital natives) have shorter attention spans and are impatient, due to the numerous enhancements found within modern day entertainment. This is an aspect validated by van Bruggen, (2005) who states that the additional visuals maybe distracting to learners with limited attention spans. This was furthered by Caspi, Gorsky, & Privman (2005) who identified that most of the students in their study had to give the videos more attention than they would when using a book. This aspect should therefore be considered when designing instructional videos. As a result of the limited literature available within the context of the current study, there is a clear recommendation for the use of shorter duration videos, although very little empirical evidence. When you compare this to literature regarding traditional lectures, this fits rather effectively as it is well documented that the attention span of learners in a traditional lecture tends to decrease after approximately 20 minutes (Chilwant, 2012; Folley, Wilkinson, & Thomson, 2009). Therefore, this could be considered the upper limit for an instructional video if designed appropriately with consideration to the multimedia principles to engage the learners wherever possible. However this duration would tend to be inadvisable and instructional videos should be as short as possible. If the content area is large the videos should be divided into smaller videos (as they will be within the context of this study) or segmented with either active or passive pauses to facilitate learning. The current study will be able to contribute to the minimal evidence-base for the duration of instructional videos, as two differing durations will be compared.

3.10 Group vs. Individual Learning

In addition to the general learning theories another dimension to consider is that of group versus individual learning as it separates the aims of the study (Hill, 1982; Laughlin, Hatch, Silver, & Boh, 2006; Lou, Abrami, & d'Apollonia, 2001). Although the students will be performing the tasks as a group, they will be processing the knowledge and be able to be tracked individually throughout the duration of each data collection period.

There are a substantial number of studies suggesting that collaborative learning improves students' achievements compared to working alone (Kirschner et al., 2009). This has several implications from a performance perspective as if the adage 'two heads are better than one' (Hill, 1982) is true, then a group based activity may not be the most conducive learning task for individuals. However, Kirschner, Ayres, & Chandler, (2011) stated that when the material being learnt is low in complexity, individual learning is superior, but when it is complex group learning is more effective. This view point is further supported by Laughlin et al., (2006) who state that cooperative groups of three to six members perform better than independent individuals on a wide range of problems, and that previous research suggests that as the group size increases, the performance also improves. Lou et al., (2001) and Laughlin et al., (2006) are in agreement that when students worked in two-person groups, there was a positive effect on both group and individual achievement i.e. the pairs did not differ from the best individuals.

Kirschner, Paas, & Kirschner, (2009) undertook a review of available literature with the aim of comparing the efficacy of individual learning environments, against group learning environments. As part of their introduction the authors state that group learning is generally considered to be more effective, when the learning outcomes of the number of learners in the group is higher, than the total of the learning outcomes of a number of comparable individuals. As a result learning can be seen to be more efficient if the learning outcomes are attained with less mental effort (Kirschner et al., 2009).

Throughout the review, the authors present the findings of many authors stating that the benefits of group learning can be seen as:

- It leads to greater active engagement in the learning process.
- Information is retained for a longer period of time.
- Higher-order skills are fostered.

- Engage in self-directed learning to a greater extent.
- Learn how to negotiate and verbalise themselves more effectively and coherently.
- Provide mutual support for each other when learning.

Whilst the negatives to group learning can be seen as:

- A resulting failure in underusing basic information.
- Over committing resources to failing projects.
- Ineffective sharing of information by group members.
- Production blocking.
- That there is no guarantee that effective collaboration that will lead to positive learning.

(Kirschner et al., 2009)

In an attempt to address the different arguments and the various supposed advantages/disadvantages to either approach, Retnowati, Ayres, & Sweller, (2010) conducted a complex study to examine the efficacy of worked examples, when used by individuals and in group work settings. The study had four experimental groups using a 2 (problem-solving vs. worked examples) x 2 (individual vs. group study) design for one hundred and seven students to teach geometry problems. Whilst the results indicated that the worked example groups scored higher than the problem-solving groups, regardless of the learning setting they were partaking in, for each and every performance measure (Retnowati et al., 2010). The authors specific hypothesis that the group work setting would be more beneficial for students than the individual settings was not confirmed (Retnowati et al., 2010). The study even included transfer problems, specifically to give the collaborative groups the best opportunity to demonstrate an advantage, which raises some additional questions about the potential bias of the authors.

Kirschner et al., (2009) states that there are some fundamental flaws that impede the analysis of groups that are found in most systematic design processes, namely that they are overly focused on the attainment of learning outcomes for individuals rather than the group. Therefore this underlying flaw will impact on the findings of all studies, which compare groups to individuals when they are only assessed in an individual setting. As a result, the question arises as to how groups could be assessed, or even how groups and then individuals can be assessed. From the brief literary overview above it is clear (and echoed by a number of studies) that a greater focus on how students learn in groups, with stricter methodologies needs to be undertaken focusing on the cognitive learning theory (Kirschner, Paas, & Kirschner, 2011). Therefore, there is little conclusive evidence as to whether groups or individuals will perform better in the context of the current study. Whilst group size maybe of critical importance there are many factors that may impede group work, so the current study will aim to contribute to this field of enquiry. It will do so by adopting a novel approach, whereby the group is assessed as a group and the scores extrapolated to individuals to compare the learning environments from a different perspective.

3.11 Effect Size

In the vast majority of scientific subject areas the almost universal approach when using inferential statistics has been to employ the null hypothesis significance test (NHST), typically the researcher would use a statistical package to produce a *p* value for a particular outcome statistic (Batterham & Hopkins, 2006). However NHST's do not provide researchers with two crucial pieces of information a) the magnitude of an effect of interest and b) the precision of the estimate of the effect (Nakagawa & Cuthill, 2007). As is commonly known, the use of NHST's is exceedingly commonplace within quantitative research; however three prominent statisticians have been very clear about their opinions of NHST's:

> "Statistical significance is the least interesting thing about the results. You should describe the results in terms of measures of magnitude –not just, does a treatment affect people, but how much does it affect them."
>
> Gene V. Glass

> "The primary product of a research inquiry is one or more measures of effect size, not *P* values."
>
> Jacob Cohen

(Sullivan & Feinn, 2012)

> "Null-hypothesis significance testing is surely the most bone-headedly misguided procedure ever institutionalised in the rote training of science students."
>
> (Rozeboom, 1997)

The opening lines from (Sullivan & Feinn, 2012) and the quotation from (Rozeboom, 1997) are prime examples of the movement supporting the use of effect sizes. Fritz, Morris, & Richler, (2012) state that effect sizes allow researchers to move away from the simple identification of statistical significance and toward a more generally interpretable, quantitative description of the size of an effect. There are multiple effect size statistics available that include Cohen's d, Hedges' *g*, and Glass's *d* and Δ (Fritz et al., 2012). From these two the most common standardised mean difference statistics used are Hedges *g* and Cohen's *d* (Durlak, 2009).

The use of effect sizes as a result is becoming increasingly commonplace, major organisations such as the American Psychological Association (APA) recommend that researchers should always provide some effect size estimate when reporting a *p* value as it is essential to good practice (Durlak, 2009; McGrath & Meyer, 2006).

In essence effect size is the main finding of a quantitative study (Sullivan & Feinn, 2012) and a non-significant result (p>0.05) does not necessarily mean that there is no worthwhile effect from the study (Batterham & Hopkins, 2006).

An effect size of d = 1.0 indicates an increase of one standard deviation on the outcome, Hattie, (2008) contextualises this by stating that it is equivalent to advancing a child's learning by two-three years, improving the rate of learning by 50% or a correlation between a variable and a form or achievement of roughly r = 0.50. The authors continue to apply the magnitude of the effect, by stating that the introduction of a new regime that had an effect size of 1.0 would mean that generally, on average, the learners that received the treatment would outperform 84% of the learners that did not receive the treatment (Hattie, 2008). Whilst the probability of achieving an effect size of 1.0 is relatively low and would indicate a dramatic difference, substantial effects can be noted at smaller effect sizes. Even a median effect size of 0.41 can be seen to have an educationally important impact as a fair proportion of those who received the intervention would outperform those who did not (Hattie, 2008; Mayer, 2014).

Effect sizes describe the observed effects; effects that are large but non-significant may suggest further research with greater power (Fritz et al., 2012). Whereas effects that are trivially small but nevertheless significant, because of large sample sizes can warn researchers against possibly overvaluing the observed effect (Fritz et al., 2012). Whilst the discussion centres around sample sizes, it is worth noting that the findings of (Weisburd, Petrosino, & Mason, 1993) found an inconsistent relationship between sample size and statistical power leading (Sherman, 2007) to term the discovery the Weisburd paradox (Nelson, Wooditch, & Dario, 2014). The paradox has been explained by Welsh & Farrington, (2012) in the following way; if all things were equal, the use of larger samples should in theory provide greater statistical power and make studies better at detecting treatment effects. However, it is generally found in a real life setting that not all things are equal and larger samples can often reduce the statistical power of the study, by increasing the heterogeneity of samples. This heterogeneity in turn increases variance, which reduces statistical power. These findings go against the widely held belief among researchers that larger sample sizes will equate to greater statistical power (Sherman, 2007). The finding is not limited to criminology as (Nelson et al., 2014) cites examples showing that similar findings have been identified in medical and educational research, hence the reason for its inclusion within this section. As a result, the interpretation of the calculated effect size must be considered alongside the sample size, failure to do so could lead to the problems described above.

3.11.1 Calculation

There are many ways to calculate effect sizes with online calculators (Becker, 2000; Ellis, 2010) and macro-enabled spreadsheets (Hopkins, 2007) available to assist the researcher. Depending upon your personal preference will determine which you favour as a researcher, regardless of choice an understanding of the equations is desirable, so that the correct equation can be used and interpreted, for the data set to be analysed appropriately (Fritz et al., 2012). Hopkins, (2015) provides a macro-enabled spreadsheet on their website for use with supporting documentation to ensure correct calculation of effect sizes. Given the author has been at the forefront of promoting effect sizes the rationale behind utilising this spreadsheet is overwhelming.

3.11.2 Effect Size Correction for Small Samples

Effect sizes as with other inferential tests are closely linked to sample size and whilst an increase in sample size is linked with a theoretical increase in power, the inverse can also be stated. When the sample size is less than ten (or there are variations in sample sizes) the use of Cohen's *d* is not recommended and a correction for *d* statistic bias should be used (Hedges & Olkin, 1985). It has been demonstrated that when the sample size is less than twenty (or less than ten in each group) the *d* statistic can show an upward bias that is relatively large in size (Nakagawa & Cuthill, 2007).

The implementation of Hedge's *g* is increasing within inferential analysis and is becoming accepted as a method of calculating an unbiased *d* statistic (Adesope & Nesbit, 2012; Durlak, 2009; Einspruch, Lynch, Aufderheide, Nichol, & Becker, 2007; Garamszegi, 2006; McGrath & Meyer, 2006; Zakzanis, 2001). However, despite the fact that most researchers will often calculate *g* rather than *d* as it is more convenient, the former is very rarely presented and it appears that *d* may well be used as a generic term for effect size currently (Fritz et al., 2012). Fritz et al., (2012) justify their comment by stating that, for most reasonably sized samples the relative difference between the methods of calculation for Cohen's *d* and Hedges' *g* will be very small. This could be seen as a reason as to why *d* is used as a generic term as it is the term that the majority of the scientific community are familiar with. When reporting the statistic, the minimal difference could lead to confusion due to the dominance of one letter being associated with the reporting of the statistic.

3.11.3 Magnitude-Based Inference Analysis

In the last decade magnitude-based inferences have emerged, been advanced and are being promoted within sport science as a novel method of analysing and providing more meaningful interpretation of data sets (Drinkwater, Pyne, & McKenna, 2008; Welsh & Knight, 2014). Magnitude-based inferences are described as being based on a calculated confidence interval for the quantity of interest, which is then categorised with the use of probability classifications into a three-level scale of magnitude: substantially positive, trivial and substantially negative (Batterham & Hopkins, 2006; Welsh & Knight, 2014). The application of the three-levels can be seen in Figure 10, which demonstrates the range from almost certainly beneficial, which is the most positive finding to very likely harmful. The illustration also provides examples of findings that would be classified as unclear due to the large spread of the probabilities related to the confidence interval (Batterham & Hopkins, 2006; Fritz et al., 2012; Hopkins, 2007). A distinct advantage of the approach is that a wide range of the population can easily interpret the findings as probabilities. Despite the apparent advantages Welsh & Knight, (2014) state that it has so far attracted little scrutiny from statisticians and cite comments from only one known source.

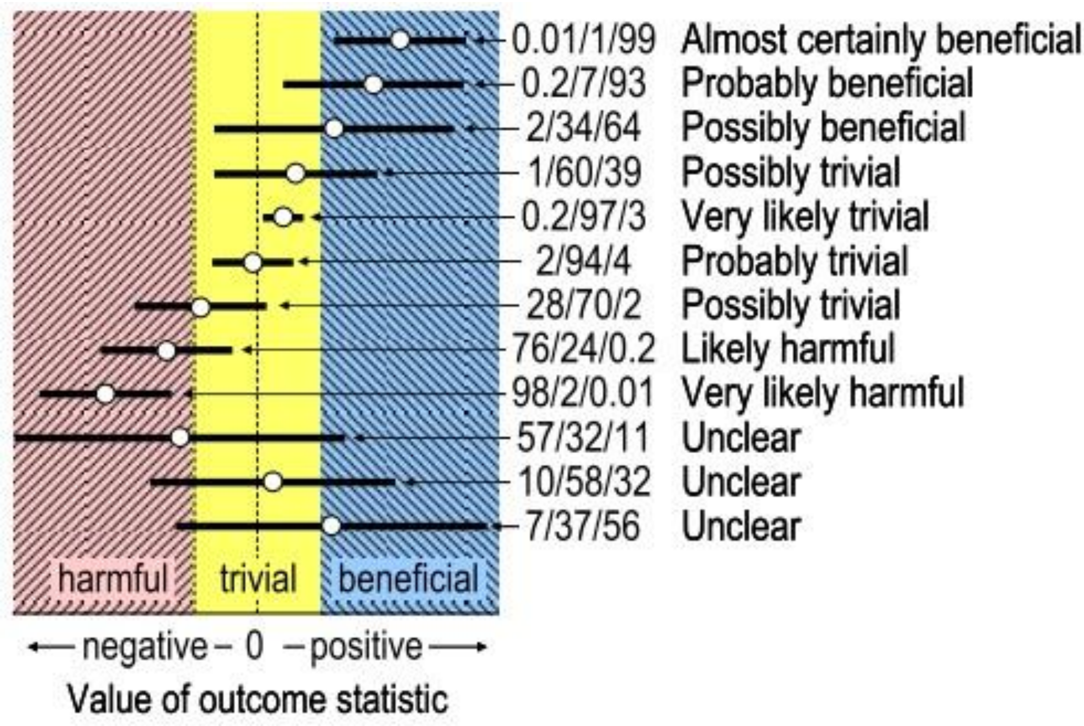

Figure 11 Magnitude-Based Inferences from Batterham & Hopkins, (2006)

3.12 Summary

Throughout the literature review some common aspects have been conveyed with a large amount of supporting evidence, principally these are that:

- Multimedia instruction whether that involves video or not is fundamentally more effective than a single method of teaching due to the dual coding and cognitive load theory.
- Observational learning is one of the most powerful strategies available to humans.
- Instructional video is becoming increasingly prevalent within higher education for a wide range of subjects and skills.

- There are a multitude of principles that should be considered/adhered to when designing instructional videos to ensure they are effective.
- Theoretically, instructional videos are aligned with behaviourist models of learning that are considered lower level; however there is a growing body of evidence that is starting to align it with constructivism.
- The underlying mechanisms that students learn by can be aligned to specific taxonomies for development and assessment of curricula and students respectively.
- The methods of assessing the cognitive domain are well established but assessing the affective and psychomotor domains is less conclusive.
- When working in groups there should theoretically be an increase in learning, but this is currently up for debate when related to instructional videos.
- Effect size statistics enable a greater and more effective interpretation of data sets than the traditional NHST's.

As a result of the literature review the background of the study and a broad range of literary sources have been introduced, which can be used to provide an effective comparison of the current study too. The current study will be informed by many of the methodological aspects such as the multimedia instructional design principles, the methods of assessment linked to Bloom's and others, the three main taxonomies of learning and the use of effective analysis to fully apply the findings of the current study, whatever they may be.

Chapter 4: Method (All years of data collection)

4.1 Research questions

The principle aim of the current study is to evaluate the effectiveness of instructional videos when used as an adjunct to traditional teaching methods, as a result the specific questions will be:

1. What effect do instructional videos have on learner's cognitive, affective and psychomotor abilities when assessed formatively for a practical task?

2. Are instructional videos suitable for use by groups when learning cognitive, affective and psychomotor skills, when assessed formatively for a practical task?

3. Are instructional videos beneficial for an individual's learning when being taught as part of a group, when assessed formatively?

4. Is a particular duration of instructional video superior to another when being used to teach cognitive, affective and psychomotor skills?

4.2 Ontological Framework

Natural philosophers throughout history have, used observation in various ways alongside their theories to support their viewpoint, to convince and to win arguments (Shadish, Cook, & Campbell, 2002). This has been an approach that has been repeatedly employed, tested and scrutinised across a range of scientific areas. However, the validity of observation has to be questioned due to the inherent nature of bias potentially affecting observers and their observations in one way or another.

Observations can be seen as objective or subjective (depending upon the observer's perspective or intentions), or to put it another way quantitative or qualitative. This has resulted in a crude assessment of the meanings of quantitative and qualitative in an attempt to distinguish between the two approaches (Howe, 1988). Generally, how they are measured and the levels of data that they create are often seen to define the two approaches. Typically, this means qualitative is seen as producing categorical data and quantitative as producing ordinal, interval or ratio data (Vincent, 2005). If viewed from an ontological sense however, all data is qualitative if it incorporates values, beliefs and intentions (intentionalist) and it is quantitative if it is non-intentionalist (Howe, 1988). This results in four kinds of data that are in theory possible relative to the specific context of the person who is viewing the data. This, therefore make further arguments about the distinctions of the two approaches more complex from a philosophical perspective. As a result, a pragmatic approach should be one that that simplifies as many of the potential philosophical conundrums as possible. For this thesis there is a clear focus on understanding causation, without necessarily understanding all the mechanisms of the causal process. This would seem to be a more valid approach to address the nature of question that the current study is aiming answer. Throughout the last few decades social scientists have given greater attention to particular methodological issues related to causation (Imai, Keele, Tingley, & Yamamoto, 2011). Causation is an encompassing term for Cause, Effect and Casual Relationships within an experimental perspective (Shadish et al., 2002). Although the debates around the importance of causation are somewhat analogous with the current application of randomised controlled trials (RCT) in education and related to the ongoing attempts to redefine the concept of an experiment (Scriven, 2008). As a result, whilst there are some generally agreed on definitions for each, the lack of absolute distinctions for each concept results in a complex archetype for any researcher.

For the current study the concept of causation is particularly relevant as it is aiming to understand the impact of videos on learner's cognitive, affective and psychomotor abilities in differing contexts. Given the study is deployed within a formal educational setting there is the a need to identify data which are as objective as possible within the limits of a small-scale doctoral inquiry. The approach will produce quantitative data that is non-intentionalist as it could have varied interpretations, particularly from the perspectives of the students. The analysis of the data aims to validate the scientific aims of the current study in a broadly positivist approach (Moore, 1985), whilst acknowledging the limitations of this paradigm. While other philosophical systems are available, or and a mixed methods approach would provide an alternative perspective, the current study seeks to confirm or deny the research aims as objectively as possible, given the educational and subject specific setting outlined above (Onwuegbuzie, 2000).

4.3 Ethics

Durham Universities School of Education and Teesside Universities Research Ethics Committees granted ethical clearance for the data collection period, between October 2010 and May 2013. The former as the lead researchers supervising institution; the latter as the institution where the research would be conducted. There were no ethical concerns from either institution, as the lead researcher was able to demonstrate that no students would be disadvantaged during the undertaking of the study (British Educational Research Association, 2004). All participants were fully informed about the study, and were asked to voluntarily complete a written consent form prior to participation in accordance with British Educational Research Association (BERA) guidelines (2004). Each participant was informed of their right to withdraw and their right to obtain a copy of the results following the completion of the study.

4.4 Procedure

The data for the thesis was collected over a period of three consecutive academic years. The research paradigm had a minor amendment between years 1 and 2, with a more substantial modification implemented for the final year of data collection.

For each of the academic years the data collection adopted a longitudinal, blinded, crossover design and took place over an 18-week period. The period included standard holidays and days when the University was closed due to severe weather conditions. However, these occurred at differing times within the period for each of the academic years and were thereby impossible to be accounted for. At no point were sessions missed due to any of these foreseen or unforeseen events.

The annual format of the module, after the initial 6 weeks of pitch side management and an introduction to rehabilitation theory, requires the students to produce rehabilitation sessions in groups. The students formulate the sessions for a specific peripheral joint (Ankle, Knee, Hip, Wrist, Elbow and Hip) at a particular stage of the rehabilitation continuum (Early, Intermediate and Late). The students have one week to devise the sessions, traditionally relying on standard resources supplied by the module leader (lecture notes, textbooks and journals). In the lesson the students will conduct the sessions and gain formative feedback from their lecturer and their peers.

Each 18-week period was divided into six, three-week blocks (Early, Intermediate and Late) to ensure parity between the times given to each peripheral joint, across all of the academic years. The sequence of the peripheral joints remained consistent across all three of the years of data collection.

4.5 Instructional Videos

Two groups of instructional videos where created that are both based upon the work of Smith, (1998) which forms the foundation of the teaching structure for the rehabilitation module. Smith, (1998) separates the rehabilitation process into three distinct sections within the rehabilitation continuum: Early, Intermediate and Late. This is used to provide examples of exercises that are appropriate for that stage and the criteria that need to be achieved in order to progress through the continuum. The continuum is designed to provide a safe progressive route for each patient, which can be individualised at any stage to ensure efficacy throughout. Each of the exercises was supported with further literature, to provide up to date definitions and information to foster an evidence-based approach.

Twenty-five short duration videos ranging from 55 to 118 seconds (1:58 minutes) were created to provide one condition for the studies independent variable. Five previously recorded and produced longer duration videos ranging from 629 seconds (10:29 minutes) to 1094 seconds (18:14 minutes) were used for the second of three conditions. The short duration videos were designed in accordance with the principles for instructional design set out by Mayer, (2008) examples of which can be seen in Figures 12 and 13. The long duration videos were produced prior to these criteria being applied and adhere to all except the redundancy and modality principle. On-screen text (Figure 15) was used in conjunction with the narration, please see Figure 16 which depicts the alignment of the videos with the multimedia principles from

Mayer, (2014). It was not deemed necessary or feasible (as the original footage was not available) to reshoot and reproduce the five long videos for the project. Additionally this would provide a good basis for comparisons against the short videos, which as can be seen in Figure 16, adhere to all the principles that they can.

Figure 12 Example one from one of the short videos

Figure 13 Example two from one of the short videos

Figure 14 Example one from one of the long videos

Figure 15 Example two from one of the long videos

Multimedia Principle	Long Videos	Short Videos
Coherence Principle	Yes – although the background is not strictly controlled (see figures 14 and 15) for examples of extraneous background content.	Yes (see figures 12 and 13) only essential background material remained.
Signaling Principle	Yes – Cues are provided in both that highlight key principles and structure. This is most evident in the short videos as the field of vision is reduced (see figures 12 and 13).	
Redundancy Principle	No – on screen text and narration is used with pictorial animation (see figure 15).	Yes – the same information is presented visually and audibly.
Split-Attention Principle	Yes – narration and pictorial representation assimilated for both.	
Spatial Contiguity Principle	Yes – text is overlaid or aligned as closely as possible to the pictorial aspects.	No – text is not used when pictorial animation is present, when text is used it is presented centrally in the field of vision with a solid background.
Temporal Contiguity Principle	Yes – relevant text is presented with corresponding pictorial components whenever appropriate.	
Segmenting Principle	No – at the time of inception the use of segmenting was unknown.	Not applicable as each video is short in duration.
Pretraining Principle	Yes – the terminology used in the videos is taught to the students prior to their viewing of the videos.	
Modality Principle	No - due to violation of redundancy principle above.	Yes – narration and pictorial components are used
Multimedia Principle	Yes – multiple sources are used in both video types.	
Personalisation Principle	Yes – a conversational style is used wherever possible during narration.	
Voice Principle	Yes – narration is used for both video types.	
Embodiment Principle	Yes – the onscreen agents are human: students and lecturers, respectively.	
Image Principle	Yes – the narrator is off-screen for both video types.	

Figure 16 Application of Multimedia Principles to the instructional videos

All of the videos where converted into windows movie video format (.wmv) to aid viewing as the University used PC's in each of its computer laboratories and library. In all years the instructional videos were presented in an embedded windows media player, the students could not download the videos and the video controls were disabled meaning that they had to watch the video from start to finish each time they did so. No limits were applied to the number of times learners could view each of the videos as well. The lead researcher attempted to provide similar resolutions and streaming rates for the participants for consistency across both types of videos to ensure that one type of video was not preferred to the other due to presentation. Unfortunately there were differences in presentation style due to the years between production and the differing styles of the editors and the instructional principles provided by Mayer, (2008). The videos were disseminated online with an adaptive release ensuring groups had access to each respective group when the crossover design dictated.

4.6 Internal Validity

During each year of the study the students were requested not to share resources with friends outside their seminar group to minimise the experimental treatment diffusion and threaten the internal validity. Students did not receive scores during the period to prevent compensatory rivalry/equalisation, resentfulness or demoralisation between groups, particularly from those who initially received an intervention and experience the control condition subsequently.

4.7 Instruments

The students were formatively assessed by each lecturer on five criteria (Motivation, Safety, Exercise, Timing and Progression) that each used a 10cm analogue scale (from 0 to 10 with no marked increments). The criterion was devised from the rehabilitation criteria described by (Smith, 1998); these encompass cognitive, affective and psychomotor domains. The students received formative feedback verbally at the end of each session; the completion of the scales did not increase the workload of the lecturers involved as it provided a mechanism for the lecturers to provide simple feedback for each area. This assessment form was also used for student assessment in each year of data collection for consistency throughout the data collection phases. The five criteria where directly derived from the end of module summative assessment and mapped to one, or a combination of the cognitive, affective or psychomotor learning domains from descriptors provided by Rovai, Wighting, Baker, and Grooms, (2009) during the development of their self-reporting instrument. The aim was to maintain the structure of the assessment form the lecturers were familiar with to enhance the validity of the assessment and the reliability to which it is completed. The assessment form was deemed to be accurate as the formative assessment form positively correlated with summative results (r=0.53, p<0.01) (Cooper and Higgins, 2014). Totals were summed to provide an aggregate score and a score for each criteria, averages were then calculated accordingly.

Although the scores were gained within groups, they were assigned to each individual (within the group) to provide an individual score for each participant throughout the duration of the study. The students worked in different groups each week ensuring no individual score would be the same. In essence, each participant had the potential for nine separate scores within each condition (experimental or control), yielding an overall total for each phase. The instrument contained the following criteria, which are defined as:

4.7.1 Motivation

This scale is used to identify components of the affective domain and was assessed according to the level of enthusiasm, motivation, professional demeanour, efficacy and clarity of the instructions given by the student to the group throughout the rehabilitation session. Although they had to perform the sessions with mock patients, students were encouraged to engage with them as if they were injured patients, which is an essential skill that is analogous with the affective domain.

4.7.2 Safety

This scale identifies how safe the programme and how controlled the environment was that the rehabilitation participants were conducting their rehabilitation exercise in. The criterion is a core component of any sports therapist's practice and straddles the cognitive and affective domain. Whilst the students will not receive specific and accurate feedback from their mock patients, they should be hesitant with their exercises and air on the side of caution at all times during their sessions so that no one is endangered or when the exercises are applied in a real world context no injury is exacerbated.

4.7.3 Exercise

The purpose of this criterion is to determine the appropriateness of the exercises the students have chosen and will straddle the cognitive and affective domain. This criterion, whilst similar to the safety criterion, encompasses more specific aspects regarding the quality of the exercise i.e. how it is demonstrated, corrected if need be and the duration it is performed for. It is imperative that the students are able to spot faults as inaccurate positioning of joints above or below the target joint could result in an incorrect exercise or a negative modification to the intended exercise.

4.7.4 Timing

This scale was used to quantify how the student was able to manage their time and implement the planned rehabilitation programme; this scale straddles the cognitive and psychomotor domain. The criterion has real world implications in terms of efficacy within the set time frame; it is particularly important for summative examination and as such is a key criterion to be assessed.

4.7.5 Progression

This scale is used to assess the ability of the students to progress the exercises that they have chosen as their mock patients perform them; as a result this scale straddles the cognitive and psychomotor domains. The latter is particularly pertinent as the patients will be ascending the psychomotor domain themselves and the students will base their progressions on how quickly the patients are able to master the exercise appropriately and safely.

4.8 Formative Assessment Reliability

Throughout the three years the lecturers that delivered the theoretical content and that formatively assessed the students were blinded to the study for each and every year, the same member of staff assessed the same students formatively each week for consistency within each year of data collection. The lecturers were briefed prior to the commencement of the data collection period and a recap was provided at the beginning of each successive year. The lecturers responsible for the academic delivery and formative assessment of the module remained consistent throughout the entire research period and there were no changes to the content or schedule of the module except for calendar variations that were beyond the control of the researcher.

4.9 Pilot Study Years 1 and 2

4.9.1 Participants

The participants for the first two years of data collection consisted of ninety-eight students, sixty females and thirty-eight males with a mean age of 20.87 years (SD = 4.53).

4.9.2 Research Design Considerations

A quasi-experimental design (repeated measurement within condition and phase) was adopted utilising existing seminar groups as control and experimental conditions for the duration of the study. A *3 x 3* crossover design was implemented for the first two years of data collection; the design was based upon a Latin square (Table 2) that ensured all students participated within each of the three conditions throughout the module to ensure no group of participants was disadvantaged.

Phase	Group A	Group B	Group C
1	Control	Long	Short
2	Long	Short	Control
3	Short	Control	Long

Table 2 3 x 3 Crossover Design

4.9.3 Experimental and Control Condition Allocation

During the phases of data collection there was a minimum of three seminar groups within the Rehabilitation module (A, B and C). In the second year of data collection there were four seminar groups, two of these were amalgamated to ensure there were three distinct groups.

Two of the specified groups will have access to the different instructional videos (classified as Long and Short) whilst those undertaking the control condition within the study will not be provided with any other additional materials than those that are already available to all of the students. The condition allocation was entirely random and assigned by the labeling of the practical sessions by central timetabling. Each of the conditions lasted for six weeks and encompassed two peripheral joints.

4.9.4 Instructional Videos Delivery

In the first year the instructional videos were housed on a virtual learning environment (VLE) alongside the other materials that were provided for the students. The VLE was used to facilitate easy access to the resources as they were contained within an additional sub-folder on the site and it provided a method of tracking the number of views each video had received.

After the first year it became apparent that the statistics from the VLE (tracking and accesses) were inaccurate, an issue that had been confirmed in discussion with the VLE support team. As a result, in the second year of data collection a bespoke internal website was created (with hyperlinks placed on the VLE to add convenience) that was designed to output viewing information into a separate Microsoft Access document.

4.9.5 Data Analysis

Each Likert item was measured with a standard 30 cm ruler to an accuracy within one millimetre, all scales where measured by the same researcher to enhance reliability. Totals where summed and averages calculated for each of the fifty-four data collection points within each year (6 joints x 3 stages x 3 conditions), one-hundred and eight in total across the two years of the experimental period. In addition to the null hypothesis statistical tests, Cohen's effect size statistics were calculated using a macro-enabled spreadsheet (Hopkins,jh 2007) and evaluated using the descriptors and magnitude-based inferences provided by (Batterham & Hopkins, 2006).

Although the scores were gained within groups, they were assigned to each individual (within the group) to provide an individual total for each participant for each experimental condition. The students worked in different groups each week, which would ensure that no individual total would be the same for any of the students. In essence each participant had the potential for six separate scores (six practical sessions within each experimental phase) within each experimental phase, yielding an overall total for each phase that could be compared between conditions. As a result individual effect sizes were calculated using the same procedures and formula's that were used for the groups with the addition of Hedges and Olin's correction for *d* statistic bias that is recommended when there are less than 10 data points in each group (Hedges & Olkin, 1985).

In addition to the procedure detailed above participants were excluded from the study during the data analysis phase if they attended less than 50% of the sessions during each of the experimental phases of the study; this resulted in twenty-one students being excluded from the data analysis over the two-year period.

4.10 Revised Methodology (Year 3)

4.10.1 Participants

The participants for the final year of data collection were 68 level 5 students (mean age 20.6 ± 2.8 years, 53% female). Students were enrolled onto the Rehabilitation and Remedial Therapy module at a University in the North East of England (2012-13).

4.10.2 Methodological Design

As per the first two years the same quasi-experimental design was applied. The experimental condition had access to instructional videos for the first half of the experimental period (9 weeks) whilst the control condition only had access to traditional teaching resources. The groups then swapped conditions for a further 9 weeks as it would be unethical for them to not have access to a resource that could potentially improve their knowledge, skills and summative assessment (2 x 2 repeated measures crossover design). The design is similar to that used in the first two experimental years with one fundamental difference; the previous 3x3 crossover design has been replaced with a 2x2 crossover design (table 3). The purpose of the modification was to increase the duration of the intervention and thereby increase the sample size for each aspect of the criteria the students were assessed against.

Phase	Group A	Group B
1	Control	Video
2	Video	Control

Table 3 2 x 2 Crossover Design

4.10.3 Condition Allocation

The Rehabilitation and Remedial Therapy module in 2012-13 contained four seminar groups. As two distinct groups were required for the study two seminar groups were combined to form the control condition and two were combined to form the experimental condition for each experimental phase. This was deemed to be the most feasible option for condition allocation, it also ensured that the entirety of each seminar group was undertaking one condition or the other minimising the potential for treatment diffusion. The allocation of conditions as per the first two years again was random and decided upon by central timetabling with their allocation of the practical sessions i.e. group 1 and 2 formed the control condition.

4.10.4 Experimental and Control Conditions

During the final year of data collection there were four seminar groups within the Rehabilitation module (A, B, C and D) and two experimental groups were formed as described above (Control and Experimental). The Experimental group will have access to the short instructional videos whilst the control condition within the study will not be provided with any other additional materials than those that are already available to all students.

4.10.5 Instructional Videos Delivery

For the final year of data collection, the videos were ported back to the VLE alongside the standard materials as although the bespoke website provided access data, accessing the website was an additional barrier for the students that they did not wish to have. Although access statistics would be unreliable it was deemed of greater importance that the students were able to view the material with ease.

4.10.6 Data Analysis

The principal outcome was the difference between the mean aggregate scores and the means of the specific criteria obtained throughout each condition by the student groups. The secondary outcome was the difference between the individual students mean aggregate score and each of the criteria assessed in the study across the two conditions.

Due to the design of the study having to conform to the academic calendar there was no distinct opportunity for a set washout period. As a result the potential for a carryover effect was analysed using the procedure and calculations contained within the article by Welleck and Blettner (2012) see appendix 3.

Chapter 5: Results

The current study encompasses two similar but sufficiently different periods of data collection; as a result the findings of each shall be presented individually initially and then combined. The first set of results pertains to the data obtained from the first two years of data collection, which included three different conditions (Long video, Short video and Control). The final year of data collection is initially presented in isolation, as it compared short videos versus a control group. In the final part of the results section, the control and short data are combined to provide an overall evaluation of the data. The data has no distinct units therefore they are presented as integers.

5.1 Results of Group Analysis (Years 1 and 2)

In order to investigate the variance between the three experimental groups a between-between Univariate Analysis of Variance (ANOVA) with Tukey post-hoc tests, was used to analyse the formative scores of the three groups. The results demonstrate that there was no statistical difference between the three groups, [$F_{2,104} = 0.17$, $p = .85$] despite the observable difference in means between the control and differing video groups in Figure 17.

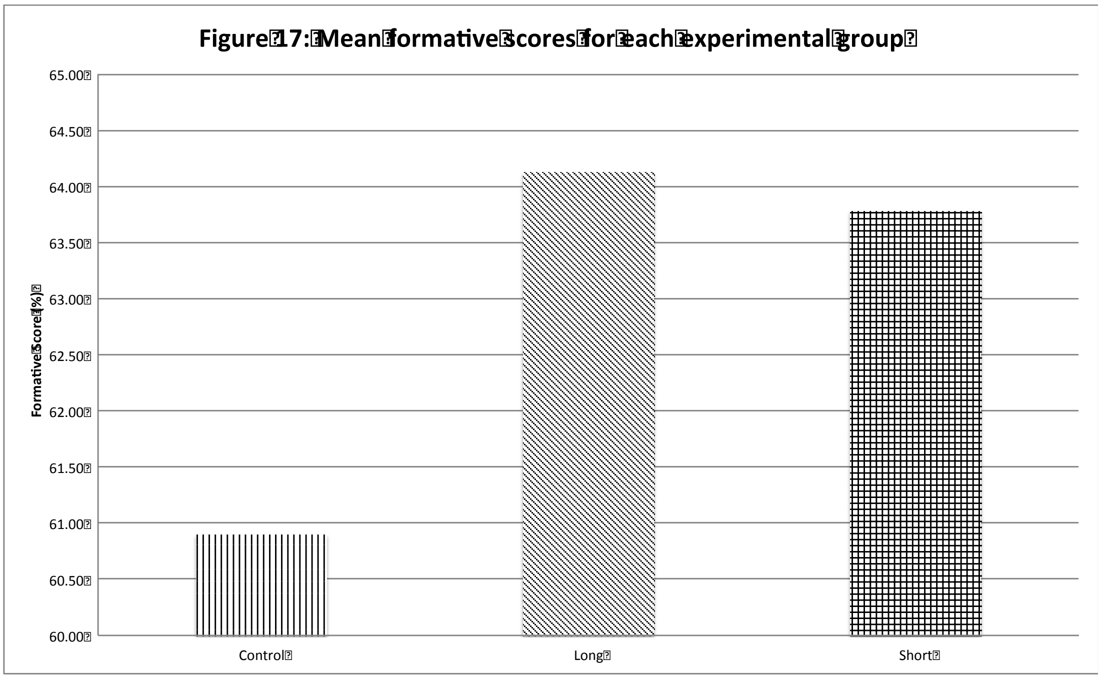

Figure 17 Mean formative scores for each experimental group

The mean difference in formative assessment between Control and Long was 3.23 (95% confidence interval -1.41 to 7.86, *p* =.25) and Control and Short was 2.88 (95% confidence interval -1.72 to 7.48, *p* =.30). The magnitude of this effect for each comparison was small (effect size 0.27: -0.20 to 0.73) and (effect size 0.21: -0.19 to 0.60), respectively. The probability (% chances) that the true population effect is beneficial/ trivial/ harmful is 61.8/ 35.8/ 2.4 and 51.1/ 46.6/ 2.2: therefore, the interventions are possibly beneficial and very unlikely to be harmful. The latter finding is of substantial relevance, as it indicates that the videos can generally yield either trivial or positive effects. When the two video groups were compared the mean difference in formative assessment between Long and Short was 0.65 (95% confidence interval -3.29 to 2.59, *p* =.84). The magnitude of this effect was insubstantial (effect size 0.05: -0.42 to 0.52). The probability that the true population effect is beneficial/ trivial/ harmful is 25.6/59.8/14.7: therefore, the difference between the two video groups is predominantly trivial.

5.2 Results of Individual Analysis (Years 1 and 2)

A Two-way *factorial* ANOVA was performed to analyse the differences, if any between the individual students shown in Figure 18. The findings demonstrate that there is a statistical difference between the participants ($F_{97,1051}$ =3.66, p <.01). It confirms the previous ANOVA by showing that there is no significant difference between the groups ($F_{2,1051}$ =1.53, p =.22) despite the observable difference in means in Figure 18. There is however a significant interaction between participants within the groups ($F_{194,1051}$ =2.11, p <.01) demonstrating the differing effects of each condition. This is further demonstrated when a Univariate ANOVA is performed to assess the variance between the participants with separate groups ($F_{97,1247}$ =3.44, p <.01, partial η^2 =0.21).

As with the experimental conditions, individual effect sizes where calculated using the same procedures as above. They were performed with the addition of Hedges and Olkin's correction for *d* statistic bias, that is recommended when there are less than 10 data points in each group (Hedges & Olkin, 1985).

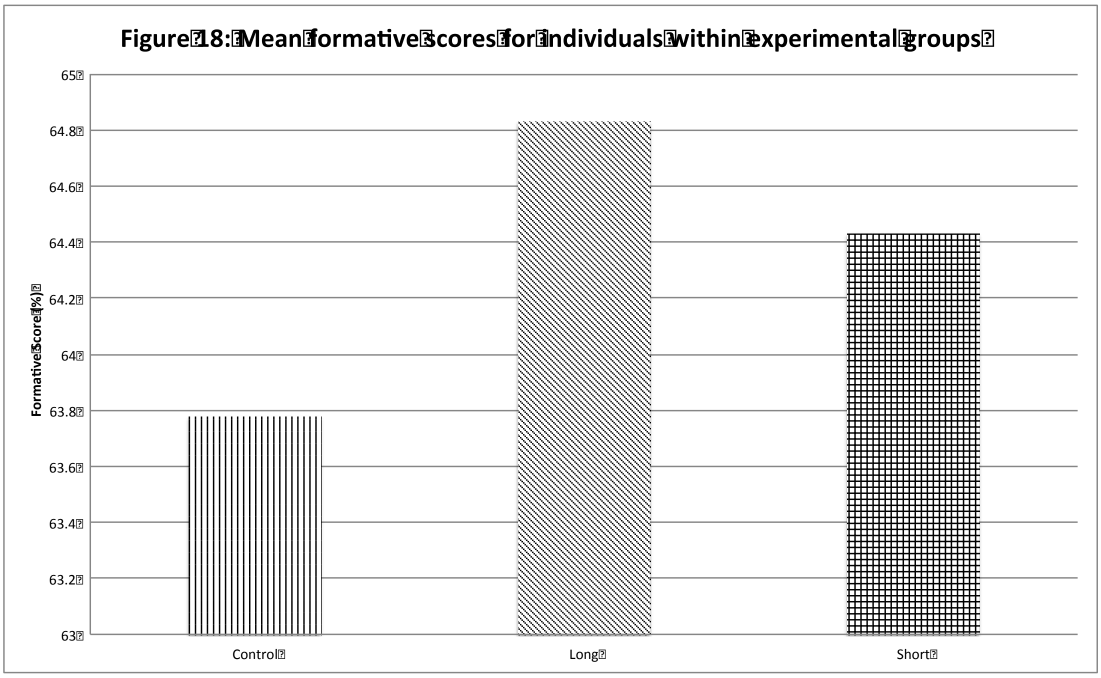

Figure 18 Mean formative scores for individuals within experimental groups.

The mean difference in formative assessment between Control and Long was 0.34 (95% confidence interval -21.72 to 22.39, *p* =.40) and Control and Short was 0.65 (90% confidence interval -21.02 to 22.32, *p* =.33). The magnitude of this effect for each comparison was small (effect size 0.49: -0.39 to 1.37) and (effect size 0.57: -0.28 to 1.42), respectively. Even though each of these were small as with the group data the quantity is greater indicating a potentially greater effect for individuals. The probability (% chances) that the true population effect is beneficial/ trivial/ harmful is 45.9/50.6/3.5 and 52.1/45.1/2.7: therefore, the interventions are possibly beneficial and most notably very unlikely to be harmful, further supporting the findings of the group data.

5.3 Accesses (Years 1 and 2)

Throughout the duration of the testing periods the accesses to the control site and amount of times each of the videos was played, within each of the video sites was monitored and recorded. The control website was accessed 365 times in total, the five long videos where viewed 453 times in total (1.24 times more than the control) and the twenty-five short videos where viewed 958 times (2.62 times more than the control and 2.11 times more than the long videos). As the students could view these videos in groups it is impossible to correlate individual accesses with formative scores.

5.4 Academic Performance (Years 1 and 2)

In conjunction with their individual formative marks for each of the three conditions and their summative mark from the rehabilitation module, summative data from the end of the student's first year practical examinations was available. This enabled a Spearman's Rho to be performed, to identify any trends between the data sets. The results of the analysis show that there is a strong positive correlation between the average mark from the student's first year practical examinations and their final rehabilitation mark (0.53, $p < .01$). There is also a strong positive correlation between their formative marks obtained during the Short and Long video conditions and their final rehabilitation mark (0.28, $p < .05$ and 0.29, $p < .01$, respectively).

5.5 Results of group-level analysis (Year 3)

The mean difference in aggregate formative scores for student groups between control and experimental conditions was 2.33 (95% confidence interval 0.57 to 4.08). The magnitude of this effect was moderate (effect size 0.68; 0.04 to 1.31) and the probability that the true population effect is beneficial/trivial/harmful is 93.9/5.4/0.6. The difference between student groups is illustrated in Figure 19. During weeks 1 to 9 there is a significant observable difference in mean aggregate formative scores between the control and video condition. The trough at week 4 is the only time that the score of the video group B, is lower than the score of the control group A. Additionally the scores of the video group B are generally more consistent than those of the control group A, the latter of which vary considerably almost each week. At week 10 when the student groups crossed over, an interesting pattern begins to emerge. There is no distinct difference as observed in weeks 1-9 for weeks 10-18, between the control group B and the video group A. The control group B maintain their performance throughout this period as the video group A increase their total formative scores typically week on week. This results in the video group A outscoring the control group B in the final week of the data collection period.

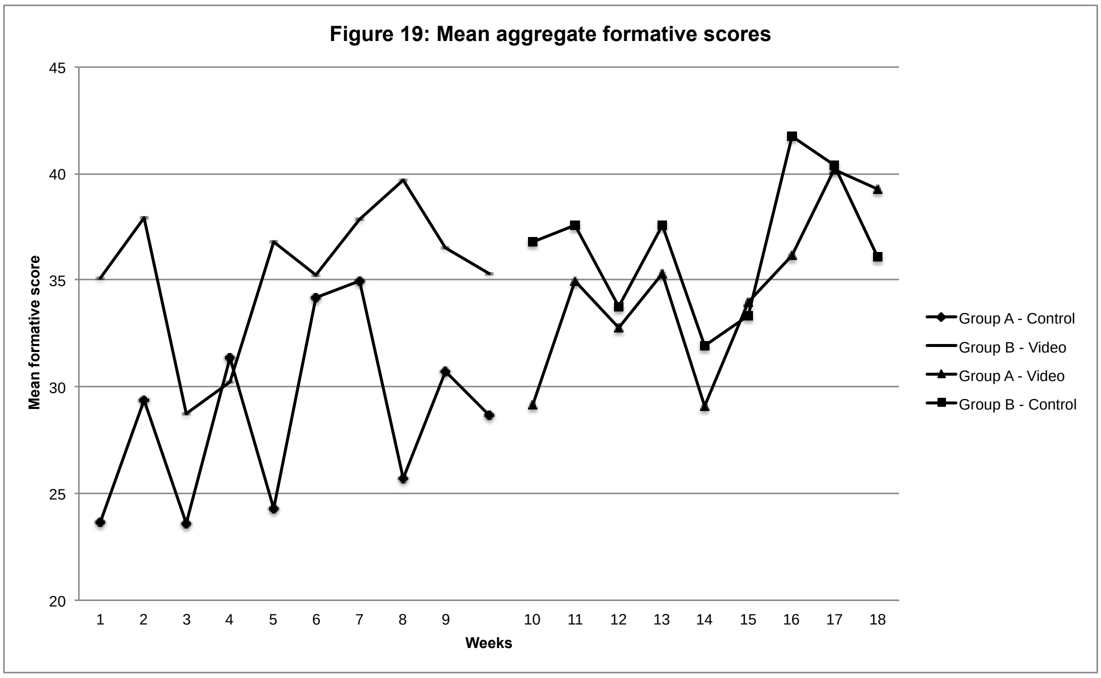

Figure 19 Mean aggregate formative scores

Table 4 details the mean difference (with 95% Confidence Intervals) and effect size (with magnitude, 95% Confidence Intervals and the beneficial/trivial/harmful probabilities) for the five individual criterion scores, for the groups between experimental conditions, allowing for a greater level of analysis to be achieved. The most notable of the five criterions are Timing and Progressions as they have a moderate effect size over the other criterions. Despite this all of the criterions have a probability of more than 80%, with the most harmful probability only being 3.2%.

Formative Component	Mean Difference	95% Confidence Interval		Effect Size	Magnitude	95% Confidence Interval		Beneficial	Trivial	Harmful
Motivation	0.27	0.02	0.52	0.50	Small	-0.07	1.08	87.0	11.8	1.1
Safety	0.41	-0.04	0.87	0.46	Small	-0.16	1.09	81.9	16.1	2.0
Exercise	0.42	-0.08	0.92	0.48	Small	-0.22	1.17	80.6	16.6	2.8
Timing	0.71	-0.08	1.50	0.72	Moderate	-0.27	1.70	86.9	9.9	3.2
Progressions	0.50	0.05	0.96	0.62	Moderate	-0.07	1.31	90.2	8.6	1.3

Table 4 Individual criterion scores for groups

5.6 Crossover

Figure 19 demonstrates how the videos effectively increased mean assessment scores, within both experimental conditions throughout the experimental phase. Figure 19 also shows how Group B experienced a carryover effect due to a minimal/no washout period. Using the calculations from Welleck and Blettner (2012) with the data set (n=68, df=134) the pre-test (to check the assumption of negligible carryover effects) T value is -0.019 and is significant ($p < 0.000$). This demonstrates that there is a carryover effect and the test for differences yields a T value of -34.76 that is again significant $p < 0.000$. The crossover calculations show that Group B retained information for the ‘control’ condition of the data collection period.

5.7 Results of Individual-level analyses (Year 3)

The mean difference in total formative assessment for individuals between control and experimental conditions was 3.92 (95% confidence interval -4.10 to 11.93). The magnitude of this effect was small (effect size 0.40; -0.40 to 1.21) and the probabilities that the true population effect is beneficial/trivial/harmful are 40.9/56.0/3.2.

Table 5 details the mean difference (with 95% Confidence Intervals) and effect size (with magnitude, 95% Confidence Intervals and the beneficial/trivial/harmful probabilities) for the five individual criterion scores for individuals, between experimental conditions. The mean differences range from 0.38-0.78 yet despite this the magnitude of each effect size was small with each ranging from 0.32-0.39. As a result, the probability of each criterion being Beneficial/Trivial/Harmful is comparable throughout.

Formative Component	Mean Difference	95% Confidence Interval		Effect Size	Magnitude	95% Confidence Interval		Beneficial	Trivial	Harmful
Motivation	0.38	-1.14	1.90	0.34	Small	-0.46	1.15	35.8	60.9	3.3
Safety	0.42	-1.01	1.86	0.35	Small	-0.41	1.10	35.9	57.7	3.3
Exercise	0.78	-1.08	2.68	0.35	Small	-0.45	1.16	37.3	59.1	3.5
Timing	0.73	-1.70	3.16	0.32	Small	-0.47	1.10	34.3	60.8	4.9
Progressions	0.74	-1.24	2.73	0.39	Small	-0.41	1.20	39.6	57.6	2.8

Table 5 Individual criterion scores for individuals

5.8 Meta-analyses of Overall Results

In order to enable an effective analysis of the overall results, the data (Control versus Short) from the first two years of data collection was processed alongside the data from the final year of data collection. Whilst it is acknowledged that the method of combing the data omits a proportion of the data from the first two years, it enables an appropriate comparison over the time period. The larger sample size yields an overall meta-analysis effect size of 0.47 (95% CI 0.33 to 0.61; $p < 0.05$), which is medium in size (Cohen, 1988) and greater than any *d* statistic identified previously. The application of the confidence intervals infers that it could range from small, to moderate. The lower value however is still comfortably greater than the 0.2 boundary for a small effect size identified by (Cohen, 1988). The data is illustrated in Figure 20 and exemplifies the differences between the two studies. It is clearly observable that the mean score from Cooper et al., (2015) is greater than the upper value of the 95% confidence interval from Cooper & Higgins, (2014). The meta-analysis statistics average the two means (factoring in sample size) and the values are within an expected range. The confidence intervals identified following the refinement of the data, are slightly smaller than anticipated considering the ranges identified by both of the periods of data collection.

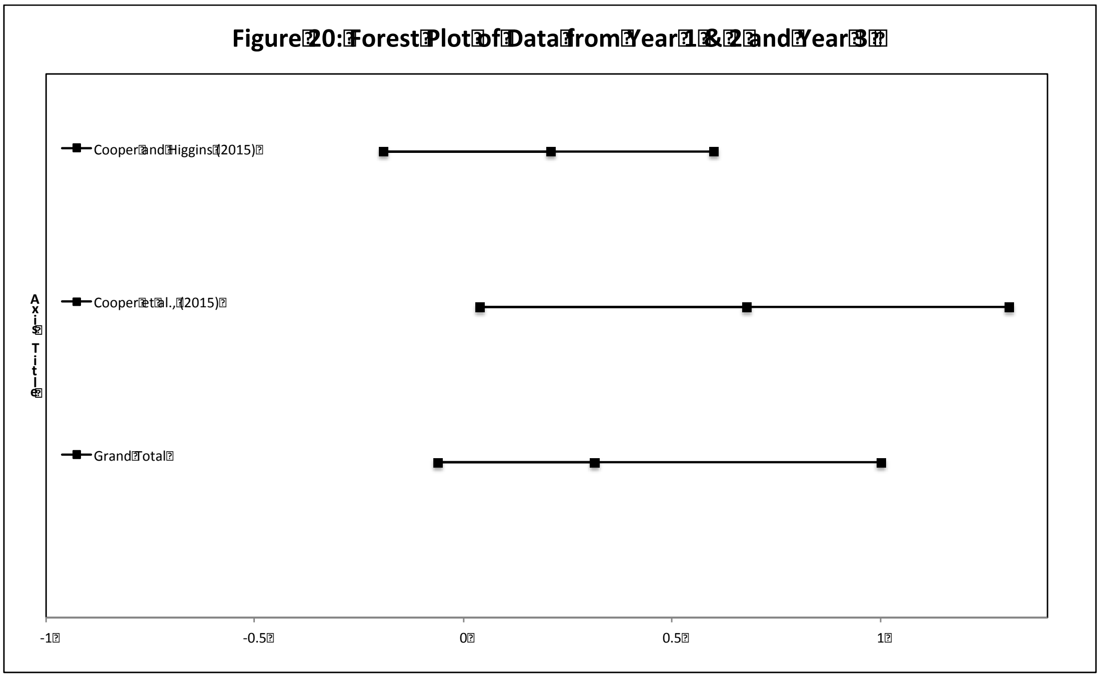

Figure 20 Forest Plot of Data from Year 1 & 2 and Year 3

Chapter 6: Discussion of findings from years 1 and 2

6.1 Introduction

The following discussion is modified from the publication (Cooper & Higgins, 2014: see Appendix 4) of the first two years of data collection. At that time a proportion of the evidence that has been cited within the literature review was not available. Therefore, the application of the first two years results to this literature will be undertaken in the overall discussion. Whilst this may appear obscure, the conclusions of the first two years' worth of data collection were reflected upon yielding the model for the final year of data collection. As such it is chronologically and developmentally important for the study in its current form.

6.2 Discussion

The data presented in sections 5.1 to 5.4 identifies the differences between experimental conditions. The experimental conditions did not show any statistically significant results for the group or the individual analysis. However, the use of effect size statistics, which enables the researcher to 'move towards a more generally interpretable, quantitative description of the size of an effect', demonstrates a small effect size when both the experimental groups were compared to the control group (Fritz et al., 2012). The real life relevance of this was an increased mean score over the control group. The implications of this 'real life relevance' should not to be understated; it is a predominant strength of the effect size approach that is only beginning to be fully appreciated. Whilst this was only a modest increase in the actual mean score, the additional couple of marks for a group or an individual could mean the difference between pass and fail; or the attainment of a higher classification boundary. The magnitude-based inferences indicate the most important findings from this study, showing that both of the video groups are almost certainly not harmful to the participants. Whilst the probabilities for a beneficial or trivial effect are both within the central quartiles (with the beneficial percentage being marginally greater than the trivial) it is clear that videos are to some degree beneficial when students are working in groups.

A comparison of the two experimental groups demonstrates a substantially reduced mean difference and an insubstantial effect size. This suggests that the different delivery durations, any infringement of the redundancy principle (see Figure 16) and the production value, made little difference to the scores achieved by the groups.

When the findings for individuals are interpreted the two-way factorial ANOVA identifies that the scores achieved by individuals are significantly different from each other. This demonstrates that individual participant data can be extrapolated from the experimental group design. Furthermore, there is a significant interaction between the groups, as just over one-fifth of the variance is explained by the combined effect of the video groups.

The effect sizes for all measures are classified as small; these are approximately twice as large for individuals than groups (with a corresponding increase in confidence intervals). This partly reflects the small scale of the study. The magnitude-based inferences for individuals follow a similar pattern to the groups with the beneficial and trivial probabilities being almost equal. Furthermore the probability that the intervention will have a harmful effect is exceedingly small, a vitally important finding of the study. As a result the confidence intervals illustrate the extent of the effect on the scores of the individuals, with dramatic effect. Within the cohorts there are clear differences in the effect of videos in a beneficial or harmful way. However, the majority of the time this is beneficial with the effect being the equivalent of an entire classification boundary for some students (10% boundaries). In addition, the findings of the Spearman's rho, indicate that the formative marks the students obtained individually, (when in the short or the long video group) is significantly correlated with the marks they obtained on an individual basis in the final summative rehabilitation examination. The formative marks obtained when they experienced the control group were not however. It is not possible to claim that the videos had a direct influence, due to the assumption that different levels of preparation were undertaken by the students in the run up to their formative, versus summative assessments. It does however demonstrate, that they were working at a level that is similar to their academic performance in the examination, making the feedback and advice more relevant and beneficial.

As a result of these findings it is possible to agree with the generalised conclusions from Chen, (2011) McNulty et al. (2009a) and Zhang et al. (2006) that e-learning can improve traditional teaching methods, with a caveat that it may not be effective for the entire cohort. The magnitude of the improvement can vary considerably across groups and individuals. Whilst the interventions did not necessarily promote learning efficiency or effectiveness consistently, in agreement with Chen et al. (2011) and Kala et al. (2010) their adoption or integration will probably be of benefit for a large proportion of the cohort. The most noteworthy finding from the first two years of data collection is that the probability of the use of videos being harmful to a group or an individual is exceedingly small. Although this does not provide conclusive evidence, of the effectiveness for teaching clinical skills that Bloomfield et al. (2010) are seeking, it should encourage educators to consider adopting the use of videos to teach clinical skills. This is particularly relevant, when you consider how important the role of observational learning is still within education to this day (Gog et al., 2008). Videos ensure that the viewing of flawless and error free movements performed by experts can be repeated and analysed an unlimited number of times, thus embedding and reinforcing the movement with no cognitive variation. The ability to pause and work through a complex movement, one frame at a time provides the learner with an incredibly powerful tool. This is a mechanism that can potentially transform the passive design of the video and promote constructivist learning. This further supports the work of Mayer, (2003) regarding the usage of instructional video within a cognitive setting. The findings expand on the initial applications of instructional media by applying it to the affective and psychomotor domains. The latter aspect is inextricably linked with the work of Bloom, (1956) and others domains demonstrating the potential that the initial findings have within a higher education context.

6.3 Limitations identified from year 1 and 2

The assessment of psychomotor skills is not as objective as it is for other domains. Whilst the reliability of assessment within this study was acceptable, the subjectivity of this element is an area that would benefit from further research. This is for the students and the examiners so that it is possible to ensure that standards and marks achieved/given are as reliable as they can be. The scale of the first two years of data collection was small; with three conditions (short video, long video and control) meaning it was not sufficiently powered to detect small effects. The aim of the final year of data collection would be to increase the scale by reducing the conditions to video versus control only. The purpose of doing so is to increase the exposure of the participants, to each condition in an attempt to quantify the effects, if any, further.

6.4 Conclusion from years 1 and 2

The role of instructional videos within traditional learning should be further encouraged, although the first two years' worth of data collection could not establish a conclusive evidence-base for their inclusion. The effects observed are potentially substantial for some students, given the real world application of the small effect size on grade boundaries. The first two years have established that instructional videos are very unlikely to be harmful to a student's learning and may have other benefits that the study has been unable to quantify. These could range from flexibility and independence of learning (any time or place with an internet connection), the ability to be repeated instantly and the consistency of delivery that could aid the acquisition of cognitive, affective and psychomotor skills further.

6.5 Reflection on years 1 and 2

As a result of the identified limitation regarding power, it was proposed that only one type of video (Short or Long) be compared against a control group for another year's worth of data collection. This would then result in two experimental durations of nine weeks (as opposed to three experimental durations of six weeks), which would hopefully increase the power of the study. This approach also reduces the complexity of the study, which may also benefit the participants, as they will be required to access materials for a more specific, less frequently alternating duration. There will not be any confusion (if there was any) between videos and content, which was potentially possible given the time difference between their creations. As they are now conducting their sessions over an increased duration, there is the possibility that they will be able to attain higher formative scores. This could be due to the extended time they will be able to use the resources for, potentially consolidating and enhancing learning to higher levels.

6.6 Discussion Introduction (Year 3)

As per year 1 and 2, the discussion below is modified from an article that has been submitted for publication (Cooper, Higgins, & Beckmann, 2015: see Appendix 5). The application of the findings from year 3 can be evaluated in isolation, which they will be initially. The findings will then be integrated and discussed to contextualise the contribution of the study and its applicability to the pertinent literature identified in chapters 2 and 3.

6.7 Discussion of group findings

The results from the third year of data collection clearly demonstrate the benefit of online instructional videos when teaching rehabilitation skills. The aggregate formative scores between the student groups, during the first 9 weeks of the study exemplify the difference that the instructional videos made. This is observed quantitatively with an effect size of 0.68 (substantially larger than any identified in year 1 and 2) and a probable beneficial effect of 93.9%, with only a possibility to harm of 0.6%. The latter is an exceptional finding that is substantially smaller than any previously identified, demonstrating the potential for harm is almost 1 in 200. In every week during the first nine weeks (except one) the initial video group B, scored far higher on the formative assessment than the control group A. When the groups changed conditions, a different trend emerged due to a potential crossover effect due to no distinct washout period. Normally a crossover contamination, were the results are not analysed independently by sequence group are of limited, if any, scientific value (Wellek & Blettner, 2012). In this case however, the crossover effect theoretically illustrates the efficacy of the instructional video. This is hypothesised as the information was retained by those in the first experimental group, and then used by them, when they were in the control condition. It is impossible to establish if there was any treatment diffusion between groups, however during the first phase of the study, the differences between mean aggregate scores are generally equal. Treatment diffusion may have been present throughout the second phase, as the students with prior knowledge of the instructional videos, would be able to communicate effectively with those who now had access to them. The use of precise information and/or specific details when communicating about the resources with the other students, could lead to the attainment of reinforcement of prior knowledge.

During the second phase of the study Group B, now the control group, continued to outscore the video group A, in each week except week 15 (marginally) and in the final week. During this period it is also apparent from Figure 19 that the video group A steadily began to increase their aggregate formative score. This resulted in a more predictable pattern between the two groups, as the relative ease or difficulty of each joint and stage began to follow a similar path.

The data is able to demonstrate how Group A (video), when using the instructional videos consistently outperformed their counterpart Group B. This then enabled them to continually perform at a comparable level even when the resource had been withdrawn, an aspect that has been directly cited as a component of constructivism (Kim & Reeves, 2007; Vogel-Walcutt et al., 2011). This is in agreement with Mayer (2002) who stated that constructivism can be fostered when cognitive activity is present within the learner, despite the media being predominantly passive. This also has the potential to demonstrate that higher-order learning can be facilitated with the instructional videos, as exemplified by both groups, an aspect that was previously unsupported (Deubel, 2003). The implications of this finding are extensive as it demonstrates that the video can positively support learning, when learners have access to the resource. It also demonstrates the positive impact of the resources for the learners, as they have retained the information for a period beyond the experimental condition.

When the individual criteria that provided the accumulative score were analysed separately, the largest effects are measured in Timing and Progressions. Each of these criteria straddles both the cognitive and psychomotor domains. This finding however is not unexpected as for the students to foster a constructivist approach, cognitive activity needs to be present. As the students perform an increasing number of rehabilitation sessions, they gain external and internal feedback relating to the Timing of exercise and knowledge, about how patients can be progressed. Overall each of the criterions demonstrates a positive effect when the students were working in groups, with the lowest effect size being 0.46 and the lowest beneficial probability being 80.6%.

Additionally, it is also worth noting from a methodological standpoint that the study demonstrates that the 2x2 crossover design is also an ethically sound model. This is concluded, as Group B (Control then Video conditions) was able to match and exceed the scores of Group A by the end of the study. The implication of this finding is sufficient, to support its further use within an educational context.

6.8 Discussion of individual findings

In addition to the student group data, each individual was attributed with a cumulative score from the different groups they participated in throughout the study. Although the scores were obtained within a group setting, it is possible with some assumptions and limitations, to track an individual using this method. The mean difference for individuals was 1.59 greater than the mean difference for student groups however, this was not all positive as the confidence intervals demonstrate clearly. Greater variance between individuals is not unexpected; the magnitude-based inference is most revealing regarding this, especially when compared to the student groups. The student groups magnitude based inference was highly beneficial, however for individuals it is predominantly trivial (56.0%) rather than beneficial (40.9%). It is however worth noting that the potential for harm is only slightly greater than that from the student groups at 3.2%. When these findings are compared to those of Laughlin et al., (2006) it is clear that some of the individuals will have stood out from the student groups. They will have in theory outperformed them as demonstrated by the mean difference, but this is the exception rather than a consistent finding. The student groupings within this study were of three or more therefore the findings identified within pairs (Lou et al., 2001) are not directly applicable. However, it is mentioned here as this demonstrates that smaller groups and individuals, are capable of outperforming groups of three or more.

The individual-level analysis of the criteria demonstrates that although the range of mean differences and the distribution is similar to that of the groups, the effect sizes are more comparable, as they only range from 0.32-0.39 and are all small in nature. Therefore, no particular criterion or domain emerges as being most beneficial to an individual over a group.

An acknowledged limitation of the study is the lack of a washout period before final assessment. However those undertaking research within an academic timetable may not have such an opportunity due to calendar constraints. It is therefore recommended that a 2x2 crossover study be used, perhaps with the winter holiday as a washout period, if possible. Another potential limitation of the study relates to the participants and their allocation. Whilst it was an effective mechanism for condition allocation it was not a random allocation and did not fully ensure that the groups were balanced academically. That said however, the studies design and the minimisation of treatment diffusion, were of paramount importance. The approach was justified as during the final phase of the study, the groups were able to produce comparable standards.

The final year of data collection has identified a far greater effect size for the use of instructional videos for student groups and individuals alike, than previously reported in from year 1 and 2 (Cooper and Higgins, 2014). In addition, the current study has revealed the specific criteria where instructional videos improved group and individual performances.

6.9 Conclusions from year 3

In summary the instructional videos have consistently proved to be beneficial at best or trivial at worst, for students when learning and performing rehabilitation skills, in both a group and individual context. Whilst there was a greater variance within the individual analysis, the probability that the instructional videos would have a harmful effect was never greater than 5%. In addition, the instructional design and dissemination of the videos in theory has led to a student centred learning approach. This has potentially fostered constructivism-allowing students to retain acquired knowledge, leading to a consistent performance without the use of the resource.

Chapter 7: Overall Discussion

7.1 Introduction and reflection on structure

As the thesis has been subdivided by two distinct stages of data collection, the structure has varied from that of a conventional thesis; however the process has been highly informative for the thesis as a whole. The submission of a publication from the first two years of data collection resulted in a full analysis of the data collected up until that point. This provided an opportunity to reflect in detail on what had been learned up to that point. If this had not happened then the final year of the study may well have simply been a continuation of the data collection from the previous two years. The final year of data collection has proved to be the most informative part of the study, identifying aspects that were not discovered in the initial two years. As a result, a particular strength has been the opportunity to reflect, this is a process that all researchers should undertake during data collection to enable refinements were ever possible.

7.2 Discussion of overall results

The findings of the overall meta-analysis results indicated an effect size of 0.47. This is greater than the typical value provided by Hattie (2008), is of medium size (Cohen, 1988) and therefore it can be inferred that the use of short videos, can be seen to have an educationally important impact when used to support the learning of practical rehabilitation skills. However, the findings on which the meta-analysis results are based are in distinct contrast to other studies. Nelson et al., (2014) state that there is a growing quantity of evidence that studies, that have a larger number of participants have lower effect sizes. The authors believe this, to be due to the theoretical aspect that the increased sensitivity derived from the larger sample size is offset by the effect size simultaneously decreasing. This, as can be observed from Figure 18 is not the case for this composition of data sets. In the second study the effect size is considerably larger than that of the first study, which had a smaller sample size than that of the second study. An aspect that the authors may wish to add to the statement could relate to study complexity, as the first study used three experimental groups whereas the second used two. Whilst other authors Morris & DeShon, (2002) have proposed methodologies for combining studies that have repeated measures and independent group designs, there appears to be no evidence or methodological procedures recommended relating to the minor modification observed between these studies. As a result, the findings from the studies follow conventional thinking, regarding increasing the sample size to increase the likelihood of identifying a difference between interventions, yielding the overall finding. The overall findings of both studies are consistent, with the second study seen as an extension of the first study. This did not identify any statistically significant findings, although some observable differences between experimental conditions are seen.

As a result, it is possible to answer the first research question positively, with evidence to supporting the finding that instructional videos generally have a beneficial effect on students when learning practical rehabilitation skills.

The reasoning behind the use of the term generally as a qualifier, relates to the use of magnitude-based inference analysis. Whilst this method has had some criticism Welsh & Knight, (2014) (although it was derived through miscalculations) the methodology ensures that findings are practically interpreted. If significance levels were solely used as the basis for interpretation then there would be no need to qualify the findings. Although the effect size is small, it was consistently positive across all years of data collection. While this is true, it is misleading, as magnitude-based inferences demonstrate that although the findings from each year are mainly beneficial, there is an inconsistent effect identified (ranging from 5.4% to 56.0%) across studies, encompassing groups and individuals. There is only a small possibility of harmful effect observed however, these values only range from 0.6% to 3.7% across both studies for groups and individuals. When the values are compared to the examples provided by Batterham & Hopkins, (2006) in Figure 10, they are closely aligned to those illustrated, as the possibly beneficial level which supports the positive finding of instructional videos on the whole.

The overall findings from this study, further support those from Attstro, (2005); Chen, (2011); McNulty et al., (2009) and Wouters et al., (2007) to cite a select few who stated that video is a well-established and accepted technique for the teaching of practical skills. In the initial part of the literature review, other authors were identified who were more skeptical by stating that there is limited consistent and conclusive evidence regarding their usage (Bloomfield et al., 2010; Samaras et al., 2006). The current study is also able to provide an evidence-base that spans three years, and encompasses one hundred and forty-five participants, stating that the use of instructional videos is generally positive for the majority of learners. The resulting sample size of this study is far greater than the majority of those cited within the instructional video section (see section 3.7) of the literature review (excluding systematic reviews and meta-analyses). It has a methodology that is considerably simpler than others, Ford et al., (2005) who used four experimental groups across two successive cohorts or Nousiainen et al., (2008) who had three experimental groups (which is similar to this study). The current study provided students with far greater exposure to the experimental condition, in terms of time that participants are able to access the intervention in comparison to some studies. Moore & Smith, (2012) for example only provided the students with videos in a single teaching session and Ford et al., (2005) only gave each group 25 minutes with the experimental resources. Whilst it is acknowledged they would have been aiming to identify differing aims, the educational impact of this exceedingly short intervention should be questioned. It has been previously identified that a human's working memory is limited in both capacity and duration. Therefore applying additional constraints to this may not create effective cognitive learning strategies, and enforce other extraneous factors.

In addition to experimental exposure time, the use of multiple accesses is another positive of the current study as exemplified by Ahn et al., (2011). They demonstrated that the use of reminder videos sent via text messages had a significant effect on retaining knowledge and practical skills. This is further supported from the findings of Stefanidis et al., (2006) regarding repetition. Although they were using a simulator, the participants in this study were able to view the videos and simulate the sessions they would be performing, even if it is only utilising an imagery approach.

The penultimate positive aspect of the current study in relation to others evaluating instructional videos, is that this study was conducted with a consistent methodological setting across all years of data collection. This is unlike Smith et al., (2006) who deployed their experimental testing over two years. They deployed it over different semesters and despite this methodological flaw, which was acknowledged the authors still pooled, analysed and published the findings of their study. Finally and possibly most significantly given the context of this study, some studies were only able to demonstrate a significant difference in cognitive scores Salyers, (2007) or mixed results such as those found by Ford et al., (2005).

Whilst there are many positives to the study (limitations will be addressed later) it is important to understand why they have come about. A predominant factor that underlies the success of the intervention has to be attributed to the power of observational learning, from instructional video's with the addition of constructivist higher order learning, as opposed to the lower level learning stated by (Deubel, 2003). It has been widely stated that observational learning, is one of the most powerful learning strategies known to man. It is typically linked with behaviourism as the skills are imitated or replicated, which forms the foundation of many of the psychomotor domains cited within this study (Dave, 1970; Dawson, 1998; Gog et al., 2008; Harrow, 1972; Wouters et al., 2007). Many authors have addressed the efficacy of delivering the ‘live demonstration’ from which the vast majority of research into observational learning has been conducted with. This has been through the use of instructional videos and those cited have found that they are as effective, if not more so in some settings, a conclusion that is in agreement with the findings of the current study (Moore & Smith, 2012; Ford et al., 2005; Smith et al., 2011; Smith et al., 2006). The efficacy of the observational video is mediated by several factors, namely the multimedia principles cited previously and it is these principles which are believed to have enhanced the efficacy of the instructional videos within this study compared to all others cited. This is due to the fact that, very few of them provided any guidelines as to how their videos were created. Some of the studies provided information regarding their dissemination to the participants of their studies and the select few who included information about the design of their instructional videos (Ford et al., 2005 for example) violated one/some of the evidence-based principles provided by Mayer, (2014). Whilst they found no significant differences the scores of those in the experimental group, were lower than those in the live demonstration group unlike the findings of this study, which could be attributable to this reason. The adoption and use of the multimedia principles are based on the cognitive load theory and could, in theory be directly linked with the cognitive, affective and psychomotor domains. The reasoning behind this statement is apparent when the wording of some of the principles that relate to fostering generative learning in particular are considered, for example:

The embodiment principle relates to the onscreen agent depicting humanlike gestures to aid learning, the current study used the two lecturers who taught the content and the course leader. All of the students were very familiar with the three onscreen agents, potentially engaging the students in the affective domain at a higher level than unknown actors or animated models. This is because they can assign a value to the material they are viewing from these onscreen agents, which could facilitate the production of germane cognitive load, enhancing learning. The embodiment principle is evidenced as having a small effect size of 0.36, however it is unclear if any of these materials sought to engage the affective domain in such a direct way. Therefore it is possible to hypothesise that the effect size within this context may indeed be greater.

It is entirely possible that the cognitive domain has been enhanced, as the students would be used to the visual demonstrations and terminology (pretraining principle) presented to them within the instructional videos. This could be seen as an overlap of the cognitive and affective domain, as students would accept knowledge from known sources due to a level of trust, which has already been developed. The overlap has been previously referred to as the cognitive affective theory of learning with media, and the fundamental premise of it is that some forms of instructional media may be more interesting than others. This results in positive learning benefits as the learner will spend more effort on the task (Moreno, 2006). This could, in theory, advance their ascension of the cognitive hierarchy, as they will be able to comprehend the information and therefore apply, analyse, synthesise and evaluate the knowledge sooner than if other onscreen agents were used. As a result, this would transform their learning within the psychomotor domain from the initial stages of imitation and replication, through constructivism, as the learners would be able to develop their own mental representation of the information at a higher stage of the cognitive and affective domains. This form of generative learning could in theory enhance their progression through Dawson's (1998) psychomotor domain, as they will have increased confidence (theoretically supported by Moreno, (2006) who showed that some instructional methods may be seen to be more supportive than others at reducing anxiety and increasing confidence). This is derived from what can be termed the conversion of lower order thinking to higher order, and which addressed the concerns of Deubel, (2003).

The personalisation and voice principles state that learning is increased when narration is presented in a conversational style by a human voice. Both of these aspects where encompassed within the instructional videos, as the lead researcher who had taught the students in the first year narrated the videos in a conversational manner. Both of these principles are evidenced as having large effect sizes of 0.79 and 0.74 respectively, which have been applied as effectively as possible within the current study. Due to many of the reasons cited above, this could result in a further increase of generative learning, with intrinsic and germane cognitive load.

An aspect that has received very little scholarly attention despite the fact that CATLM was initially proposed in 2007 is the interaction between both the visual and auditory channels, in the context of the affective domain. It has been exceedingly well evidenced that the use of the visual and auditory channels has a beneficial effect upon learning (Paivio, 1991). The interaction of the information acquired within the working memory generates mental schemas, which are more effective than relying on one channel alone. Authors have referred to the function of the working memory as a cognitive process for obvious reasons. During the learning process, learners are able to begin to self-regulate, apply problem-solving strategies or modify their cognitive behaviour (Wouters et al., 2007). The mechanisms behind how this is facilitated are not clear and even the definition of germane cognitive load centres on certain cognitive elements. It is proposed that the methodology employed in the design of the instructional videos used in the current study with specific attention to the embodiment, personalisation and voice principles engaged the students in the affective domain. This resulted in an increased amount of germane cognitive load, with a simultaneous reduction in extraneous cognitive load. This was facilitated as the students are used to the communication methods (both auditory and visual, encompassing body language which could be thought of as an affective application of the pretraining principle) of the information that was presented to them. This when combined with the intrinsic load facilitated learning to a higher level, a distinctly important implication of the results from this study. This is also in line with CATLM, although the research that has been conducted to evidence the CATLM predominantly involves the use of enhanced graphics. This could be seen as increasing the extraneous load, through violation of the coherence principle due to the increased seductive detail, which is acknowledged as a potential limitation by Mayer & Estrella, (2014). As a result, there are no empirical studies that evaluate the use of the cognitive affective theory, with human onscreen agents and narrators in the way depicted within this study.

Whilst this is purely theoretical, it could explain the variation in results that have been obtained by authors investigating the efficacy of instructional videos. If they have used animated onscreen agents or unfamiliar models/instructors/narrators, then there may not be the affective connection which has been demonstrated by Shen et al., (2009) and Snelson & Elison-Bowers, (2009). Although this is within different contexts, both of them highlight the importance of emotion within instructional videos.

Additionally it is acknowledged that there is no direct way of measuring this (the only option would be through a self-report scale and the problems of reliability concerning these scales are well known (Mayer & Estrella, 2014)). It would, in theory be possible to investigate this with a comparison between two sets of identical videos, with differing onscreen agents and narrators i.e. a familiar set versus a unknown set.

Returning to observational learning, the efficacy of the videos could have been enhanced with the adherence to other instructional principles, namely the coherence, redundancy, split-attention, temporal-contiguity, pretraining, modality and multimedia principles. The other principles will be addressed following this section, although brief reasons have already been presented regarding their use in the instructional videos in Figure 11.

The coherence principle relates to the reduction of extraneous words and sounds, this principle was adhered to as much as possible within the long videos and fully in the short videos (Ely et al., 2014; Mayer & Fiorella, 2014). The long videos had a differing focus and involved larger group sizes when the exercises were being demonstrated. As a result they were filmed in larger rooms than the short videos (see figures 14 and 15) which had the resulting effect of increasing the potential for random sounds from the participants or other users of the rooms. The short videos were filmed in very controlled environments minimising the likelihood of any unwanted sounds (see figures 12 and 13).

Although there were variations between the filming of the two sets of videos, there were no significant differences identified between them to which the principle could be supported by. As a result, the adherence to the principle can only be seen as being positive, as both sets of videos, as exemplified within the results section increased mean formative scores for groups and individuals across all years of the study.

The redundancy principle suggests that the use of pictorial animation and narration is superior to the use of pictorial animation, narration and onscreen text (Austin, 2009; Samaras et al., 2006). This particular principle can be evaluated within the context of this study as it was violated and adhered to between the two sets of videos. Whilst there were other differences, namely duration it is possible to infer some effect of the principle from the results. The principle was violated for the long videos, as the use of the onscreen text was adopted to enhance viewing of the videos when it was not possible to have the sound at an audible volume, in the library for example.

It has been mentioned within the literature review that not all studies are in agreement with the theoretical foundations of this principle, as Montali & Lewandowski, (1996) identified that the use of on-screen text was advantageous for less skilled readers. The findings from the first two years (from which the comparison is possible) agree in principle with their contradiction of the principle. This is because the mean difference in formative assessment between the control group and the long group was 0.35 points higher, than the difference between the control group and the short group when students were assessed in groups. This resulted in a non-significant difference with a tiny effect size of 0.05 (-0.42 to 0.52 95% CI). Which when analysed with magnitude based inferences, identified a trivial difference due to the 95% confidence interval crossing zero. Interestingly the opposite was seen within the individual findings. As although it was not calculated specifically, the mean difference in formative scores was 0.31 higher, in favour of the short videos. This could result in a similar effect size, and magnitude based inference analysis.

Therefore, it is possible to imply that despite the differences in video type, which are acknowledged as a limitation, the findings should not be used definitively. The use of onscreen text in addition to the pictorial animation and narration is useful for some groups but not individuals. The underlying rationale as to why this may be the case is particular difficult to provide, as this aspect was not exclusively assessed and further research into this area would be warranted, especially for less capable readers.

Following the undertaking of an eye-tracking study in a module for the taught doctorate, the importance of the split-attention principle was enforced. It refers to the learner having to divide their attention between multiple sources of information before it can be appropriately processed (Ayres & Sweller, 2014; Mayer, 2014; Sweller, 2003). The principle as alluded to by Sweller, (2003) can only be adhered to if other principles relating to the presentation and timing of the sources of information are applied correctly. The potential for reducing the efficacy of the learning, by splitting the learner's attention was given careful consideration, to ensure that the differing sources were physically and temporally integrated wherever possible. When evaluating the resources from as objective a perspective as possible, it is possible to see that this is the case for the short videos. Whilst it was attempted, the execution of the principle for the long videos may not be as effective, due to the limitations of the production software used at the time. The long videos did include text on the screen and the range of options available regarding size, position, font, and colour was minimal and therefore the application was not ideal. When this is related to the split-attention and spatial contiguity principle it is possible to hypothesise that some attention would have been diverted away from the exercise. This is due to the less than ideal location of the text; however as has been proven above, there were negligible differences between the videos. An aspect that appears to have been omitted from any of the studies cited within this study thus far, is that none of the authors state the size of the viewer used to deliver the content. As the videos in the current study were optimised for streaming they only occurred a small part of the screen in the embedded media player. Whilst this may have had a negative effect on the ability of the learners to read the text i.e. it was too small, it may also have meant that their field of vision was able encompass all of the salient points of the video. Although their attention would have been divided it is possible that the learners could have been reading the text and were able to observe the exercise in the background in their periphery. This could be sufficient to form a representative mental schema in their working memory for storage in their long-term memory. The only effective way to prove this would be with the use of eye-tracking software quantifying the number of fixations and saccades, as they view the content.

Furthermore, the age of the participants within studies may have an impact on their ability to observe, process and combine information. This is an aspect that has not been discussed at any particular length within any studies, despite it being a valid construct that more experienced learners may potentially be able to process disparate sources more effectively than novices. There is however, no apparent literature to evidence this so the theory is purely conjecture.

The result of all three years' worth of data collection support the findings from the literature review of the temporal contiguity principle (Baggett, 1984; Eitel et al., 2013; Ginns, 2006). Although due to the large amount of information that can be conveyed in a video, the corresponding narration whilst initially synchronised with the pictures, may have had a short delay, as it cannot all be presented in a second. This differs to onscreen text (although it cannot all be read and processed by a learner in this time period); this however, is in line with the literature. Baggett, (1984) identified within their study, that the learners who received the narration either synchronised or with a 7 second delay performed to a superior level when compared to the other groups. The importance of this finding directly relates to the capacity of the working memory. As although it has been cited that the working memory can retain information for 15-30 seconds (Young et al., 2014), it appears that this may be reduced within the context of instructional video for reasons unknown. As those in the groups with a 14 or 21 second delay, should be encompassed within this theoretical time frame, which has important implications regarding the enforcement of the principle. As a result, it would appear that the large effect size of 1.22 might be an underestimation of the actual effect this principle could have upon learners. Although it was not directly investigated within this study, it would potentially be able to contribute to the evidence-base regarding the principle by stating that despite the duration of the video synchronization of the narration with the onscreen content is of paramount importance. This has also been identified by other authors when collating instructional principles (Mayer, 2014b; Meij & Meij, 2013).

From all of the principles identified, the pre-training principle is possibly the one of least relevance to the current study. This is because the experimental phase is contained within a module whereby all of the concepts, terminology, aims and objectives were highlighted and provided in the form of pre-reading for the students. Despite this fact the importance of the principle should be emphasised, particularly for learners attempting to acquire unfamiliar knowledge from an instructional video, when contextualised from a learner's perspective. If the learner hears a word, term or phrase, that they are unfamiliar with, then they would have to devote a proportion of their cognitive processing to this new knowledge with very little (other than possibly a pictorial reference) to aid them in this interpretation. If they successfully place the word, term or phrase, it would have at best resulted in additional and unwanted cognitive processing. At the opposing end of the spectrum, the learner could devote a huge amount of cognitive processing to the comprehension of a word, term or phrase. They could also miss key concepts of the instructional video due to this aspect occupying part of the limited capacity within their working memory. The consequences of this are clearly apparent and the learner's progress, within all three of the learning domains could be impeded. As a result no new terms were intentionally introduced within the content of the instructional videos. The only aspects that may have countered this, involved the use of specific equipment, however each of them were very obvious i.e. a non-motorised treadmill, were the name provides all of the information a learner would need. Mayer, (2014b) identified that the pretraining effect had a large effect size that can easily be appreciated. Therefore as a result of the evidence-base and the underling theory, this is a principle that although the current study did not investigate, can support as pre-training was used prior to the delivery of the videos. Although this is non-conventional the theoretical application of it would enable learners a greater duration to process the material, form mental schemas and store them in the long-term memory. They can then be retrieved and combined with new information from the instructional videos when required.

The theoretical underpinnings of the modality principle indicate that using pictorial animation and spoken narration, is the most effective use of the visuo-spatial sketchpad and phonological loop (Austin, 2009; Schmidt-Weigand et al., 2010; Schüler et al., 2011) and has been identified as having a large effect size (Mayer, 2014b). Whilst the underpinnings relate to the limited capacity of the working memory, the findings from the current study (that delivered pictorial animation, narration and onscreen text in the long videos) agree with the findings from Cheon et al., (2014) and Jong, (2009). These studies state that the efficacy of the principle is directly related to the lack of control the learner may have when viewing the instructional video. Both of the authors then state, that the use of segmenting may neutralise the modality principle. Whilst segmenting was not used directly in this study (further explanation below) the learners were able to view the video's as many times as they wished. As a result it appears that the modality principle may be losing its relevance due to the emerging use of the segmenting principle. Segmenting is used both in active or passive forms and the use of online technologies that support the flexible, repetitive type of viewing that some learner's desire. Therefore it would be prudent to state that this is one principle that potentially should not be recommended. Instructional designers should focus on others i.e. pretraining, segmenting and temporal contiguity to optimise cognitive processing.

The multimedia principle is a puzzling instructional principle given the context that it is found within. It appears that the principle is detailed to ensure that educators do not use videos purely as a source to provide scrolling text, with or without the use of static pictures or clips. As a result, the principle was typically adhered to with ease, although there were times were it was felt that a slide with text would have been appropriate to summarise key points. The slide however was narrated, thereby utilising both the visual and audio channels for the acquisition of information.

The signaling principle was not included, as although the instructional videos zoomed in on particular activities focusing the attention of the viewer to a certain extent, no specific aspects were highlighted in line with the principle. As a result, it would be remiss to include this within the evaluation of the videos, although it is certainly an aspect that could be integrated into any future work. Additionally the spatial-contiguity principle was not included, as although it was adhered to for the longer duration videos used in the first two years of data collection it was not possible to adhere to this principle. This is because text was not presented when pictorial animations were on screen to avoid violating any other principles. When text was presented, it was central to the viewer's field of vision and synchronous with narration so that both channels were receiving the information.

The segmenting principle poses a great challenge for interpretation, as the principle was not adhered to in the long duration videos. The content was continuously delivered for the duration or consciously adhered to in the short duration videos, however there may have been a role of segmentation within the videos, as a result of the production of the videos. This potential segmentation took the form of the transitional method (fading to black and then graduating the content in) that the editor used when moving from video, to the text based learning points at the end of the video. The method of fading to black is similar to the darkening of the screen employed by Spanjers et al., (2011). The duration of this unintentional pause was not as long as that used by Spanjers et al., (2011) it may have had a beneficial effect, however that is particularly difficult to infer as it was not intentional. As a result, it would be rather bold of the current study to state any claims about this principle. However the application of the principle does require further clarification, as the principle was not applied on the premise that the short videos would only have a small duration. It is unclear if this short duration is sufficient to engage the segmenting effect or not, if it is then the findings of the current study should be re-evaluated.

7.3 Overview of instructional principles

The multimedia principles cited throughout this study, were easy to understand and apply by those experienced with instructional video design. Whilst the majority of them have a sound evidence-base and the theoretical underpinnings of them are rationale, some of the principles do not appear to be of great use or are automatically adhered to through the adherence to other principles. As a result the simplification by Meij & Meij, (2013) would be more than sufficient and appropriate for most instructional designers. The use of the multimedia principles from Mayer, (2014b) provides a level of detail, which may well overwhelm those who are unfamiliar with the topic and potentially put them off designing and developing instructional videos. Particularly if the designer seeks to implement all of the advanced principles of multimedia learning as well, as this could potentially create a resource that would be confusing to the learner. Regardless of the principles used however, it is clear that there is greater need to convey a recommended set of instructional principles to the educational world. This is because there was distinct variability in the videos used within the empirical studies cited in this study. Whilst this study was by no means perfect (particularly the long videos) the short videos adhered to as many of the principles as possible, they demonstrated a clear educational benefit to the learners across the duration of the study. This finding is contrary to the multitude of other studies that identified, at best that there was no difference between the control and instructional video conditions. This demonstrates that they were an effective alternate method, but did not substantially improve students learning. This can in part be attributed to the use of the multimedia principles and the different educational setting, which shall be elaborated on further below. Earlier within the study it was questioned what the effect of adopting a large number of principles would be, as the vast majority of empirical studies investigated the use of a single or combination of a few instructional principles. Whilst it would have been implausible for them to have an additive effect (as it would result in an extremely large effect size) the overall effect size identified in this study, falls somewhere in the middle of the range that has been identified for each of the principles. Therefore, this is one of the first studies to demonstrate that the use of a large number of principles can also have a beneficial effect on learning.

7.4 The effect of video duration

As identified in the literature review there is very little evidence regarding the optimal duration for instructional videos, however it is clear that an upper limit of 20 minutes can be applied. The two durations of videos in this study; short (55 to 118 seconds) and long (10.29 to 18.14 minutes) contrasted substantially in duration. Whilst the long videos were not any longer than 20 minutes, the longest was almost twenty times longer than the shortest duration short video. As detailed in the results section, the first two years of data collection in which both durations were used identified no significant difference between the short or long videos. The longer duration videos obtained a mean formative score 0.65 (95% CI -3.29 to 2.59) higher than the short duration videos. The finding is surprising as the longer duration videos violated a number of instructional principles. They should have, in theory resulted in a lower score than the short duration videos, but this adds to the limited evidence-base. As a result it is not yet possible to answer the research question relating to duration, or to contribute definitively to the evidence-base regarding the optimal duration. However, the use of the short duration videos in the final year of data collection would in theory support the use of shorter videos, given the magnitude of the effect identified.

7.5 Level of learning

The main argument against the theoretical use of instructional videos within higher education, has been that it is not possible to obtain higher level learning from a passive ‘behaviourist’ media (Deubel, 2003). However the range of the literature identified in the instructional videos section demonstrated that they are used at differing educational levels from undergraduate (McCutcheon et al., 2014) to doctoral level students (Moore & Smith, 2012; Duijn et al., 2014). Although it is acknowledged that the majority of skills taught within this range are at the lower level; the acquisition of hand washing techniques by Bloomfield et al., (2010) as an example, some are not, for instance the assessment and treatment of the cervical spine injuries by Duijn et al., (2014). When the use of the psychomotor domains is reviewed, it is possible to generalise and state that all instructional videos will be targeted at the lower level of the hierarchy initially. They can all be seen through the simplified lens of behaviorism, as the educators will be looking for an observable change in behaviour (Skinner, 1974). Despite this however the context, educational level of the learner and the required final application of the instructional video has potentially the most significant bearing on the level of learning that is obtained. Particularly if the instructional video is directly linked to a summative assessment, given the trend for learners to become assessment driven. As a result, it would be of great interest if two groups of students, one at undergraduate and one at postgraduate with the same (theoretical) background knowledge were shown the same instructional videos and evaluated on the same measure. Hypothetically it would be expected that the postgraduate students would ascend the psychomotor hierarchy faster, due to the differing educational approach and experiences they have undertaken to get to the current level. This could be seen as an application of constructivism, as a problem-solving approach is most likely to be used with this level of student. Whilst the undergraduates may not fully engage with the challenge, the context of this problem may only be seen as the need to achieve a successful summative assessment. Therefore the motivation of the two educational levels may also differ. Although this is only hypothetical, the argument from Deubel, (2003) should *not* be sufficient reason not to use instructional videos, as there are a number of other factors that have to be considered. The current study engaged the students in the cognitive, affective and psychomotor domains, an approach that was clearly effective and one that is believed to have resulted in constructivist learning, exemplified by the carryover effect from the final year of data collection. If the learner was not able to construct their own appropriate mental representations from the instructional videos, the resulting effect should have produced a decrease or at best a plateau in formative scores. However as can be observed from Figure 19, Group B – Control generally increased their mean formative scores throughout the second phase of the data collection period. It is this fact that demonstrates that the learners had reached appropriate levels in all three domains, that they had engaged with the information and that they were able to construct mental schemas that could be stored and retrieved for later use. In some cases, up to nine weeks after the last time they were able to access the resource. As a result it seems that it is entirely appropriate to refute the statement by Deubel (2003), and to state that higher-level learning can be achieved from instructional videos. This is stated with the caveat that they adhere to multimedia instructional principles, that encompass the cognitive, affective and psychomotor domains, and with the further caveat that this is potentially context specific.

7.6 Groups versus Individuals

Across all years of data collection, it is apparent that the instructional videos had an effect, with scores obtained by groups and individuals in the experimental group(s) being greater than those who were in the control group.

In the first two years, the trend across the three conditions, between groups and individuals is comparable with a larger mean difference being observed in favour of the long videos for individuals (Figures 13 and 14). The trend within the third year between groups and individuals is also comparable, but yet again larger mean differences are observed for individuals (tables 4 and 5).

As this study is unique, in that it assessed groups, rather than individuals as previously employed by studies (Kirschner, Paas, & Kirschner, 2009b) the findings cannot be directly compared to those who assessed individual achievement in a group setting. Nevertheless it is apparent from inferential observation, that in the first two years, achievement levels whether in groups or as an individual did not vary dramatically. As each individual participated in different groups throughout the course of each condition of the study it would be expected that there would be a greater degree of variation between the mean scores obtained. This is only apparent however when the confidence intervals are observed. When in groups the confidence intervals between control and long were -1.41 to 7.86, and between control and short were -1.72 to 7.48. The individuals confidence intervals between control and long were -21.72 to 22.39 and for control and short were -21.02 to 22.32.

In the third year of data collection, the differences are just as apparent, particularly due to the increased level of detail achieved with the analysis of each component of the score, rather than just the total formative score. As above the mean differences did not vary dramatically between groups and individuals. However the confidence intervals are substantially larger, with the individuals having a much greater range that can be seen between tables 4 and 5. This vast increase in confidence intervals for the mean differences for individuals, demonstrates the extent that the individual achievement varied amongst the population within the study.

When contextualised, the findings do not provide evidence to support the benefit of groups or individuals, as the results of the mean differences generally did not vary sufficiently. The one aspect that this study highlights is that when groups are assessed it may not be a true reflection of the abilities of all. This may explain why traditionally groups are assessed in terms of their effectiveness with the use of individual assessment items. As such this should be sufficient evidence to justify the systematic design processes of studies, although the use of a dual form or assessment (group based and individual) could be warranted (Kirschner et al., 2009b).

As a result, it is possible to answer the remaining research questions; instructional videos do have an effect (generally beneficial) on group learning when assessed formatively. On the whole the effect is beneficial resulting in a net increase in mean formative scores, the magnitude of the effect is also educationally significant.

Instructional videos do have an effect on individuals, when being taught as part of a group, when being assessed in a group setting formatively; however the magnitude of this effect is predominantly trivial. This is concluded as the instructional videos will substantially enhance the achievement of some learners, but provide the opposite effect for others, as exemplified with the substantial variation between confidence intervals.

7.7 Ethical considerations

An important aspect of educational studies is that it is ethically sound and conducted in a way that does not advantage or disadvantage any student or groups of students. The current study whilst being more complex than Duijn et al., (2014) with a greater number of videos and skills presented over an extended duration, identified similar ethical aspects following a cross over design. Duijn et al., (2014) identified that the order that students experienced the experimental groups in their study, had no significant effect on the students, which emphasises the ethical efficacy of the crossover design. This finding can also be seen within this study. Especially during the final year of data collection seen in Figure 18, when the video group began to attain the same scores as those who had received the instructional videos in the first phase of the study. As a result, the use of a crossover design in educational studies should continue to be recommended. However, as exemplified again by the final year of data collection, there is a necessity to incorporate a washout period between experimental conditions if feasibly possible. This is an aspect that very few of the studies identified within the instructional video section employed, and is a methodological limitation across each of them. One of the principal reasons that the omission of a washout period could be attributed to, is the time-pressured environment within an academic year as there are only a finite number of teaching weeks. Even a short investigation such as the two week study conducted by Moore & Smith, (2012) could be significantly influenced by a washout period. The students would typically be taught other material or have a week without teaching in between experimental conditions. Both of which could impact in an unpredictable way, as they could in theory either lose or consolidate the information delivered in the first week, which would then bias any follow up tests. As a result, it should be stated that in an ideal scenario a crossover design should be employed with a washout period. Participants should therefore be assessed at the end of the first phase and after the washout period to re-establish the baseline so that any bias could be accounted for if identified (Wellek & Blettner, 2012).

7.8 Limitations

The current study has identified a range of limitations from assessment to dissemination; the former has had the largest effect on the study although how much of an effect is unknown. The study used an unvalidated assessment scale that was based on criterion taken from the end of year summative assessment, for the module in which the study was being conducted (appendix 1). The assessment method relied on the subjective quantification by the lecturers teaching the module to assess the students on five separate criteria. Although measures such as blinding and ensuring the same lecturers assessed the same groups throughout were employed; the use of a subjective measure is not without error.

The dissemination method of the videos varied throughout the study, which could have had an effect on viewing figures, and frequency. Although this could not be objectively quantified as the access data from all years of data collection was unreliable. In the first year the VLE was generating entirely random access data, and as such the use of a bespoke website for year two was sought. Unfortunately, the access file corrupted due to the volume, frequency and overlapping of traffic generated by the learners. In the final year the VLE was reinstated however, it was not able to generate reliable access statistics and therefore an aspect of data analysis was not possible.

In the first two years, summative performance data was obtained and correlated against the formative scores. This was not possible in the final year as the lead researcher left the University where the study was being conducted at the end of the data collection period, prior to the summative assessments being conducted.

The use of the quasi-experimental design, particularly the amalgamation of seminar groups to form groups applies a degree of non-randomness, which does not appropriately account for differing abilities between students. However, the allocation of the students into seminar groups by alphabetical classification, does apply some form of randomisation but this is an assumption, as it was not tested prior to data collection each year.

7.9 Conclusion

7.9.1 Research questions

The study has been able to answer all of the research questions posed to varying degrees. The study is able to state that the use of instructional videos (designed to adhere to multimedia instructional principles) when used in addition to traditional resources has an educationally significant impact for learners on their cognitive, affective and psychomotor performance in a practical setting. The extent of this impact varies between groups and individuals, with groups demonstrating better consistency in terms of performance than individuals.

The study has demonstrated that it is possible to foster higher levels of generative learning from instructional videos. Through the literature review it is also possible to state that instructional videos can be used across all levels in higher education, across a wide range of subject areas.

The study has shown that there is no distinct difference between differing durations of instructional videos; however the use of shorter videos has a substantially greater evidence-base, be that through short video clips or segmentation.

7.9.2 Recommendations

When designing instructional videos, multimedia principles should be adhered to, to maximise the efficacy of the resource regardless of the domain(s) they are targeted at. The guidelines by Meij & Meij, (2013) should be designers first port of call given their clarity and simplicity, should designers wish to explore elements further then they should engage with the detailed principles from Mayer, (2014b).

Whilst instructional videos typically were only used for cognitive or knowledge content their usage should be encouraged for every aspect of education. Particularly in the acquisition of psychomotor skills, as the benefits of videos over a live demonstration has been shown from a learning and University perspective. Wherever possible the use of multiple domains should be encouraged, particularly for the medical or health subjects were the use of the cognitive-affective model could be employed.

When applying a crossover design for any educational experiment the use of washout period should be included if possible, to enable the control group to act as a true control in the final phase of the study.

7.12 Concluding remarks

The current study has contributed additional evidence to the use of instructional videos in a non-traditional setting, with a novel assessment method to be beneficial to learners for cognitive, affective and psychomotor domains. Although the control groups were solely reliant on traditional resources and were not receiving an intervention of any description, the comparison enables researchers to clearly identify the potential that instructional videos have. This is an aspect that has been lacking from most previous studies, as they have proven videos to be as effective as another methodology, which does not truly, demonstrate their potential. It is hoped that as a result of the publications which have/will stem from this study, that further research into the use of instructional videos will be undertaken in differing subject areas, to support the use of instructional videos in higher education.

References

Adesope, O. O., & Nesbit, J. C. (2012). Verbal redundancy in multimedia learning environments: A meta-analysis. *Journal of Educational Psychology*, *104*(1), 250–263. doi:10.1037/a0026147

Ahn, J. Y., Cho, G. C., Shon, Y. D., Park, S. M., & Kang, K. H. (2011). Effect of a reminder video using a mobile phone on the retention of CPR and AED skills in lay responders. *Resuscitation*, *82*(12), 1543–1547. doi:10.1016/j.resuscitation.2011.08.029

Allen Moore, W., & Russell Smith, A. (2012). Effects of video podcasting on psychomotor and cognitive performance, attitudes and study behaviour of student physical therapists. *Innovations in Education and Teaching International*. doi:10.1080/14703297.2012.728876

Alonso, F., Lopez, G., Manrique, D., & Vines, J. M. (2008). Learning objects, learning objectives and learning design. *Innovations in Education and Teaching International*, *45*(4), 389–400. doi:10.1080/14703290802377265

Alonso, F., López, G., Manrique, D., & Viñes, J. M. (2005). An instructional model for web-based e-learning education with a blended learning process approach. *British Journal of Educational Technology*, *36*(2), 217–235.

Amer, A. (2006). Reflections on Bloom's revised taxonomy. *Electronic Journal of Research in Educational Psychology*, *4*(1), 213–230. Retrieved from http://www.investigacion-psicopedagogica.org/revista/articulos/8/english/Art_8_94.pdf

Anderson, L., Krathwohl, R., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., … Wittrock, M. (2001). *Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy.* New York: Longman.

Applefield, J., & Huber, R. (2000). Constructivism in theory and practice: Toward a better understanding. *The High School Journal*, *84*(2), 35–53. Retrieved from http://www.jstor.org/stable/10.2307/40364404

Artino, a. R. J. (2008). Cognitive load theory and the role of learner experience: An abbreviated review for educational practitioners. *Association for the Advancement of Computing In Education Journal (AACE) Journal*, *16*(4), 425–439. doi:10.1016/S0959

Atkinson, R. C., Brelsford, J. W., & Shiffrin, R. M. (1967). Multiprocess models for memory with applications to a continuous presentation task. *Journal of Mathematical Psychology*, *4*(2), 277–300. doi:10.1016/0022-2496(67)90053-3

Atkinson, R. C., & Shiffrin, R. M. (1965). Mathematical models for memory and learning. In D. A. Kimble (Ed.), *Third conference on Learning, remembering and forgetting*.

Attstro, R. (2005). Computer-mediated instructional video: a randomised controlled trial comparing a sequential and a segmented instructional video in surgical hand wash. *European Journal of Dental Education*, *9*, 53–58.

Attstro, R., Schittek Janda, M., Tani Botticelli, A., Mattheos, D., Nebel, D., Wagner, A., … Attstrom, R. (2005). Computer-mediated instructional video: a randomised controlled trial comparing a sequential and a segmented instructional video in surgical hand wash. *European Journal of Dental Education*, *9*, 53–58. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0579.2004.00366.x/full

Austin, K. A. (2009). Multimedia learning: Cognitive individual differences and display design techniques predict transfer learning with multimedia learning modules. *Computers & Education*, *53*(4), 1339–1354. doi:10.1016/j.compedu.2009.06.017

Ayres, P. (2015). State-of-the-Art Research into Multimedia Learning: A Commentary on Mayer's Handbook of Multimedia Learning. *Applied Cognitive Psychology*, n/a–n/a. doi:10.1002/acp.3142

Ayres, P., Marcus, N., Chan, C., & Qian, N. (2009). Learning hand manipulative tasks: When instructional animations are superior to equivalent static representations. *Computers in Human Behavior*, *25*(2), 348–353. doi:10.1016/j.chb.2008.12.013

Ayres, P., & Paas, F. (2012). Cognitive load theory: New directions and challenges. *Applied Cognitive Psychology*, *26*(6), 827–832. doi:10.1002/acp.2882

Ayres, P., & Sweller, J. (2005). The Split-Attention Principle in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 135–146). New York: Cambridge University Press.

Ayres, P., & Sweller, J. (2014). The Split-Attention Principle in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handboook of Multimedia Learning* (Second edi., pp. 206–226). Cambridge University Press.

Baddeley. (1979). The concept of working memory: a view of its current state and probable future development. *Cognition*, *10*(1-3), 17–23. doi:10.1016/0010-0277(81)90020-2

Baddeley, a D. (2001). Is working memory still working? *The American Psychologist*, *56*(11), 851–864. doi:10.1037/0003-066X.56.11.851

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. doi:10.1016/S0021-9924(03)00019-4

Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In G. A. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 47–89). Academic Press.

Baggett, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations. *Journal of Educational Psychology*, *76*(3), 408–417. doi:10.1037/0022-0663.76.3.408

Baker, C. (2010). The impact of instructor immediacy and presence for online student affective learning, cognition, and motivation. *The Journal of Educators Online*, *7*(1), 1–30.

Batterham, A. M., & Hopkins, W. G. (2006). Making meaningful inferences about magnitudes. *International Journal of Sports Physiology and Performance*, *1*(1), 50–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19114737

Becker, L. (2000). Effect size calculator. *College of Letters, Arts, and Sciences, University of Colorado, Colorado Springs*, (1993). Retrieved from http://web.uccs.edu/lbecker/Psy590/es.htm

Benedict, L., & Pence, H. E. (2012). Teaching Chemistry Using Student-Created Videos and Photo Blogs Accessed with Smartphones and Two-Dimensional Barcodes. *Journal of Chemical Education*, *89*(4), 492–496. doi:10.1021/ed2005399

Berk, R. A. (2009). Multimedia Teaching with Video Clips : TV , Movies , YouTube , and mtvU in the College Classroom. *International Journal of Technology in Teaching and Learning*, *5*(1), 1–21. doi:10.1016/j.sbspro.2010.12.326

Bloom, B. S. (1956). *Taxonomy of Educational Objectives*. (J. P. Keeves, Ed.)*Educational and Psychological Measurement* (Vol. 16). Longman London. doi:10.1177/001316445601600310

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: David McKay.

Bloomfield, J., Roberts, J., & While, A. (2010). The effect of computer-assisted learning versus conventional teaching methods on the acquisition and retention of handwashing theory and skills in pre-qualification nursing students: a randomised controlled trial. *International Journal of Nursing Studies*, *47*(3), 287–294. doi:10.1016/j.ijnurstu.2009.08.003

Boucheix, J.-M., & Lowe, R. K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction*, *20*(2), 123–135. doi:10.1016/j.learninstruc.2009.02.015

Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*(10), 489–95. doi:10.1016/j.tics.2005.08.007

British Educational Research Association. (2004). *Revised Ethical Guidelines for Educational Research (2004). Educational Research*. British Educational Research Association. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:REVISED+ETHICAL+GUIDELINES+FOR+EDUCATIONAL+RESEARCH#0

Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., & Freund, H. J. (2004). Neural Circuits underlying imitation learning of hand actions: an event related fMRI study. *Neuron*, *42*(2), 323–334.

Burton, J. K., Moore, D. M., & Magliaro, S. G. (2003). BEHAVIORISM AND

INSTRUCTIONAL TECHNOLOGY. In *Handbook of research on educational communications and technology* (pp. 3–36). Retrieved from http://www.aect.org/edtech/ed1/01.pdf

Butcher, K. (2014). The Multimedia Principle. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 174–205). Cambridge University Press.

Cannon, H. M., & Feinstein, A. H. (2005). Bloom beyond Bloom: Using the revised taxonomy to develop experiential learning strategies. *Developments in Business Simulations and Experiential Exercises*, *32*, 348–356.

Caspi, A., Gorsky, P., & Privman, M. (2005). Viewing comprehension: Students' learning preferences and strategies when studying from video. *Instructional Science*, *33*(1), 31–47. doi:10.1007/s11251-004-2576-x

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2014). Handbook of Human Centric Visualization, 551–580. doi:10.1007/978-1-4614-7485-2

Castro-alonso, J. C., Ayres, P., & Paas, F. (2014). Handbook of Human Centric Visualization, 551–580. doi:10.1007/978-1-4614-7485-2

Cate, T. J. Ten. (2000). Summative assessment of medical students in the affective domain. *Medical Teacher*, *22*(1), 40–43. doi:10.1080/01421590078805

Chan, Y. M. (2010). Video instructions as support for beyond classroom learning. *Procedia - Social and Behavioral Sciences*, *9*, 1313–1318. doi:10.1016/j.sbspro.2010.12.326

Chandler, P. (2004). The crucial role of cognitive processes in the design of dynamic visualizations. *Learning and Instruction*, *14*(3), 353–357. doi:10.1016/j.learninstruc.2004.06.009

Chen, E. (2009). Empowering College Students' Research Skills via Digital Media Elaine. In *Brick and Click Libraries Symposium Proceedings*.

Chen, J.-L. (2011). The effects of education compatibility and technological expectancy on e-learning acceptance. *Computers & Education*, *57*(2), 1501–1511. doi:10.1016/j.compedu.2011.02.009

Chen, S. (2007). Instructional design strategies for intensive online courses: An objectivist-constructivist blended approach. *Journal of Interactive Online Learning*, *6*(1), 72–86. Retrieved from http://www.unhas.ac.id/hasbi/LKPP/Hasbi-KBK-SOFTSKILL-UNISTAFF-SCL/Mental Model/konstruktivisme2.pdf

Chenkin, J., Lee, S., Huynh, T., & Bandiera, G. (2008). Procedures can be learned on the Web: a randomized study of ultrasound-guided vascular access training. *Academic Emergency Medicine : Official Journal of the Society for Academic Emergency Medicine*, *15*(10), 949–54. doi:10.1111/j.1553-2712.2008.00231.x

Cheon, J., Crooks, S., & Chung, S. (2014). Does segmenting principle counteract modality principle in instructional animation? *British Journal of Educational Technology*, *45*(1), 56–64. doi:10.1111/bjet.12021

Chi, D. L., Pickrell, J. E., & Riedy, C. a. (2014). Student learning outcomes associated with video vs. paper cases in a public health dentistry course. *Journal of Dental Education*, *78*(1), 24–30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24385521

Chilwant, K. S. (2012). Comparison of two teaching methods , structured interactive lectures and conventional lectures . *Biomedical Research*, *23*(3), 363–366.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. doi:10.1007/BF01320076

Clark, R. E., & Feldon, D. F. (2005). Five common but questionable principles of multimedia learning. In *Cambridge Handbook of Multimedia Learning* (pp. 1–22). doi:10.1017/CBO9781139547369.009

Cleveland-Innes, M., & Ally, M. (2013). Affective learning outcomes in workplace training: A test of synchronous vs. asynchronous online learning environments. *Canadian Journal of …*, *30*(1), 15–35. Retrieved from http://wigan-ojs.library.ualberta.ca/index.php/cjuce-rcepu/article/view/20052

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Retrieved from http://books.google.com/books?hl=en&lr=&id=Tl0N2lRAO9oC&oi=fnd&pg=PR11&dq=Statistical+Power+Analysis+for+the+Behavioral+Sciences&ots=dp0CVmk0Wr&sig=LYVnEDt7c3mNR1ljNCdFWbGeb7Q

Cooper, D., & Higgins, S. (2014). The effectiveness of online instructional videos in the acquisition and demonstration of cognitive, affective and psychomotor rehabilitation skills. *British Journal of Educational Technology*, *46*(4), 768–779. doi:10.1111/bjet.12166

Cooper, D., Higgins, S., & Beckmann, N. (2015). Online instructional videos as a complimentary method of teaching practical rehabilitation skills for groups and individuals. *British Journal of Educational Technology*, *Under Revi*.

Croft, D. E. (2010). The Safe and Proper Use of Wrenches: An Instructional Video. In *Educational Technology* (pp. 1–11). Technology, Colleges, and Community Worldwide Online Conference. Retrieved from http://scholarspace.manoa.hawaii.edu/handle/10125/15327

Dave, R. H. (1970). *Developing and Writing Behavioural Objectives.* (R. J. Armstrong, Ed.). Tucson: Educational Innovators Press.

Dawson, W. (1998). *Extensions to Bloom's taxonomy of educational objectives*.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, *20*(2), 111–122. doi:10.1016/j.learninstruc.2009.02.010

Deubel, P. (2003). An Investigation of Behaviorist and Cognitive Approaches to Instructional Multimedia Design. *Journal of Educational Multimedia and Hypermedia*, *12*(1), 63–90.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, *91*, 176–180.

Dijksterhuis, A., Nordgren, L. F., Doolittle, P. E., Sweller, J., Merrienboer, J. Van, FGWC, … Van Gerven, P. W. M. (2009). Research on cognitive load theory and its design implications for e-learning. *Learning and Instruction*, *38*(2), 1–10. doi:10.1002/acp

Doolittle, P. (2001). Multimedia learning: Empirical results and practical applications. *Retrieved August*. Retrieved from http://scr.csc.noctrl.edu/courses/edn509/resources/readings/multimediaLearningEmpericalResults.pdf

Drinkwater, E. J., Pyne, D. B., & McKenna, M. J. (2008). Design and interpretation of anthropometric and fitness testing of basketball players. *Sports Medicine (Auckland, N.Z.)*, *38*(7), 565–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18557659

Duijn, A. J. Van, Swanick, K., & Donald, E. K. (2014a). Student Learning of Cervical Psychomotor Skills Via Online Video Instructio...: EBSCOhost, *28*(1). Retrieved from http://web.b.ebscohost.com/ehost/detail?vid=4&sid=16ae2643-ae98-4038-9c4c-3733637e105d@sessionmgr110&hid=118&bdata=Jmxhbmc9ZXMmc2l0ZT1laG9zdC1saXZl#db=rzh&AN=2012452516

Duijn, A. J. Van, Swanick, K., & Donald, E. K. (2014b). Student Learning of Cervical Psychomotor Skills Via Online Video Instructio...: EBSCOhost, *28*(1). Retrieved from http://web.b.ebscohost.com/ehost/detail?vid=4&sid=16ae2643-ae98-4038-9c4c-3733637e105d%40sessionmgr110&hid=118&bdata=Jmxhbmc9ZXMmc2l0ZT1laG9zdC1saXZl#db=rzh&AN=2012452516

Durlak, J. a. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology*, *34*(9), 917–28. doi:10.1093/jpepsy/jsp004

Eccles, D. W., & Feltovich, P. J. (2008). Implications of Domain-General “Psychological Support Skills” for Transfer of Skill and Acquisition of Expertise. *Performance Improvement Quarterly*, *21*(1), 43–60. doi:10.1002/piq

Educanon. (2014). *No Title*. Retrieved from https://www.educanon.com

Einspruch, E. L., Lynch, B., Aufderheide, T. P., Nichol, G., & Becker, L. (2007). Retention of CPR skills learned in a traditional AHA Heartsaver course versus 30-min video self-training: a controlled randomized study. *Resuscitation*, *74*(3), 476–86. doi:10.1016/j.resuscitation.2007.01.030

Eitel, A., Scheiter, K., Schüler, A., Nyström, M., & Holmqvist, K. (2013). How a picture facilitates the process of learning from text: Evidence for scaffolding. *Learning and Instruction*, *28*, 48–63. doi:10.1016/j.learninstruc.2013.05.002

Ellis, P. D. (2010). *The Essential Guide to Effect Sizes: Statistical Power,* *Meta-Analysis, and the Interpretation of Research Results*. Cambridge University Press. Retrieved from www.cambridge.org/9780521142465

Ely, E., Pullen, P. C., Kennedy, M. J., Hirsch, S. E., & Williams, M. C. (2014). Use of instructional technology to improve teacher candidate knowledge of vocabulary instruction. *Computers and Education*, *75*, 44–52. doi:10.1016/j.compedu.2014.01.013

Evans, C. (2008). The effectiveness of m-learning in the form of podcast revision lectures in higher education. *Computers & Education*, *50*(2), 491–498. doi:10.1016/j.compedu.2007.09.016

Fee, S., & Fee, L. (2003). Pedagogical Approaches for the Use of Digital Video. In *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 1407–1414). Retrieved from http://www.editlib.org/index.cfm?fuseaction=Reader.ViewAbstract& paper_id=18186

Ferris, T. L. J., & Aziz, S. M. (2005). A Psychomotor Skills Extension to Bloom's Taxonomy of Education Objectives for Engineering Education. In *Exploring Innovation in Education and Research* (pp. 1–5). Tainan, Taiwan: iCEER.

Folley, D., Wilkinson, S., & Thomson, S. (2009). Rebooting the student's attention span by using Personal Response Systems, (6), 21–25. Retrieved from http://repository-intralibrary.leedsmet.ac.uk/IntraLibrary?command=open-preview&learning_object_key=i6568n36631t

Ford, G. S., Mazzone, M. a, & Taylor, K. (2005). Effect of computer-assisted instruction versus traditional modes of instruction on student learning of musculoskeletal special tests. *Journal of Physical Therapy Education*, *19*(2), 22–30. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=2009060991&site=ehost-live

Fotheringham, D. (2010). Triangulation for the assessment of clinical nursing skills: a review of theory, use and methodology. *International Journal of Nursing Studies*, *47*(3), 386–91. doi:10.1016/j.ijnurstu.2009.09.004

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of Experimental Psychology. General*, *141*(1), 2–18. doi:10.1037/a0024338

Fuller, U., Johnson, C. G., Ahoniemi, T., Cukierman, D., Hernán-Losada, I., Jackova, J., … Thompson, E. (2007). Developing a Computer Science-specific Learning Taxonomy, 152–170. doi:10.1145/1345375.1345438

Gallagher, A. M., Gilligan, R., & Mcgrath, M. (2014). The effect of DVD training on the competence of occupational therapy students in manual handling: A pilot study. *International Journal of Therapy and Rehabilitation*, *21*(12), 575–583.

Galy, E., Cariou, M., & Mélan, C. (2011). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*. doi:10.1016/j.ijpsycho.2011.09.023

Garamszegi, L. Z. (2006). Comparing effect sizes across variables: generalization without the need for Bonferroni correction. *Behavioral Ecology*, *17*(4), 682–687. doi:10.1093/beheco/ark005

Gastaut, H. J., & Bert, J. (1954). EEG changes during cinematographic presentation; moving picture activation of the EEG. *Electroencephalography and Clinical Neurophysiology*, *6*(3), 433–444. doi:10.1016/0013-4694(54)90058-9

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, *16*(6), 511–525. doi:10.1016/j.learninstruc.2006.10.001

Gog, T., Paas, F., Marcus, N., Ayres, P., & Sweller, J. (2008). The Mirror Neuron System and Observational Learning: Implications for the Effectiveness of Dynamic Visualizations. *Educational Psychology Review*, *21*(1), 21–30. doi:10.1007/s10648-008-9094-3

Gog, T., & Rummel, N. (2010). Example-Based Learning: Integrating Cognitive and Social-Cognitive Research Perspectives. *Educational Psychology Review*, *22*(2), 155–174. doi:10.1007/s10648-010-9134-7

Green, S. (2003). Evaluating the use of streaming video to support student learning in a first-year life sciences course for student nurses. *Nurse Education Today*, *23*(4), 255–261. doi:10.1016/S0260-6917(03)00014-5

Hanna, W. (2007). The New Bloom's Taxonomy: Implications for Music Education. *Arts Education Policy Review*, *108*(4), 7–17.

Harrow, A. (1972). *A Taxonomy of the Psychomotor Domain: A Guide for Developing Behavioral Objectives.* New York: McKay.

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology*, *21*(6), 713–729. doi:10.1002/acp.1345

Hattie, J. A. C. (2008). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London: Routledge.

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

Hill, G. (1982). Group versus individual performance: Are N+ 1 heads better than one? *Psychological Bulletin*, *91*(3), 517–539. Retrieved from http://psycnet.apa.org/journals/bul/91/3/517/

Hodges, N. J., Ong, N. T., Larssen, B. C., & Lim, S. B. (2011). What Observation of Motor Skills Does and Does Not Teach Us. *BIO Web of Conferences*, *1*, 00034. doi:10.1051/bioconf/20110100034

Hollender, N., Hofmann, C., Deneke, M., & Schmitz, B. (2010a). Integrating cognitive load theory and concepts of human-computer interaction. *Computers in Human Behavior*, *26*(6), 1278–1288. doi:10.1016/j.chb.2010.05.031

Hollender, N., Hofmann, C., Deneke, M., & Schmitz, B. (2010b). Integrating cognitive load theory and concepts of human–computer interaction. *Computers in Human Behavior*, *26*(6), 1278–1288. doi:10.1016/j.chb.2010.05.031

Hopkins, W. G. (2007). A Spreadsheet for Deriving a Confidence Interval , Mechanistic Inference and Clinical Inference from a P Value. *North*.

Hopkins, W. G. (2015). *Sportscience*. *sportsci.org*. Retrieved from http://sportsci.org/index.html

Howe, K. R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher*, *17*(8), 10. Retrieved from http://edr.sagepub.com/content/17/8/10.short

Hung, D. (2001). Theories of learning and computer-mediated instructional technologies. *Educational Media International*. doi:10.1080/0952398011010511

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews. Neuroscience*, *7*(12), 942–51. doi:10.1038/nrn2024

Ibrahim, M. (2012). Implications of designing instructional video using cognitive theory of multimedia learning. *Critical Questions in Education*, *3*(2), 83–104. Retrieved from http://ts.isil.westga.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=90486720&site=eds-live&scope=site

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the Black Box of Causality: Learning About Causal Mechanisms from Experimental and Observational Studies. *American Political Science Review*, *105*(04), 765–789. doi:10.1017/S0003055411000414

Isaacs, G. (1996). *Bloom's taxonomy of educational objectives*. *University of Queensland, Teaching and Educational Development Institute*.

Iserbyt, P., Liesbet, M., Jan, E., & Daniel, B. (2012). Multimedia design principles in the psychomotor domain: The effect of multimedia and spatial contiguity on students' learning of Basic Life Support with task cards. *Journal of Educational Multimedia and Hypermedia*, *21*(2), 111–125. Retrieved from http://www.editlib.org/p/38611/

Issa, N., Mayer, R. E., Schuller, M., Wang, E., Shapiro, M. B., & DaRosa, D. a. (2013). Teaching for understanding in medical classrooms using multimedia design principles. *Medical Education*, *47*(4), 388–96. doi:10.1111/medu.12127

Issa, N., Schuller, M., Santacaterina, S., Shapiro, M., Wang, E., Mayer, R. E., & Darosa, D. a. (2011). Applying multimedia design principles enhances learning in medical education. *Medical Education*, *45*(8), 818–26. doi:10.1111/j.1365-2923.2011.03988.x

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, *17*(02), 187. doi:10.1017/S0140525X00034026

Jelovsek, J. E., Kow, N., & Diwadkar, G. B. (2013). Tools for the direct observation and assessment of psychomotor skills in medical trainees: a systematic review. *Medical Education*, *47*(7), 650–73. doi:10.1111/medu.12220

Jong, T. (2009). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science*, *38*(2), 105–134. doi:10.1007/s11251-009-9110-0

Jowett, N., LeBlanc, V., Xeroulis, G., MacRae, H., & Dubrowski, A. (2007). Surgical skill acquisition with self-directed practice using computer-based video training. *American Journal of Surgery*, *193*(2), 237–242. doi:10.1016/j.amjsurg.2006.11.003

Kala, S., Isaramalai, S.-A., & Pohthong, A. (2010). Electronic learning and constructivism: a model for nursing education. *Nurse Education Today*, *30*(1), 61–66. doi:10.1016/j.nedt.2009.06.002

Kalyuga, S. (2011). Cognitive Load Theory: How Many Types of Load Does It Really Need? *Educational Psychology Review*, *23*(1), 1–19. doi:10.1007/s10648-010-9150-7

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, *38*(1), 23–31. doi:10.1207/S15326985EP3801_4

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*(4), 351–371. doi:10.1002/(SICI)1099-0720(199908)13:4<351::AID-ACP589>3.0.CO;2-6

Kalyuga, S., Chandler, P., & Sweller, J. (2011). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *25*(SUPPL. 1). doi:10.1002/acp.1773

Keijzer, F. (2005). Theoretical behaviorism meets embodied cognition: two theoretical analyses of behavior. *Philosophical Psychology*, *18*(1), 123–143. doi:10.1080/09515080500085460

Kelly, M., Lyng, C., McGrath, M., & Cannon, G. (2009). A multi-method study to determine the effectiveness of, and student attitudes to, online instructional videos for teaching clinical nursing skills. *Nurse Education Today*, *29*(3), 292–300. doi:10.1016/j.nedt.2008.09.004

Kennedy, D., Hyland, A., & Ryan, N. (2007). *Writing and using learning outcomes: a practical guide*. *Moon*. Retrieved from http://sss.dcu.ie/afi/docs/bologna/writing_and_using_learning_outcomes.pdf

Kim, B., & Reeves, T. C. (2007). Reframing research on learning with technology: in search of the meaning of cognitive tools. *Instructional Science*, *35*(3), 207–256. doi:10.1007/s11251-006-9005-2

Kirschner, F., Paas, F. G. W. C., & Kirschner, P. a. (2009a). A cognitive-load approach to collaborative learning: United brains for complex task. *Educational Psychology Review*, *21*, 31–42. doi:10.1007/s10648-008-9095-2

Kirschner, F., Paas, F., & Kirschner, P. a. (2009b). Individual and Group-Based Learning. *Effects on Retention and Transfer Efficiency*, *25*(2), 1–38.

Kirschner, F., Paas, F., & Kirschner, P. a. (2011). Task complexity as a driver for collaborative learning efficiency: The collective working-memory effect. *Applied Cognitive Psychology*, *25*(4), 615–624. doi:10.1002/acp.1730

Kirschner, P. (2002). Cognitive load theory: implications of cognitive load theory on the design of learning. *Learning and Instruction*, *12*(1), 1–10. doi:10.1016/S0959-4752(01)00014-7

Kirschner, P. a, Ayres, P., & Chandler, P. (2011). Contemporary cognitive load theory research: The good, the bad and the ugly. *Computers in Human Behavior*, *27*(1), 99–105. doi:http://dx.doi.org/10.1016/j.chb.2010.06.025

Kirschner, P. a., Ayres, P., & Chandler, P. (2011). Contemporary cognitive load theory research: The good, the bad and the ugly. *Computers in Human Behavior*, *27*(1), 99–105. doi:10.1016/j.chb.2010.06.025

Kovacs, G. (1997). Procedural Skills in Medicine : Linking Theory To Practice. *The Journal of Emergency Medecine*, *15*(3), 387–391. Retrieved from /citations?view_op=view_citation&continue=/scholar?hl=fr&start=110&as_sdt=0,5&scilib=4&citilm=1&citation_for_view=okt2lBoAAAAJ:Wp0gIr-vW9MC&hl=fr&oi=p

Krathwohl, D. R. (2002). A Revision Bloom's Taxonomy: An Overview. *Theory into Practice*, *41*(4), 212–218.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: The classification of educational goals. Handbook II: Affective domain*. Boston, Mass.: Allyn and Bacon.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1973). *Taxonomy of Educational Objectives, the Classification of Educational Goals. Handbook II: Affective Domain.* New York: David McKay.

Kruse, N. B., & Veblen, K. K. (2012). Music teaching and learning online: Considering YouTube instructional videos. *Journal of Music, Technology and Education*, *5*(1), 77–87. doi:10.1386/jmte.5.1.77_1

Ladyshewsky, R. K. (2013). Instructor Presence in Online Courses and Student Satisfaction. *International Journal for the Scholarship of Teaching & Learning*, *7*(1), 1–23. Retrieved from http://navigator-edinboro.passhe.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=91025571&site=eds-live

Lau, K. H. V. (2014). Computer-based teaching module design: Principles derived from learning theories. *Medical Education*, *48*(3), 247–254. doi:10.1111/medu.12357

Laughlin, P. R., Hatch, E. C., Silver, J. S., & Boh, L. (2006). Groups perform better than the best individuals on letters-to-numbers problems: effects of group size. *Journal of Personality and Social Psychology*, *90*(4), 644–51. doi:10.1037/0022-3514.90.4.644

Lawlor, B., & Donnelly, R. (2010). Using podcasts to support communication skills development: A case study for content format preferences among postgraduate research students. *Computers & Education*, *54*(4), 962–971. doi:10.1016/j.compedu.2009.09.031

Lehman, S., Schraw, G., McCrudden, M. T., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology*, *32*(4), 569–587. doi:10.1016/j.cedpsych.2006.07.002

Leutner, D. (2014). Motivation and emotion as mediators in multimedia learning. *Learning and Instruction*, *29*, 174–175. doi:10.1016/j.learninstruc.2013.05.004

Liaw, S., Huang, H., & Chen, G. (2007). Surveying instructor and learner attitudes toward e-learning. *Computers & Education*, *49*(4), 1066–1080. doi:10.1016/j.compedu.2006.01.001

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small Group and Individual Learning with Technology: A Meta-Analysis. *Review of Educational Research*, *71*(3), 449–521. doi:10.3102/00346543071003449

Lusk, D. L., Evans, A. D., Jeffrey, T. R., Palmer, K. R., Wikstrom, C. S., & Doolittle, P. E. (2009). Multimedia learning and individual differences: Mediating the effects of working memory capacity with segmentation. *British Journal of Educational Technology*, *40*(4), 636–651. doi:10.1111/j.1467-8535.2008.00848.x

Lynch, D. R., Russell, J. S., Evans, J. C., & Sutterer, K. G. (2009). Beyond the Cognitive: The Affective Domain, Values, and the Achievement of the Vision. *Journal of Professional Issues in Engineering Education and Practice*, *135*(1), 47–56. doi:10.1061/(ASCE)1052-3928(2009)135:1(47)

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, *88*(1), 49–63. doi:10.1037/0022-0663.88.1.49

Mayer, R. (1999). Multimedia aids to problem-solving transfer. *International Journal of Educational Research*, *31*(7), 611–623. doi:10.1016/S0883-0355(99)00027-0

Mayer, R. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, *13*(2), 125–139. doi:10.1016/S0959-4752(02)00016-6

Mayer, R. (2005). Introduction to multimedia learning. *The Cambridge Handbook of Multimedia Learning*. Retrieved from http://books.google.com/books?hl=en&lr=&id=SSLdo1MLIywC&oi=fnd&pg=PR17&dq=Introduction+to+Multimedia+Learning&ots=uRB6J8X6Lw&sig=cqMFvAyNxjczbUlmoRsdJCAVDls

Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology*, *81*(2), 240–246. doi:10.1037/0022-0663.81.2.240

Mayer, R. E. (2002). Cognitive Theory and the Design of Multimedia Instruction: An Example of the Two-Way Street between Cognition and Instruction. *New Directions for Teaching and Learning*, (89), 55–71. doi:10.1002/tl.47

Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. *The Cambridge Handbook of Multimedia Learning*, 31–48.

doi:10.1207/s15326985ep4102_2

Mayer, R. E. (2008). Applying the science of learning: evidence-based principles for the design of multimedia instruction. *The American Psychologist*, *63*(8), 760–9. doi:10.1037/0003-066X.63.8.760

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.

Mayer, R. E. (2009). Research---Based Principles for Designing Multimedia Instruction Overview of Multimedia Instruction.

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction*, *20*(2), 167–171. doi:10.1016/j.learninstruc.2009.02.012

Mayer, R. E. (2011). *Applying the Science of Learning to Multimedia Instruction*. (J. P. Mestre & B. H. Ross, Eds.)*The Psychology of Learning and Cognition in Education*. Elsevier, Amsterdam, The Netherlands. doi:10.1016/B978-0-12-387691-1.00011-9

Mayer, R. E. (2014a). Incorporating motivation into multimedia learning. *Learning and Instruction*, *29*, 171–173. doi:10.1016/j.learninstruc.2013.04.003

Mayer, R. E. (2014b). *The Cambridge Handbook of Multimedia Learning*. (R. E. Mayer, Ed.) (Second Edi.). New York, USA: Cambridge University Press.

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), 444–452. doi:10.1037/0022-0663.84.4.444

Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction*, *33*, 12–18. doi:10.1016/j.learninstruc.2014.02.004

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), 389–395. doi:10.1037/0022-0663.96.2.389

Mayer, R. E., & Fiorella, L. (2014). Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity and Temporal Principles. In R. E. Mayer (Ed.), *Cambridge Handbook of Multimedia Learning* (Second., pp. 279–315). Cambridge University Press.

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), 187–198. doi:10.1037/0022-0663.93.1.187

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*(2), 312–320. doi:10.1037/0022-

0663.90.2.312

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review*, *14*(1), 87–99. doi:10.1023/A:1013184611077

Mayer, R. E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, *91*(4), 638–643. doi:10.1037/0022-0663.91.4.638

Mayer, R., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*(1), 107–119. doi:10.1016/S0959-4752(01)00018-4

Mayer, R., & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, *38*(1), 43–52. doi:10.1207/S15326985EP3801_6

McCombs, S., & Liu, Y. (2007). The efficacy of podcasting technology in instructional delivery. *International Journal of Technology in Teaching and …*, *3*(2), 123–134. Retrieved from http://www.sicet.org/journals/ijttl/specialIssue/youmei.pdf

Mccutcheon, K., Lohan, M., Traynor, M., & Martin, D. (2014). A systematic review evaluating the impact of online or blended learning vs. face-to-face learning of clinical skills in undergraduate nurse education. *Journal of Advanced Nursing*. doi:10.1111/jan.12509

McGrath, R. E., & Meyer, G. J. (2006). When effect sizes disagree: the case of r and d. *Psychological Methods*, *11*(4), 386–401. doi:10.1037/1082-989X.11.4.386

McNulty, J. a, Hoyt, A., Gruener, G., Chandrasekhar, A., Espiritu, B., Price, R., & Naheedy, R. (2009). An analysis of lecture video utilization in undergraduate medical education: associations with performance in the courses. *BMC Medical Education*, *9*, 1–6. doi:10.1186/1472-6920-9-6

Mcnulty, J. A., Sonntag, B., & Sinacore, J. M. (2009). Evaluation of Computer-Aided Instruction in a Gross Anatomy Course : A Six-year Study. *Anatomical Sciences Education*, *8*(2), 2–8. doi:10.1002/ase.66

Meij, H. Van Der, & Meij, J. Van Der. (2013). Eight Guidelines for the Design of Instructional Videos for Software Training. *Technical Communication*, *60*(3), 205–228.

Merkt, M., Weigand, S., Heier, A., & Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction*, *21*(6), 687–704. doi:10.1016/j.learninstruc.2011.03.004

Merriënboer, J. J. G., & Ayres, P. (2005). Research on cognitive load theory and its design implications for e-learning. *Educational Technology Research and Development*, *53*(3), 5–13. doi:10.1007/BF02504793

Merriënboer, J. J. G., & Sweller, J. (2005a). *Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions*. *Educational Psychology Review* (Vol. 17). doi:10.1007/s10648-005-3951-0

Merriënboer, J. J. G., & Sweller, J. (2005b). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, *17*(2), 147–177. doi:10.1007/s10648-005-3951-0

Montali, J., & Lewandowski, L. (1996). Bimodal reading: benefits of a talking computer for average and less skilled readers. *Journal of Learning Disabilities*, *29*(3), 271–279. doi:10.1177/002221949602900305

Moore, J. (1985). Some historical and conceptual relations among logical positivism, operationism, and behaviorism. *The Behavior Analyst*, *8*(1), 53–63. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2741771&tool=pmcentrez&rendertype=abstract

Moreno, R. (2006). When worked examples don’t work: Is cognitive load theory at an Impasse? *Learning and Instruction*, *16*(2 SPEC. ISS.), 170–181. doi:10.1016/j.learninstruc.2006.02.006

Moreno, R. (2010). Cognitive load theory: more food for thought. *Instructional Science*, *38*(2), 135–141. doi:10.1007/s11251-009-9122-9

Moreno, R., & Mayer, R. (1999a). Visual presentations in multimedia learning: Conditions that overload visual working memory. *Visual Information and Information Systems*, *1614*, 798–805. Retrieved from http://www.springerlink.com/index/01117565685641m8.pdf

Moreno, R., & Mayer, R. (2007). Interactive Multimodal Learning Environments. *Educational Psychology Review*, *19*(3), 309–326. doi:10.1007/s10648-007-9047-2

Moreno, R., & Mayer, R. E. (1999b). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*(2), 358–368. doi:10.1037/0022-0663.91.2.358

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*(1), 105–125. doi:10.1037/1082-989X.7.1.105

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews of the Cambridge Philosophical Society*, *82*(4), 591–605. doi:10.1111/j.1469-185X.2007.00027.x

Nelson, M. S., Wooditch, A., & Dario, L. M. (2014). Sample size, effect size, and statistical power: a replication study of Weisburd’s paradox. *Journal of Experimental Criminology*. doi:10.1007/s11292-014-9212-9

Nousiainen, M., Brydges, R., Backstein, D., & Dubrowski, A. (2008). Comparison of expert instruction and computer-based video training in teaching fundamental surgical skills to medical students. *Surgery*, *143*(4), 539–44. doi:10.1016/j.surg.2007.10.022

Olson, V. D. (2008). Instruction Of Competent Psychomotor Skill. *College Teaching Methods & Styles Journal*, *4*(9), 27–30.

Onwuegbuzie, A. J. (2000). Positivists, Post-Positivists, Post-Structuralists, and Post-Modernists: Why Can't We All Get Along? Towards a Framework for Unifying Research Paradigms. *Annual Meeting of the Association for the Advancement of Educational Research*, 25.

Ouyang, J. R., & Stanley, N. (2014). Theories and research in educational technology and distance learning instruction through Blackboard. *Universal Journal of Educational Research*, *2*(2), 161–172. doi:10.13189/ujer.2014.020208

Oztop, E., Kawato, M., & Arbib, M. (2006). Mirror neurons and imitation: a computationally guided review. *Neural Networks : The Official Journal of the International Neural Network Society*, *19*(3), 254–71. doi:10.1016/j.neunet.2006.02.002

Öztürk, D., & Dinç, L. (2014). Effect of web-based education on nursing students' urinary catheterization knowledge and skills. *Nurse Education Today*, *34*(5), 802–8. doi:10.1016/j.nedt.2013.08.007

Paas, F., Camp, G., & Rikers, R. (2001). Instructional compensation for age-related cognitive declines: Effects of goal specificity in maze learning. *Journal of Educational Psychology*, *93*(1), 181–186. doi:10.1037/0022-0663.93.1.181

Paas, F., Gog, T., & Sweller, J. (2010). Cognitive Load Theory: New Conceptualizations, Specifications, and Integrated Research Perspectives. *Educational Psychology Review*, *22*(2), 115–121. doi:10.1007/s10648-010-9133-8

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, *38*(1), 1–4. doi:10.1207/S15326985EP3801_1

Paas, F., & Sweller, J. (2011). An Evolutionary Upgrade of Cognitive Load Theory: Using the Human Motor System and Collaboration to Support the Learning of Complex Cognitive Tasks. *Educational Psychology Review*, 27–45. doi:10.1007/s10648-011-9179-2

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*, *38*(1), 63–71. doi:10.1207/S15326985EP3801_8

Paivio, A. (1991). Dual Coding Theory: Retrospect and Current Status. *Canadian Journal of Psychology*, *45*(3), 255–287.

Pavlov, I. P. (1927). *Conditioned reflexes*. London: Clarendon Press.

Perry, S. (1999). Strategies for teaching clinical decision-making skills in neurologic content: Perspective from the field. *Journal of Neurologic Physical Therapy*, *23*(4), 170–177. Retrieved from http://journals.lww.com/jnpt/Abstract/1999/23040/Strategies_for_Teaching_Clinical_Decision_Making.12.aspx

Piaget, J. (1959). *The language and thought of the child*. Psychology Press.

Pickard, M. (2007). The new Bloom's taxonomy: An overview for family and consumer sciences. *Journal of Family and Consumer Sciences Education*, *25*(1), 45–55. Retrieved from http://natefacs.org/JFCSE/v25no1/v25no1Pickard.pdf

Plaisant, C., & Shneiderman, B. (2005). Show me! Guidelines for recorded demonstration. In *2005 IEEE Symposium on Visual Languages and Human-Centric Computing*. Dallas, Texas.

Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction*, *29*, 128–140. doi:10.1016/j.learninstruc.2013.02.006

Pociask, F., & Morrison, G. (2004). The Effects of Split-Attention and Redundancy on Cognitive Load When Learning Cognitive and Psychomotor Tasks. *Association for Educational Communications*, 707–718. Retrieved from http://eric.ed.gov/ERICWebPortal/recordDetail?accno=ED485075

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, *12*(1), 61–86. doi:10.1016/S0959-4752(01)00016-0

Radin Salim, K., Puteh, M., & Mohd Daud, S. (2011). Levels of practical skills in basic electronic laboratory: Students' perceptions. *2011 IEEE Global Engineering Education Conference, EDUCON 2011*, 231–235. doi:10.1109/EDUCON.2011.5773142

Reed, S. (2006). Cognitive Architectures for Multimedia Learning. *Educational Psychologist*, *41*(2), 87–98. doi:10.1207/s15326985ep4102_2

Reeves, T. (2006). How do you know they are learning? The importance of alignment in higher education. *International Journal of Learning Technology*, *2*(4), 294–309. Retrieved from http://inderscience.metapress.com/index/FL4XEQ54FW3T46PY.pdf

Reid, J. (2012). The dynamics of certification. *Radiologic Technology*, *83*(4), 414–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22563616

Retnowati, E., Ayres, P., & Sweller, J. (2010). Worked example effects in individual and group work settings. *Educational Psychology*, *30*(3), 349–367. doi:10.1080/01443411003659960

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review*, *7*(3), 216–237. doi:10.1016/j.edurev.2012.05.003

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–92. doi:10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Neuroscience*, *2*(September), 1–10. doi:10.1038/35090060

Rovai, A. (2004). A constructivist approach to online college learning. *The Internet and Higher Education*, *7*(2), 79–93. doi:10.1016/j.iheduc.2003.10.002

Rovai, A. P., Wighting, M. J., Baker, J. D., & Grooms, L. D. (2009).

Development of an instrument to measure perceived cognitive, affective, and psychomotor learning in traditional and virtual classroom higher education settings. *The Internet and Higher Education*, *12*(1), 7–13. doi:10.1016/j.iheduc.2008.10.002

Royer, J. (1979). Theories of the transfer of learning. *Educational Psychologist*, *14*(1), 53–69. doi:10.1080/00461527909529207

Rozeboom, W. W. (1997). Good science is abductive, not hypothetico-deductive. In S. A. Harlow & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 335–392). Mahwah, Erlbaum.

Rupani, C. M. (2011). EVALUATION OF EXISTING TEACHING LEARNING PROCESS ON BLOOM’S TAXONOMY. *Population (English Edition)*, *1*, 119–128.

Sahasrabudhe, V., & Kanungo, S. (2014). Appropriate media choice for e-learning effectiveness: Role of learning domain and learning style. *Computers and Education*, *76*, 237–249. doi:10.1016/j.compedu.2014.04.006

Salyers, V. (2007). Teaching psychomotor skills to beginning nursing students using a web-enhanced approach: A quasi-experimental study. *International Journal of Nursing Education Scholarship*, *4*(1), 1–12. Retrieved from http://works.bepress.com/dr_vincent_salyers/8/

Samaras, H., Giouvanakis, T., Bousiou, D., & Tarabanis, K. (2006). Towards a new generation of multimedia learning research. *AACE Journal*, *14*(1), 3–30. Retrieved from http://www.editlib.org/index.cfm?fuseaction=Reader.ViewFullText& paper_id=5858

Schaal, S. (1997). Learning From Demonstration. *Processing*, 1040–1046.

Scheiter, K., & van Gog, T. (2009). Using eye tracking in applied research to study and stimulate the processing of information from multi-representational sources. *Applied Cognitive Psychology*, *23*(December 2008), 1209–1214. doi:10.1002/acp

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. doi:10.1016/j.learninstruc.2009.02.011

Schnotz, W. (2002). Towards an Integrated View of Learning From Text and Visual Displays. *Educational Psychology*, *14*(1), 101–120.

Schnotz, W., & Kürschner, C. (2007). A Reconsideration of Cognitive Load Theory. *Educational Psychology Review*, *19*(4), 469–508. doi:10.1007/s10648-007-9053-4

Schüler, A., Scheiter, K., & Gerjets, P. (2013). Is spoken text always better? Investigating the modality and redundancy effect with longer text presentation. *Computers in Human Behavior*, *29*(4), 1590–1601. doi:10.1016/j.chb.2013.01.047

Schüler, A., Scheiter, K., Rummer, R., & Gerjets, P. (2011). Explaining the modality effect in multimedia learning: Is it due to a lack of temporal contiguity with written text and pictures? *Learning and Instruction*, *22*(2), 92–102. doi:10.1016/j.learninstruc.2011.08.001

Scriven, M. (2008). A summative evaluation of RCT methodology: and an alternative approach to causal research. *Journal of MultiDisciplinary Evaluation,* *5*(9), 11–24. Retrieved from http://survey.ate.wmich.edu/jmde/index.php/jmde_1/article/view/160/186

Seddon, G. M. (1978). The Properties of Bloom's Taxonomy of Educational Objectives for the Cognitive Domain. *Review of Educational Research*, *48*(2), 303–323.

Seilstad, B. (2012). Using tailor-made YouTube videos as a preteaching strategy for English language learners in Morocco: Towards a hybrid language learning course. *Teaching English with Technology*, *12*(4), 31–49. Retrieved from http://www.ceeol.com/aspx/getdocument.aspx?logid=5&id=6db4a4a0db514a8d87cdaafaf21d03bf

Seitz, C., Orsini, M., & Gringle, M. (2011). YouTube: An international platform for sharing methods of cheating. *International Journal for …*, *7*(1). Retrieved from http://www.ojs.unisa.edu.au/index.php/IJEI/article/view/744

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and Quasi-Experimental for Generalized Designs Causal Inference. *Handbook of Industrial and Organizational Psychology*, *223*, 623. doi:10.1198/jasa.2005.s22

Shelbourne, D. K. (1990). Accelerated rehabilitation after anterior cruciate ligament reconstruction. *The American Journal of Sports Medicine*, *18*(3), 292–299.

Shen, L., Wang, M., & Shen, R. (2009). Affective e-Learning: Using "Emotional" Data to Improve Learning in Pervasive Learning Environment Related Work and the Pervasive e-Learning Platform. *Educational Technology & Society*, *12*(2), 176–189.

Sherman, L. W. (2007). The power few: Experimental criminology and the reduction of harm : t McCord Prize Lecture. *Journal of Experimental Criminology*, *3*(4), 299–321. doi:10.1007/s11292-007-9044-y

Simpson, E. (1972). *The classification of educational objectives in the psychomotor domain: The psychomotor domain.* (3rd Volume.). Washington: Gryphon House.

Simpson, E. J. (1966). *The Classification of Educational Objectives, Psychomotor Domain*.

Skinner, B. (1974). *About behaviorism*. New York: Knopf.

Smith, A. R., Cavanaugh, C., & Moore, W. A. (2011). Instructional multimedia: an investigation of student and instructor attitudes and student study behavior. *BMC Medical Education*, *11*(1), 2–38. doi:10.1186/1472-6920-11-38

Smith, A. R., Jones, J., Cavanaugh, C., Venn, J., & Wilson, W. (2006). Effect of interactive multimedia on basic clinical psychomotor skill performance by physical therapist students. *Journal of Physical Therapy Education*, *20*(2), 61–67. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=2009376714&site=ehost-live

Smith, G. (1998). Return to fitness. In *Orthopaedic Physiotherapy* (Vol. 270, pp. 255– 270). Mosby. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20048467

Snelson, C., & Elison-Bowers, P. (2009). Using YouTube videos to engage the affective domain in e-learning. *Research, Reflections and Innovations in Integrating ICT in Education*, *3*, 1481–1485. Retrieved from http://www.formatex.org/micte2009/book/1481-1485.pdf

Sorden, S. D. (2005). A Cognitive Approach to Instructional Design for Multimedia Learning. *Informing Science Journal*, *8*.

Spanjers, I. a E., Wouters, P., Van Gog, T., & Van Merriënboer, J. J. G. (2011). An expertise reversal effect of segmentation in learning from animated worked-out examples. *Computers in Human Behavior*, *27*(1), 46–52. doi:10.1016/j.chb.2010.05.011

Stefanidis, D., Korndorffer, J. R., Markley, S., Sierra, R., & Scott, D. J. (2006). Proficiency maintenance: Impact of ongoing simulator training on laparoscopic skill retention. *Journal of the American College of Surgeons*, *202*(4), 599–603. doi:10.1016/j.jamcollsurg.2005.12.018

Sullivan, G., & Feinn, R. (2012). Using effect size-or why the P value is not enough. *Journal of Graduate Medical Education*, (September), 279–282. Retrieved from http://www.jgme.org/doi/full/10.4300/JGME-D-12-00156.1

Swarts, J. (2012). New Modes of Help : Best Practices for Instructional Video, *59*(3), 195–206.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. doi:10.1207/s15516709cog1202_4

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design✫. *Learning and Instruction*, *4*(4), 295–312. doi:10.1016/0959-4752(94)90003-5

Sweller, J. (2003). Evolution of human cognitive architecture. *The Psychology of Learning and Motivation*, *43*, 215–266. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0079742103010156

Sweller, J., & Chandler, P. (1991). Evidence for Cognitive Load Theory Evidence for Cognitive Load Theory, *8*(4), 351–362. doi:10.1207/s1532690xci0804

Sweller, J., Merrienboer, J. Van, & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology*, *10*(3), 251–296. Retrieved from http://www.springerlink.com/index/vj4917q523256673.pdf

Tallent-Runnels, M. K., Lan, W. Y., Fryer, W., Thomas, J. a., Cooper, S., & Wang, K. (2005). The relationship between problems with technology and graduate students' evaluations of online teaching. *The Internet and Higher Education*, *8*(2), 167–174. doi:10.1016/j.iheduc.2005.03.005

Thoirs, K., & Coffee, J. (2012). Developing the clinical psychomotor skills of musculoskeletal sonography using a multimedia DVD: A pilot study. *Australasian Journal of Educational Technology*, *28*(4), 703–718.

Thomas, K. (2005). *Learning Taxonomies in the Cognitive, Affective and Psychomotor Domain*. *Rocky Mountain Alchemy*.

Thorndike, E. L. (1913). *Educational psychology: The psychology of learning.* New York: Teachers College Press.

Triola, M. M., Huwendiek, S., Levinson, A. J., & Cook, D. a. (2012). New directions in e-learning research in health professions education: Report of two symposia. *Medical Teacher*, *34*(1), e15–e20. doi:10.3109/0142159X.2012.638010

Turella, L., Pierno, A. C., Tubaldi, F., & Castiello, U. (2009). Mirror neurons in humans: consisting or confounding evidence? *Brain and Language*, *108*(1), 10–21. doi:10.1016/j.bandl.2007.11.002

Urbach, F., Singer, R. N., Simpson, E., Greer, G., & Fleishman, E. A. (1972). *The Psychomotor Domain*.

van Bruggen, J. (2005). *Theory and practice of online learning*. *British Journal of Educational Technology* (Vol. 36). doi:10.1111/j.1467-8535.2005.00445_1.x

Van Bruwaene, S., De Win, G., & Miserez, M. (2009). How much do we need experts during laparoscopic suturing training? *Surgical Endoscopy and Other Interventional Techniques*, *23*(12), 2755–2761. doi:10.1007/s00464-009-0498-z

van Gog, T., Paas, F., & van Merriënboer, J. J. G. (2006). Effects of Process-Oriented Worked Examples on Troubleshooting Transfer Performance. *Learning and Instruction*, *16*, 154–164.

Van Gog, T., Verveer, I., & Verveer, L. (2014). Learning from video modeling examples: Effects of seeing the human model's face. *Computers and Education*, *72*, 323–327. doi:10.1016/j.compedu.2013.12.004

Van Gog, T., Paas, F., Marcus, N.,, Ayres, P, Sweller, J. (2009). The Mirror-Neuron System and Observational Learning. *Educational Psychology Review*, *21*, 21–30.

van Merriënboer, J. J. G., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology*, *20*(3), 343–352. doi:10.1002/acp.1250

Van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: Design principles and strategies. *Medical Education*, *44*(1), 85–93. doi:10.1111/j.1365-2923.2009.03498.x

Veneri, D. (2011). The role and effectiveness of computer-assisted learning in physical therapy education: a systematic review. *Physiotherapy Theory and Practice*, *27*(4), 287–298. doi:10.3109/09593985.2010.493192

Verhoeven, L., Schnotz, W., & Paas, F. (2009). Cognitive load in interactive knowledge construction. *Learning and Instruction*, *19*(5), 369–375.

doi:10.1016/j.learninstruc.2009.02.002

Vogel-Walcutt, J. J., Gebrim, J. B., Bowers, C., Carper, T. M., & Nicholson, D. (2011). Cognitive load theory vs. constructivist approaches: which best leads to efficient, deep learning? *Journal of Computer Assisted Learning*, *27*(2), 133–145. doi:10.1111/j.1365-2729.2010.00381.x

Vygotsky, L. S. (1962). *Thought and language*. Cambridge MA: MIT Press.

Wang, Y., & Crooks, S. M. (2015). Does Combining the Embodiment and Personalization Principles of Multimedia Learning Affect Learning the Culture of a Foreign Language? *Journal of Educational Multimedia and Hypermedia*, *24*(2), 161–177.

Weisburd, D., Petrosino, A., & Mason, G. (1993). Design Sensitivity in Criminal Justice Experiments. *Crime and Justice*, *17*, 337–379.

Wellek, S., & Blettner, M. (2012). On the proper use of the crossover design in clinical trials: part 18 of a series on evaluation of scientific publications. *Deutsches Ärzteblatt International*, *109*(15), 276–81. doi:10.3238/arztebl.2012.0276

Welsh, A. H., & Knight, E. J. (2014). *“Magnitude-Based Inference”: A Statistical Review. Medicine and science in sports and exercise*. doi:10.1249/MSS.0000000000000451

Welsh, B. C., & Farrington, D. P. (2012). The Importance of Randomized Experiments. In *The OXford Handbook of Crime Prevention* (pp. 446–465). Oxford University Press.

White, I. (2010). A simple, low-cost stereographic video capture and viewing solution for teaching psychomotor skills using online delivery. *British Journal of Educational Technology*, *41*(3), 420–431. doi:10.1111/j.1467-8535.2009.00958.x

Willis, R. E., Richa, J., Oppeltz, R., Nguyen, P., Wagner, K., Van Sickle, K. R., & Dent, D. L. (2012). Comparing three pedagogical approaches to psychomotor skills acquisition. *American Journal of Surgery*, *203*(1), 8–13. doi:10.1016/j.amjsurg.2011.07.002

Wittrock, M. C. (1989). Generative Processes of Comprehension. *Educational Psychologist*, *24*(4), 345–376. doi:10.1207/s15326985ep2404_2

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, *22*(6), 449–457. doi:10.1016/j.learninstruc.2012.05.004

Wouters, P., Paas, F., & van Merrienboer, J. J. G. (2008). How to Optimize Learning From Animated Models: A Review of Guidelines Based on Cognitive Load. *Review of Educational Research*, *78*(3), 645–675. doi:10.3102/0034654308320320

Wouters, P., Paas, F., & van Merriënboer, J. J. G. (2008). Observational learning from animated models: Effects of modality and reflection on transfer. *Contemporary Educational Psychology*, *34*, 1–8.

Wouters, P., Tabbers, H. K., & Paas, F. (2007). Interactivity in Video-based Models. *Educational Psychology Review*, *19*(3), 327–342.

doi:10.1007/s10648-007-9045-4

Xeroulis, G. J., Park, J., Moulton, C. A., Reznick, R. K., LeBlanc, V., & Dubrowski, A. (2007). Teaching suturing and knot-tying skills to medical students: A randomized controlled study comparing computer-based video instruction and (concurrent and summary) expert feedback. *Surgery*, *141*(4), 442–449. doi:10.1016/j.surg.2006.09.012

Young, J. Q., Van Merrienboer, J., Durning, S., & Ten Cate, O. (2014). Cognitive Load Theory: implications for medical education: AMEE Guid No. 86. *Medical Teacher*, *36*(5), 371–84. doi:10.3109/0142159X.2014.889290

Zakzanis, K. K. (2001). Statistics to tell the truth, the whole truth, and nothing but the truth: formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Archives of Clinical Neuropsychology : The Official Journal of the National Academy of Neuropsychologists*, *16*(7), 653–67. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14589784

Zhang, D., Zhou, L., Briggs, R., & Nunamakerjr, J. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management*, *43*(1), 15–27. doi:10.1016/j.im.2005.01.004

Appendices

Appendix 1: Formative Feedback Sheet

Rehabilitation and Remedial Therapy

Formative Feedback Sheet

Group names:

Demonstrates enthusiasm, motivational skills and professional demeanour, giving clear and effective instructions to athlete/group

0 ______________________________ 10

Safe and controlled environment

0 ______________________________ 10

Appropriateness of exercises for the stage of rehabilitation and specific joint concerned

0 ______________________________ 10

Effective timing and implementation of program

0 ______________________________ 10

Demonstrating appropriate exercise progression

0 ______________________________ 10

Seminar time: Seminar date:
Lecturer:

Appendix 2: Simpson's Classification of Educational Objectives, Psychomotor Domain: A Tentative System

CHAPTER IV. CLASSIFICATION OF EDUCATIONAL OBJECTIVES, PSYCHOMOTOR DOMAIN: A TENTATIVE SYSTEM

The following schema for classification of educational objectives in the psychomotor domain is presented with the full knowledge that it is still in a very tentative form. Even if the investigator felt quite confident about the system she would, at best, have to admit that it could not be sufficiently tried, in any of its versions, during the one-year funding period.

The major organizational principle operative is that of complexity, with attention to the sequence involved in the performance of a motor act.

1.0 <u>Perception</u> – This is an essential first step in performing a motor act. It is the process of becoming aware of objects, qualities, or relations by way of the sense organs. It is the central portion of the situation – interpretation – action chain leading to purposeful motor activity.

The category of perception has been divided into three subcategories indicating three different levels with respect to the perception process. It seems to the investigator that this level is a parallel of the first category, receiving or attending, in the affective domain.

1.1 <u>Sensory stimulation</u> – Impingement of a stimulus (i) upon one or more of the sense organs.

1.11 <u>Auditory</u> – Hearing or the sense or organs of hearing

1.12 <u>Visual</u> – Concerned with the mental pictures or images obtained through the eyes

1.13 <u>Tactile</u> – Pertaining to the sense of touch

1.14 <u>Taste</u> – Ascertain the relish or flavor of by taking a portion into the mouth

1.15 <u>Smell</u> – To perceive by excitation of the olfactory nerves

1.? Kinesthetic - The muscle sense; pertaining to sensitivity from activation of receptors in muscles, tendons, and joints

The preceding categories are not presented in any special order of importance, although, in Western cultures, the visual cues are said to have dominance, whereas in some cultures, the auditory and tactile cues may pre-empt the high position we give the visual. Probably no sensible ordering of these is possible. It should also be pointed out that "the cues that guide action may change for a particular motor activity as learning progresses (e.g., kinesthetic cues replacing visual cues)" [51].

1.1 Sensory stimulation - Illustrative educational objectives.

Sensitivity to auditory cues in playing a musical instrument as a member of a group.

Awareness of difference in "hand" of various fabrics.

Sensitivity to flavors in seasoning food.

1.2 Cue selection - Deciding to what cues one must respond in order to satisfy the particular requirements of task performance.

This involves identification of the cue or cues and associating them with the task to be performed. It may involve grouping of cues in terms of past experience and knowledge. Cues relevant to the situation are selected as a guide to action; irrelevant cues are ignored or discarded.

1.2 Cue selection - Illustrative educational objectives.

Recognition of operating difficulties with machinery through the sound of the machine in operation.

Sensing where the needle should be set in beginning machine stitching.

Recognizing factors to take into account in batting in a softball game.

1.3 Translation - Relating of perception to action in performing a motor act. This is the mental process of determining the meaning of the cues received for action. It involves symbolic translation, that is, having an image or being reminded of something, "having an idea," as a result of cues received. It may involve insight which is essential in solving a problem through perceiving the relationships essential to solution. Sensory translation is an aspect of this level. It involves "feedback," that is, knowledge of the effects of the process; translation is a continuous part of the motor act being performed.

1.3 Translation - Illustrative educational objectives.

Ability to relate music to dance form.

Ability to follow a recipe in preparing food.

Knowledge of the "feel" of operating a sewing machine successfully and use of this knowledge as a guide in stitching.

2.0 Set - Set is a preparatory adjustment or readiness for a particular kind of action or experience.

Three aspects of set have been identified: mental, physical, and emotional.

2.1 Mental set - Readiness, in the mental sense, to perform a certain motor act. This involves, as prerequisite, the level of perception and its subcategories which have already been identified. Discrimination, that is, using judgment in making distinctions is an aspect.

2.1 Mental set - Illustrative educational objectives.

Knowledge of steps in setting the table.

Knowledge of tools appropriate to performance of various sewing operations.

2.2 Physical set - Readiness in the sense of having made the anatomical adjustments necessary for a motor act to be performed. Readiness, in the physical sense, involves receptor set, that is, sensory attending, or focusing the attention of the needed sensory organs and postural set, or positioning of the body.

2.2 Physical set - Illustrative educational objectives.

Achievement of bodily stance preparatory to bowling.

Positioning of hands preparatory to typing.

2.3 Emotional set - Readiness in terms of attitudes favorable to the motor act's taking place. Willingness to respond is implied.

2.3 Emotional set - Illustrative educational objectives.

Disposition to perform sewing machine operation to best of ability.

Desire to operate a production drill press with skill.

3.0 Guided response - This is an early step in the development of skill. Emphasis here is upon the abilities which are components of the more complex skill. Guided response is the overt behavioral act of an individual under the guidance of the instructor. Prerequisite to performance of the act are readiness to respond, in terms of set to produce the overt behavioral act and selection of the appropriate response. Selection of response may be defined as deciding what response must be made in order to satisfy the particular requirements of task performance. There appear to be two major subcategories, imitation and trial and error.

3.1 Imitation - Imitation is the execution of an act as a direct response to the perception of another person performing the act.

3.1. Imitation - Illustrative educational objectives.

Imitation of the process of stay-stitching the curved neck edge of a bodice.

Performing a dance step as demonstrated.

Debeaking a chick in the manner demonstrated.

3.2 Trial and error - Trying various responses, usually with some rationale for each response, until an appropriate response is achieved. The appropriate response is one which meets the requirements of task performance, that is, "gets the job done" or does it more efficiently. This level may be defined as multiple-response learning in which the proper response is selected out of varied behavior, possibly through the influence of reward and punishment.

3.2 Trial and error - Illustrative educational objectives.

Discovering the most efficient method of ironing a blouse through trial of various procedures.

Ascertaining the sequence for cleaning a room through trial of several patterns.

4.0 **Mechanism** - Learned response has become habitual. At this level, the learner has achieved a certain confidence and degree of skill in the performance of the act. The act is a part of his repertoire of possible responses to stimuli and the demands of situations where the response is an appropriate one. The response may be more complex than at the preceding level; it may involve some patterning of response in carrying out the task. That is, abilities are combined in action of a skill nature.

4.0 Mechanism - Illustrative educational objectives.

Ability to perform a hand-hemming operation.

Ability to mix ingredients for a butter cake.

Ability to pollinate an oat flower.

5.0 **Complex overt response** - At this level, the individual can perform a motor act that is considered complex because of the movement pattern required. At this level, a high degree of skill has been attained. The act can be carried out smoothly and efficiently, that is, with minimum expenditure of time and energy. There are two subcategories: resolution of uncertainty and automatic performance.

5.1 Resolution of uncertainty - The act is performed without hesitation of the individual to get a mental picture of task sequence. That is, he knows the sequence required and so proceeds with confidence. The act is here defined as complex in nature.

5.1 Resolution of uncertainty - Illustrative educational objectives.

Skill in operating a milling machine.

Skill in setting up and operating a production band saw.

Skill in laying a pattern on fabric and cutting out a garmet.

5.2 Automatic performance - At this level, the individual can perform a finely coordinated motor skill with a great deal of ease and muscle control.

5.2 Automatic performance - Illustrative educational objectives.

Skill in performing basic steps of national folk dances.

Skill in tailoring a suit.

Skill in performing on the violin.

Appendix 3: Carryover calculations from Wellek et al., (2012)

BOX 1

Steps in confirmatory statistical analysis of a crossover trial ([1], sect. 2.3; [10], sect. 4.1)

Symbols:

- X_{1i}, X_{2i} = result for patient i of sequence group A–B in period 1 or 2 respectively
- Y_{1j}, Y_{2j} = result for patient j of sequence group B–A in period 1 or 2 respectively
- $C_i(X) = X_{1i} + X_{2i}$, $C_j(Y) = Y_{1j} + Y_{2j}$ [within-subject sums of the results from both periods]
- $D_i(X) = X_{1i} - X_{2i}$, $D_j(Y) = Y_{1j} - Y_{2j}$ [within-subject differences of the results from period 1 and period 2]
- m, n = number of patients in sequence group A–B and sequence group B–A respectively
- N = m + n [total number of patients]

Note: for the example in *Box 3* we have:

m = 7, n = 6;
$X_{11} = 310$, $X_{21} = 270$, $C_1(X) = 310{+}270 = 580$, $D_1(X) = 310{-}270 = 40$;
$Y_{11} = 370$, $Y_{21} = 385$, $C_1(Y) = 370{+}385 = 755$, $D_1(Y) = 370{-}385 = -15$;
And so on for the remaining patients.

1. Pre-test to check the assumption of negligible carryover effects

Has to be performed like an "ordinary" unpaired t-test (see [13]) with $C_1(X), \ldots, C_m(X)$ and $C_1(Y), \ldots, C_n(Y)$ as the two samples. The test statistic thus has to be calculated according to the following equation:

$$T = \sqrt{\frac{mn}{N}}\,\frac{\bar{C}(X) - \bar{C}(Y)}{\sqrt{(SQ_{CX} + SQ_{CY})/(N-2)}},$$

with $\bar{C}(X) = (C_1(X) + \ldots + C_m(X))/m$,
$SQ_{CX} = (C_1(X) - \bar{C}(X))^2 + \ldots + (C_m(X) - \bar{C}(X))^2$
And analogously for $\bar{C}(Y)$, SQ_{CY}.

The (two-sided) p value (see [14]) is then determined as always in the unpaired t-test, namely as the probability that the absolute value of a (centrally) t-distributed parameter with N-2 degrees of freedom exceeds the calculated absolute value of the test statistic T.

2. Test for differences between treatment effects

Formally, this test is carried out according to exactly the same scheme as the pre-test.
The crucial difference is that the customary formula for the unpaired t-test are now applied to the within-subject differences $D_1(X), \ldots, D_m(X)$ and $D_1(Y), \ldots, D_n(Y)$.

British Journal of Educational Technology *Vol 46 No 4 2015* 768–779
doi:10.1111/bjet.12166

The effectiveness of online instructional videos in the acquisition and demonstration of cognitive, affective and psychomotor rehabilitation skills

Darren Cooper and Steve Higgins

Darren Cooper is a course leader in Sports Therapy at the Institute of Sport & Exercise Science, University of Worcester. He is a practicing Graduate Sports Therapist with elite athletes from all levels. His research interests include the cognitive, affective and psychomotor education of students and athletes. Steve Higgins is the director of Research at the School of Education, Durham University. His research interests include the areas of effective use of digital technologies for learning in schools and how teachers can be supported in developing the quality of teaching and learning in their classrooms. Address for correspondence: Mr Darren Cooper, Institute of Sport & Exercise Science, The University of Worcester, Henwick Grove, Worcester WR2 6AJ, UK. Email: D.Cooper@worc.ac.uk

Abstract

The use of instructional videos to teach clinical skills is an ever growing area of e-learning based upon observational learning that is cited as one of the most basic yet powerful learning strategies. The aim of the current study is to evaluate the effectiveness of online instructional videos for the acquisition and demonstration of cognitive, affective and psychomotor skills among undergraduate students, throughout formative assessments with two different durations of instructional videos. The research suggests that the use of videos to support traditional learning should be encouraged. While a conclusive evidence—base for their usage has not yet been established they are a medium which is likely to benefit a proportion of a cohort, and it is very unlikely that they will be harmful to students' learning.

Introduction

The adoption of e-learning technologies for the education of health professionals has increased progressively throughout the past five decades (Triola, Huwendiek, Levinson & Cook, 2012). Principally due to the literature regarding e-learning citing the method as a promising alternative to the traditional classroom setting, it is generally concluded that e-learning will improve learning and is equivalent to or in some cases is superior to traditional teaching methods (Chen, 2011; McNulty, Sonntag & Sinacore, 2009a; Zhang, Zhou, Briggs & Nunamakerjr, 2006).

The adoption of e-learning within a health setting presents a unique set of problems as the complexity of teaching a hands-on profession in a traditional academic setting, rather than a purely theoretical or philosophical subject have been documented in previous literature (see Perry, 1999), and it would be a fair assumption that these challenges carry over into the realm of e-learning. As a result of the challenges posed, educators have responded by adopting various e-learning methods in an attempt to standardise and enhance the efficiency of teaching (Veneri, 2011). Among these methods video has been highlighted as one of the most consistently used methods; however, video is a medium that has been used in academic health for a long time and is well established as an accepted technique (Schittek Janda *et al*, 2005). Newer technologies currently available have enabled the digital distribution of videos that can illustrate abstract cognitive processes or concepts, and this is responsible for bringing video back to the forefront of learning and teaching within medical education (Chen, 2011; McNulty *et al*, 2009b; Wouters,

Practitioner Notes

What is already known about this topic

- e-Learning is increasingly popular within educational domains.
- Observational learning is one of the most powerful learning strategies for humans that can be enabled further by technology.
- As technology becomes ubiquitous, there are less barriers for the delivery of e-learning.

What this paper adds

- Instructional videos can be used to successfully aid traditional learning for cognitive, affective and psychomotor skill acquisition.
- There is a trivial difference between two different durations of instructional videos, both having positive effects overall.
- Formative assessment throughout the experimental phase each year significantly correlated with the student's end of year summative assessment.

Implications for practice and/or policy

- The use of instructional videos is probably beneficial and very unlikely to be harmful to students.
- Instructional videos should be designed with a working knowledge of CLT and MMLT wherever possible.
- Effect size statistics reveal substantially more about a finding than the *p* value alone. Although the effect may not be significant, the increases observed in this study have real-world implications on student attainment.

Tabbers & Paas, 2007). Although videos are still being used within their typical settings to present theoretical content, there is an emerging area pertaining to the use of instructional videos to teach psychomotor skills (McNulty *et al*, 2009b; Smith, Cavanaugh & Moore, 2011).

The traditional strategies used to teach psychomotor skills include the use of lectures, textbooks, self-instruction and live demonstration (Smith *et al*, 2011). Learning motor skills by watching a live demonstration has long been recognised as a successful and well-researched instructional method for over 30 years (Wouters *et al*, 2007). Previous instructional videos of demonstrations were either broadcast "off-line" through TV or stored on CD-ROMS and were not generally made available to students and included only when the teacher felt it was appropriate for the course (Green, 2003; Schittek Janda *et al*, 2005; Zhang *et al*, 2006). However, even though the technology now makes wider distribution possible, introducing videos through e-learning does not necessarily ensure effective and efficient learning (Chen, 2011; Kala, Isaramalai & Pohthong, 2010). More significantly there is limited conclusive evidence demonstrating its effectiveness for teaching clinical skills (Bloomfield, Roberts & While, 2010). The acquisition of psychomotor skills among students and the relationship to teaching methodologies is a relatively new area of research in healthcare education (Veneri, 2011), and in general multimedia educational research has been characterised by inconsistency that have led to inconclusive findings with relation to the effect that multimedia resources can have on the learning process (Samaras, Giouvanakis, Bousiou & Tarabanis, 2006). Tallent-Runnels *et al* (2005) expand on this further stating that there is a serious mismatch between the technological features that are emerging and the pedagogical teaching principles relating to e-learning. Pedagogical principles provide a theoretical structure that guides strategies and activities that in turn form the foundation of good educational practice; within e-learning these are principally focused and implemented by the instructional design (Alonso, Lopez, Manrique & Vines, 2008; Alonso, López, Manrique & Viñes, 2005; Kala *et al*, 2010). Instructional design has developed in combination with three basic learning theories: behaviourism, cognitivism and constructivism (Alonso *et al*, 2005, 2008). Each of the three theories focuses on different perspectives of learning, behaviourist theory centres around observable behaviour (objectivity), cognitivism on unobservable behaviour (subjectivity) and constructivism identifies the importance of new knowledge construction and learner-centred experiences (Kala *et al*, 2010). Although it is generally acknowledged that the cognitive and constructivist approaches have dominated the evolution of learning theories, behaviourism is still perhaps the most pertinent to the fundamental acquisition of psychomotor skills.

Observational learning

Learning by observing and imitating others has long been recognised as one of the most basic yet powerful learning strategies for humans (see Gog, Paas, Marcus, Ayres & Sweller, 2008). Skill acquisition takes place when the behaviour, strategies or thoughts of observers are moulded after viewing experts who perform flawless and error-free movements (Wouters *et al*, 2007).

A significant challenge within multimedia learning concerns the notion as to whether it is possible to promote constructivist learning from passive media (Mayer, 2002). Mayer, Moreno, Boire and Vagge's (1999) Multimedia Learning Theory (MMLT) provides the most definitive and well-evidenced instructional design approach (see Moreno & Mayer, 1999) and is founded on all three of the main pedagogical theories and the extensively tested and substantiated Cognitive Load Theory (Paas, Gog & Sweller, 2010). Therefore, to maximise the instructional effectiveness of multimedia, it is useful to be guided by the research-based theory of how people learn and to connect applications with evidence and theory (Mayer, 2002).

Psychomotor classification

A dominant theoretical classification of learning objectives that incorporates a psychomotor element that is ecologically valid given it is derived to provide taxonomy for the analysis of university-level education and has construct validity given its foundations in behaviourism is Bloom's taxonomy (Ferris & Aziz, 2005). Bloom's taxonomy divided educational objectives into three domains (cognitive, affective and psychomotor) each with an internal hierarchy.

The psychomotor domain was omitted from the original publication and the team led by Bloom never published a hierarchical taxonomy (Ferris & Aziz, 2005) and although it is acknowledged that no single theoretical foundation exists for instructional design, a blending of behaviourist and cognitive approaches seems inevitable (Deubel, 2003). Many authors have sought to extend and complete Bloom's taxonomy so that it fully encompasses a psychomotor domain; however, there has not been an agreed consensus of opinion regarding the classification of this domain. Simpson (1967) attributes this to the involvement of the other two domains within the psychomotor domain that create a "special problem of complexity" when developing a classification system for it (Dave, 1970). Dave proposed a taxonomy that has five levels and bases its foundations clearly upon behaviourist theory, with the first level being *imitation* and the ascending pathways leading to *naturalisation* of the activity and skill.

There has been a dearth of proposed psychomotor classifications since; however, one exception to this is a classification by Dawson (1998), who sought to extend Bloom's taxonomy. Dawson's taxonomy is similar in nature to that of Dave's (1970) and provides a hierarchy from observation through to mastery. Dawson (1998) aligns the pertinence of the psychomotor domain to sport, craft and trade training as well as the objectives of the performing arts and those professions that require manual dexterity such as laboratory practice and surgery. The domain builds on the behaviourist approach as there is a clear pathway for the development of higher order learning given the mastery component of the hierarchy; however, this is purely theoretical as the paper by Deubel (2003) stated that it is still yet to be proven. As a result, despite the extensive range of available taxonomies to draw from the work of Dave (1970) and Dawson (1998) are the most appropriate classifications of the psychomotor domain when compared with the structure, purpose and intended recipients of Bloom's taxonomies.

Instructional video

The use of instructional video is growing considerably and while there are a plethora of papers (ranging from the safe use of wrenches to assembly of electrical circuitry) that investigate the role of an instructional video with a practical task, only a limited proportion of research articles quantify the psychomotor component. From the range of studies conducted a selection relate to the topic of investigation within this proposed study. Salyers (2007) conducted a quasi-experimental study with nursing students to determine the effects of traditional and web-enhanced instructional approaches that included instructional videos. Although their sample size was limited to 36 students, the researchers quantified the student's cognitive, psychomotor and satisfactory scores from the two approaches. The study demonstrated that the students in the web-enhanced group achieved significantly higher scores on the final cognitive examination and achieved better scores on the psychomotor skills examination although not significantly so but were not as satisfied with the course as their traditional delivery colleagues. The results of the study yield interesting conclusions that support the previous learning presented with the exception of the affective domain that the authors accredit to the technological learning environment; however, the fact remains that those who participated in the web-enhanced format did perform better (Salyers, 2007). Another study involving nursing students was conducted by Bloomfield *et al* (2010) to assess the efficacy of an instructional hand-washing video when compared with a control group with a traditional instructional format. The study had a far higher sample size than the previous study with 242 students of mixed gender, age, educational background and first language recruited (Bloomfield *et al*, 2010). Due to the nature of the task, the study was brief and the follow-up measures (knowledge and skill performance) were conducted at 2 and 8 weeks to determine the retention of knowledge. There were no significant differences in the cognitive and psychomotor skill performance scores between the two groups at the 2-week follow-up with significant differences present at the 8-week follow-up (Bloomfield *et al*, 2010). Although the findings of the study were not as robust as the study conducted by Salyers (2007), the findings reveal that the instructional video is a viable alternative for instructing the hand-washing skill and further supports MMLT effectively by demonstrating increased retention due to the deeper learning theorised by Mayer (2002). The application of the findings by Bloomfield *et al* (2010) are given further educational scope as Ahn, Cho, Shon, Park and Kang (2011) conducted a study examining the effect of a "reminder video" viewed on their mobile phones on retention of cardiopulmonary resuscitation (CPR) and automated external defibrillation (AED) skills. Seventy-five male students were randomly assigned to the video-reminded or the control group, and those in the experimental group received text messages every 2 weeks encouraging them to view the 180-second long "reminder video." Three months following their initial training, students in the video-reminded group demonstrated significantly more accurate CPR and AED skills than their control counterparts. As an additional measure, the video-reminder group also demonstrated significantly higher self-assessed confidence and willingness to perform CPR skills than those in the control group (Ahn *et al*, 2011). The research presented is consistent with the conclusions of literature cited initially, and while it can be viewed as an appropriate substitute, it is evident that instructional multimedia for cognitive and psychomotor learning within higher education appears to be most effective when it is delivered in addition to traditional classroom instruction (Smith *et al*, 2011). Although the literature presented earlier was based upon the broad range across all e-learning approaches, the acquisition and demonstration of psychomotor skills has a growing evidence base. However, the conclusions of a systematic review conducted by Veneri (2011) supported by the work of medical and allied health educators state that the usage and effectiveness of computer assisted learning (e-learning) particularly with regard to higher order thinking and psychomotor skills requires further research.

The current study

The current study evaluated the efficacy of online instructional videos (of differing durations) for undergraduate students to determine the effect that they have on the acquisition of cognitive, affective and psychomotor skills. The students were assessed formatively and evaluated on a group and individual basis.

Methodology

The study adopted a longitudinal, blinded, crossover design and took place over an 18-week period within two consecutive academic years. The 18-week period was divided into three, 6-week blocks to ensure parity between the three experimental conditions: control, short and long videos. Lecturers who were blinded to the study assessed students formatively each week.

Ethics

Ethical clearance was granted by Durham University School of Education, Ethics Committee and Teesside University Research Ethics Committee, the former as the lead researcher's supervising institution and the latter as the institution where the research would be conducted.

Participants

Ninety-eight students participated in the study across two academic years, 60 females and 38 males with a mean age of 20.87 ($SD = 4.53$). All participants were fully informed about the study and were asked to voluntarily complete a written consent form prior to participation in accordance with British Educational Research Association guidelines (British Educational Research Association, 2004). Each participant was informed of their right to withdraw and their right to obtain a copy of the results following the completion of the study.

Research design considerations

The crossover design was based upon a Latin square (Table 1) that ensured all students participated within each of the three conditions throughout the module to ensure no group of participants was disadvantaged. Students were requested not to share resources with friends outside their seminar group to minimise the experimental treatment diffusion and threaten the internal validity. Students did not receive scores during the period to prevent compensatory rivalry/equalisation, resentfulness or demoralisation between groups, particularly from those who begin with a resource and experience the control condition subsequently. The lecturers responsible for the academic delivery and formative assessment of the module remained consistent throughout both of the research periods.

Instructional videos

Two groups of instructional videos were created that are both based upon the work of Smith (1998) that forms the foundation of the teaching structure for the rehabilitation module. Smith

Table 1: Crossover design

Phase	*Group A*	*Group B*	*Group C*
1	Control	Long	Short
2	Long	Short	Control
3	Short	Control	Long

(1998) separates the rehabilitation process into three distinct sections within the rehabilitation continuum: early, intermediate and late, and provides examples of exercises that are appropriate for that stage and the criteria that need to be achieved in order to progress through the continuum. Each of the exercises was supported with further literature to provide up-to-date definitions and information to foster an evidence-based approach.

Twenty-five short duration videos ranging from 55 to 118 (1.58 minutes) seconds were created for one group and five previously recorded and produced longer duration videos ranging from 629 (10.29 minutes) seconds to 1094 (18.14 minutes) seconds were used for the second group. The short duration videos were designed in accordance with the principles for instructional design set out by Mayer (2008); the long duration videos were produced prior to these criteria being applied and adhere to all except the redundancy principle, as on-screen text was used with narration. It was not deemed feasible to reshoot and reproduce these five videos. The production value of the videos was not a distinct criterion within the research design; however, it is important to acknowledge that the videos had differing production values due to the software and resources available at the time of production. The short videos benefitted from more recent software and greater expertise throughout their production.

All videos were converted into the same format (.wmv) and with as similar resolutions and streaming rates as possible for consistency. The videos were distributed online with an adaptive release ensuring groups had access to each respective group when the crossover design dictated.

Experimental and control conditions

During the phases of data collection, there was a minimum of three seminar groups within the rehabilitation module (A, B and C). In the second year of data collection, there were four seminar groups, two of these were amalgamated to ensure there were three distinct groups. Two of these groups had access to the different instructional videos, whereas the control condition within the study was not provided with any additional materials other than those that are already available to all students.

Procedure

The annual format of the module after the initial 6 weeks of pitch side management and introduction to rehabilitation theory requires the students to produce rehabilitation sessions in groups, for a specific peripheral joint at a particular stage of the rehabilitation continuum. The students have 1 week to formulate the sessions traditionally relying on standard resources. In the lesson, the students conduct the sessions and gain formative feedback from the tutor and their peers.

Instruments

The students were assessed formatively by each lecturer using a 10 cm Likert scale devised from Smith (1998) with no numerical or written reference points. The five criteria assessed were identical to those used in the end of year practical examination; these encompass cognitive, affective and psychomotor domains. The students received formative feedback verbally at the end of each session; the completion of the scales did not increase the workload of those involved.

Data analysis

Each Likert item was measured with a standard 30 cm ruler to an accuracy within 1 mm; all scales were measured by the same person enhancing reliability. Totals were summed, and averages were calculated in Excel (Microsoft Excel for Mac 2011, Microsoft Corporation, Redmond, WA, USA) for each of the 54 data collection points within each year, 108 in total across the experimental period. In addition to the null hypothesis statistical tests, Cohen's effect size statistics were calculated using a macro-enabled spreadsheet (Hopkins, 2006) and evaluated using the descriptors and magnitude-based inferences provided by Batterham and Hopkins (2006).

Although the scores were gained within groups, they were assigned to each individual (within the group) to provide an individual total for each participant for each experimental condition. The students worked in different groups each week ensuring no individual total would be the same. In essence, each participant had the potential for six separate scores within each experimental phase, yielding an overall total for each phase. As a result, participants were excluded from the study during data analysis if they attended less than 50% of the sessions during each of the experimental conditions of the study; this resulted in 19 students being excluded from the data analysis.

Results

In order to investigate the variance between the three experimental groups a between–between univariate analysis of variance (ANOVA) with Tukey post hoc tests was used to analyse the formative scores of the three groups. The results demonstrate that there was no statistical difference between the three experimental groups ($F_{2,104} = 0.17$, $p = 0.85$) despite the observable difference in means between the control and videos groups in Figure 1.

The mean difference in formative assessment between control and long was 3.23 (95% confidence interval −1.41 to 7.86, $p = 0.25$) and control and short was 2.88 (95% confidence interval −1.72 to 7.48, $p = 0.30$). The magnitude of this effect for each comparison was small (effect size 0.27: −0.20 to 0.73) and (effect size 0.21: −0.19 to 0.60), respectively. The probability (% chances) that the true population effect is beneficial/trivial/harmful is 61.8/35.8/2.4 and 51.1/46.6/2.2, therefore the interventions are possibly beneficial and very unlikely to be harmful. When the two video groups where compared the mean difference in formative assessment between long and short was 0.65 (95% confidence interval −3.29 to 2.59, $p = 0.84$). The magnitude of this effect was insubstantial (effect size 0.05: −0.42 to 0.52). The probability that the true population effect is beneficial/trivial/harmful is 25.6/59.8/14.7, therefore the difference between the two video groups is predominantly trivial.

Individuals

A two-way *factorial* ANOVA was performed to analyse the differences, if any, between the individual students. The findings demonstrate that there is a statistical difference between the

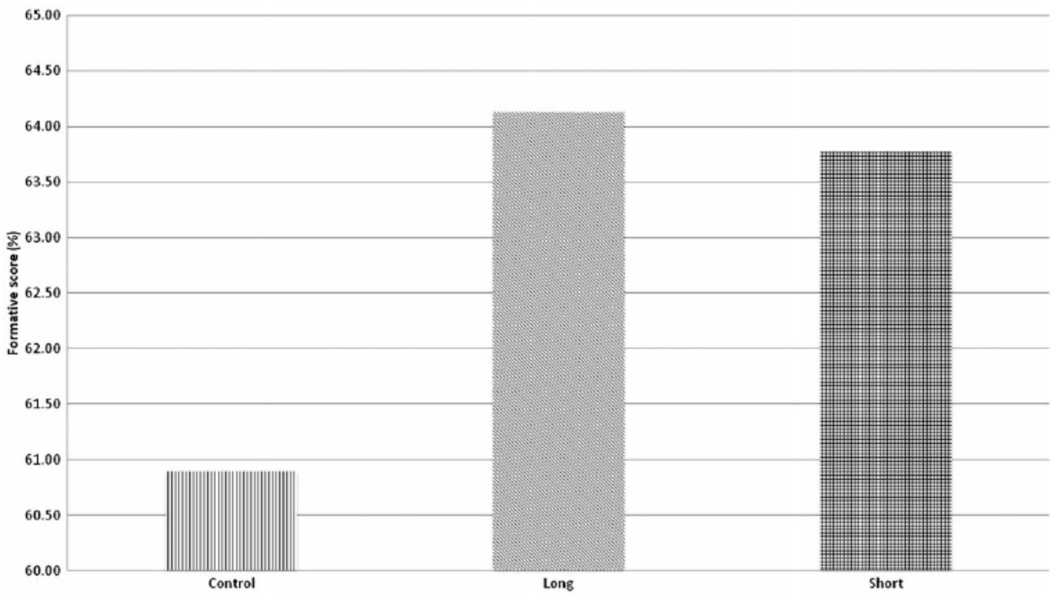

Figure 1: Mean formative scores for each experimental group

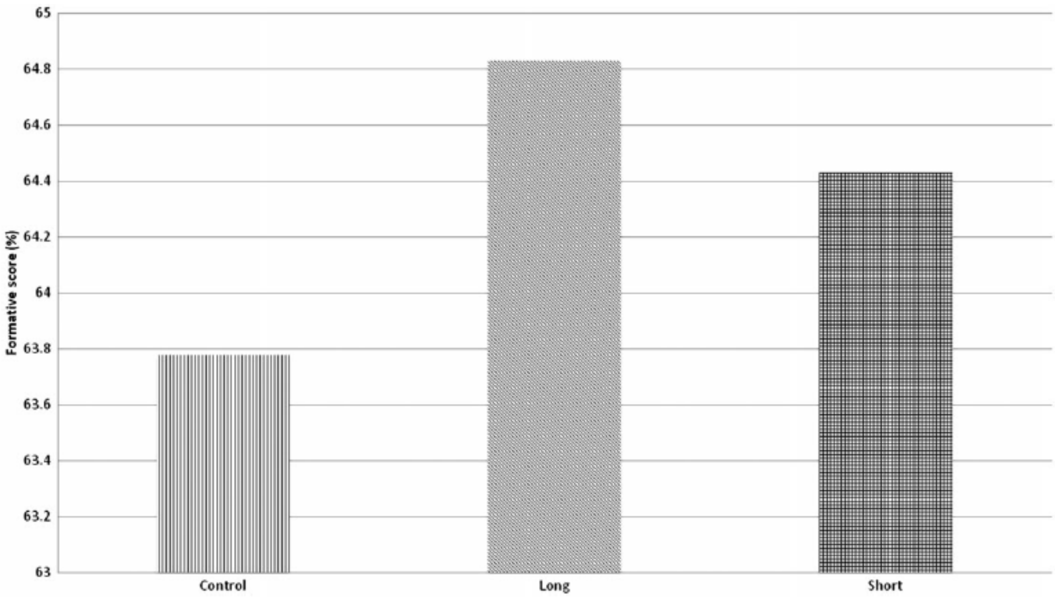

Figure 2: Mean formative scores for individuals within experimental groups

participants ($F_{97,1051} = 3.66$, $p < .01$); it confirms the previous ANOVA by showing that there is no significant difference between the groups ($F_{2,1051} = 1.53$, $p = 0.22$) despite the observable difference in means in Figure 2; however, there is a significant interaction between participants within the groups ($F_{194,1051} = 2.11$, $p < 0.01$). This is further demonstrated when a univariate ANOVA is performed to assess the variance between the participants with separate groups ($F_{97,1247} = 3.44$, $p < 0.01$, partial $\eta^2 = 0.21$).

As with the experimental conditions, individual effect sizes were calculated using the same procedures as above with the addition of Hedges and Olkin's correction for *d* statistic bias that is recommended when there are less than 10 in each group (Hedges & Olkin, 1985).

The mean difference in formative assessment between control and long was 0.34 (95% confidence interval –21.72 to 22.39, $p = 0.40$) and control and short was 0.65 (90% confidence interval –21.02 to 22.32, $p = 0.33$). The magnitude of this effect for each comparison was small (effect size 0.49: –0.39 to 1.37) and (effect size 0.57: –0.28 to 1.42), respectively. The probability (% chances) that the true population effect is beneficial/trivial/harmful is 45.9/50.6/3.5 and 52.1/45.1/2.7, therefore the interventions are possibly beneficial and most notably very unlikely to be harmful.

Accesses

Throughout the duration of the testing periods, the accesses to the control site and amount of times each of the videos was played within each of the video sites was monitored and recorded. The control website was accessed 365 times in total, the five long videos were viewed 453 times in total (1.24 times more than the control) and the 25 short videos were viewed 958 times (2.62 times more than the control and 2.11 times more than the long videos). As the students could view these videos in groups, it is impossible to correlate individual accesses with formative scores.

Academic performance

In conjunction with their individual formative marks for each of the three conditions and their summative mark from the rehabilitation module, summative data from the end of the first-year students' practical examinations was available that enabled a Spearman's Rho to be performed to identify any trends between the data sets. The results of the analysis show that there is a strong positive correlation between the average mark from the first-year students' practical examinations and their final rehabilitation mark (0.53, $p < 0.01$) and between their formative marks obtained during the short and long video conditions and their final rehabilitation mark (0.28, $p < 0.05$ and 0.29, $p < 0.01$, respectively).

Discussion

The differences between experimental conditions did not show any statistically significant results for the group or the individual analysis; however, the use of effect size statistics that enables the researcher to "move towards a more generally interpretable, quantitative description of the size of an effect" (Fritz, Morris & Richler, 2012) demonstrates a small effect size when both the experimental groups were compared with the control group. The real-life relevance of this was an increased mean score over the control group, while this was only a modest increase in the mean score the additional couple of marks for a group or an individual could mean the difference between pass and fail or the attainment of a higher classification boundary. The magnitude-based inferences indicate the most important findings from this study, showing that both of the video groups are almost certainly not harmful to the participants. Although the probabilities for a beneficial or trivial effect are both within the central quartiles, with beneficial being marginally greater than trivial, it is clear that videos are to some degree beneficial when students are working in groups.

A comparison of the two experimental groups demonstrate a substantially reduced mean difference and an insubstantial effect size, suggesting the different delivery durations, any infringement of the redundancy principle and the production value made little difference to the scores achieved by the groups.

When the findings for individuals are interpreted, the two-way factorial ANOVA identifies that the scores achieved by individuals are significantly different from each other demonstrating that individual data can be extrapolated from the experimental group design. Furthermore, there is a significant interaction between the groups, as just over one-fifth of the variance is explained by the combined effect of the video groups.

The effect sizes are classified as small and are approximately twice as large for individuals than groups (with a corresponding increase in confidence intervals). This partly reflects the small scale of the study. The magnitude-based inferences for individuals follow a similar pattern to the groups with the beneficial and trivial probabilities being almost equal and the probability that the intervention will have a harmful effect being exceedingly small. As a result, the confidence intervals illustrate the extent of the effect on the scores of the individuals with a dramatic effect. Within the cohorts, there are clear differences in the effect of videos in a beneficial or harmful way; however, the majority of the time is beneficial with the effect being the equivalent of an entire classification boundary for some students (10% boundaries). In addition, the findings of the Spearman's rho indicate that the marks the students obtained individually when in the short or the long video group is significantly correlated with the marks they obtained on an individual basis in the final summative rehabilitation examination, whereas the marks obtained when they experienced the control group were not. It is not possible to claim that the videos had a direct influence due to the assumption that different levels of preparation undertaken in the run up to formative assessments opposed to summative; however, it demonstrates that they were working at a level that is similar to their academic performance in the examination making the feedback and advice more relevant and beneficial.

As a result of these findings, it is possible to agree with the generalised conclusions from Chen (2011), McNulty *et al* (2009a) and Zhang *et al* (2006) that e-learning can improve traditional teaching methods, with a caveat that it may not be effective for the entire cohort. The magnitude of the improvement can vary considerably across groups and individuals, and while the interventions did not necessarily promote learning efficiency or effectiveness consistently in agreement with Chen (2011) and Kala *et al* (2010), their adoption or integration will probably be of benefit for a large proportion of the cohort. The most noteworthy finding of this study is that the probability of the use of videos being harmful to a group or an individual is exceedingly small and although this does not provide conclusive evidence of the effectiveness for teaching clinical skills that Bloomfield *et al* (2010) are seeking, it should encourage educators to consider adopting the use of videos to teach clinical skills. This is particularly relevant when you consider how important the role of observational learning still is within education to this day (Gog *et al.*, 2008). Videos ensure that the viewing of flawless and error-free movements performed by experts can be repeated and analysed an unlimited number of times embedding and reinforcing the movement with no cognitive variation. The ability to pause and work through a complex movement one frame at a time provides the learner with an incredibly powerful tool, a mechanism that can transform the passive design and promote constructivist learning.

Limitations

The assessment of psychomotor skills is not as objective as it is for other domains. Although the reliability of assessment within this study was acceptable, the subjectivity of this element is an area that would benefit from further research for the students and the examiners to ensure standards and marks achieved/given are as reliable as possible. The scale of this exploratory study was small with three conditions (short video, long video and control) meaning it was not sufficiently powered to detect small effects. A larger scale study could be undertaken to establish the effectiveness of video use in this context.

Conclusion

The role of instructional videos within traditional learning should be further encouraged; although the current study could not establish a conclusive evidence base for their inclusion, the effects observed are potentially substantial for some students given the real-world application of the small effect size on grade boundaries. The study has established that instructional videos are very unlikely to be harmful to a student's learning and may have other benefits that the study has been unable to quantify such as flexibility and independence of learning (any time or place with an internet connection), the ability to be repeated instantly and the consistency of delivery that could aid the acquisition of cognitive, affective and psychomotor elements within each taxonomy further. Future research into the use of instructional videos should be undertaken on a larger scale, perhaps with a less complicated research design to improve power. Such research should aim to include a rigorous method of objectifying each taxonomy and seek to control and monitor the usage of content more precisely.

Acknowledgements

The authors would like to thank Gary Henderson, Jason Hynd, Gok Kandasamy, Helen Llewellyn, Steve Smelt and the participants from Teesside University.

References

Ahn, J. Y., Cho, G. C., Shon, Y. D., Park, S. M. & Kang, K. H. (2011). Effect of a reminder video using a mobile phone on the retention of CPR and AED skills in lay responders. *Resuscitation*, *82*, 12, 1543–1547.

Alonso, F., López, G., Manrique, D. & Viñes, J. M. (2005). An instructional model for web-based e-learning education with a blended learning process approach. *British Journal of Educational Technology*, *36*, 2, 217–235.

Alonso, F., Lopez, G., Manrique, D. & Vines, J. M. (2008). Learning objects, learning objectives and learning design. *Innovations in Education and Teaching International*, *45*, 4, 389–400.

Batterham, A. M. & Hopkins, W. G. (2006). Making meaningful inferences about magnitudes. *International Journal of Sports Physiology and Performance*, *1*, 1, 50–57.

Bloomfield, J., Roberts, J. & While, A. (2010). The effect of computer-assisted learning versus conventional teaching methods on the acquisition and retention of handwashing theory and skills in pre-qualification nursing students: a randomised controlled trial. *International Journal of Nursing Studies*, *47*, 3, 287–294.
British Educational Research Association (2004). Revised ethical guidelines for educational research (2004). Educational Research. British Educational Research Association.
Chen, J.-L. (2011). The effects of education compatibility and technological expectancy on e-learning acceptance. *Computers & Education*, *57*, 2, 1501–1511.
Dave, R. H. (1970). Psychomotor levels. In R. J. Armstrong (Ed.), *Developing and writing behavioural objectives* (pp 33–34). Tucson, AZ: Educational Innovators Press.
Dawson, W. R. (1998). *Extensions to Bloom's taxonomy of educational objectives*. Sydney, Australia: Putney Publishing.
Deubel, P. (2003). An investigation of behaviorist and cognitive approaches to instructional multimedia design. *Journal of Educational Multimedia and Hypermedia*, *12*, 1, 63–90.
Ferris, T. L. J. & Aziz, S. M. (2005). A psychomotor skills extension to Bloom's taxonomy of education objectives for engineering education. In *Exploring innovation in education and research* (pp 1–5). Tainan, Taiwan: iCEER.
Fritz, C. O., Morris, P. E. & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of Experimental Psychology. General*, *141*, 1, 2–18.
Gog, T., Paas, F., Marcus, N., Ayres, P. & Sweller, J. (2008). The mirror neuron system and observational learning: implications for the effectiveness of dynamic visualizations. *Educational Psychology Review*, *21*, 1, 21–30.
Green, S. (2003). Evaluating the use of streaming video to support student learning in a first-year life sciences course for student nurses. *Nurse Education Today*, *23*, 4, 255–261.
Hedges, L. V & Olkin, I. (1985). *Statistical methods for meta-analysis*. London: Academic Press.
Hopkins, W. G. (2006). Spreadsheets for analysis of controlled trials, with adjustment for a subject characteristic. *Sportscience*, *10*, 46–50.
Kala, S., Isaramalai, S.-A. & Pohthong, A. (2010). Electronic learning and constructivism: a model for nursing education. *Nurse Education Today*, *30*, 1, 61–66.
Mayer, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*, 1, 107–119.
Mayer, R. E. (2008). Applying the science of learning: evidence-based principles for the design of multimedia instruction. *The American Psychologist*, *63*, 8, 760–769.
Mayer, R. E., Moreno, R., Boire, M. & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, *91*, 4, 638–643.
Moreno, R. & Mayer, R. (1999). Visual presentations in multimedia learning: conditions that overload visual working memory. *Visual Information and Information Systems*, *1614*, 798–805.
McNulty, J. A., Sonntag, B. & Sinacore, J. M. (2009a). Evaluation of computer-aided instruction in a gross anatomy course: a six-year study. *Anatomical Sciences Education*, *8*, 2, 2–8.
McNulty, J. A., Hoyt, A., Gruener, G., Chandrasekhar, A., Espiritu, B., Price, R. *et al* (2009b). An analysis of lecture video utilization in undergraduate medical education: associations with performance in the courses. *BMC Medical Education*, *9*, 1–6.
Paas, F., Gog, T. & Sweller, J. (2010). Cognitive load theory: new conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review*, *22*, 2, 115–121.
Perry, S. (1999). Strategies for teaching clinical decision-making skills in neurologic content: perspective from the field. *Journal of Neurologic Physical Therapy*, *23*, 4, 170–177.
Salyers, V. (2007). Teaching psychomotor skills to beginning nursing students using a web-enhanced approach: a quasi-experimental study. *International Journal of Nursing Education Scholarship*, *4*, 1, 1–12.
Samaras, H., Giouvanakis, T., Bousiou, D. & Tarabanis, K. (2006). Towards a new generation of multimedia learning research. *AACE Journal*, *14*, 1, 3–30.
Schittek Janda, M., Tani Botticelli, A., Mattheos, D., Nebel, D., Wagner, A., Nattestad, A. *et al* (2005). Computer-mediated instructional video: a randomised controlled trial comparing a sequential and a segmented instructional video in surgical hand wash. *European Journal of Dental Education*, *9*, 53–58.
Simpson, E. (1967). Educational objectives in the psychomotor domain. In M. B. Kapfer (Ed.), *Behavioural objectives in curriculum development* (pp 60–67). New Jersey: Educational Technology Publications. 1971.
Smith, A. R., Cavanaugh, C. & Moore, W. A. (2011). Instructional multimedia: an investigation of student and instructor attitudes and student study behavior. *BMC Medical Education*, *11*, 1, 2–38.
Smith, G. (1998). Return to fitness. In M. E. Tidswell (Ed.), *Orthopaedic physiotherapy* Vol. 270 (pp 255–270). London: Mosby.

Tallent-Runnels, M. K., Lan, W. Y., Fryer, W., Thomas, J. A., Cooper, S. & Wang, K. (2005). The relationship between problems with technology and graduate students' evaluations of online teaching. *The Internet and Higher Education*, *8*, 2, 167–174.
Triola, M. M., Huwendiek, S., Levinson, A. J. & Cook, D. A. (2012). New directions in e-learning research in health professions education: report of two symposia. *Medical Teacher*, *34*, 1, e15–e20.
Veneri, D. (2011). The role and effectiveness of computer-assisted learning in physical therapy education: a systematic review. *Physiotherapy Theory and Practice*, *27*, 4, 287–298.
Wouters, P., Tabbers, H. K. & Paas, F. (2007). Interactivity in video-based models. *Educational Psychology Review*, *19*, 3, 327–342.
Zhang, D., Zhou, L., Briggs, R. & Nunamakerjr, J. (2006). Instructional video in e-learning: assessing the impact of interactive video on learning effectiveness. *Information & Management*, *43*, 1, 15–27.

Online instructional videos as a complimentary method of teaching practical rehabilitation skills for groups and individuals.

Darren Cooper, Steve Higgins and Nadin Beckmann

Darren Cooper is a Course Leader in Sports Therapy at the Institute of Sport & Exercise Science, University of Worcester. He is a practicing Graduate Sports Therapist with elite athletes from all levels; his research interests include the cognitive, affective and psychomotor education of students and athletes.
Address for correspondence: Darren Cooper, Institute of Sport & Exercise Science, The University of Worcester, Henwick Grove, Worcester, WR2 6AJ. Email: *D.Cooper@worc.ac.uk*. Prof. Steve Higgins is a Professor at the School of Education, Durham University. His research interests include the areas of effective use of digital technologies for learning in schools, and how teachers can be supported in developing the quality of teaching and learning in their classrooms. Dr Nadin Beckmann is a Lecturer at the School of Education, Durham University. Her current research projects focus on the dynamic, process-oriented approaches to personality in the context of adult learning and development.

Abstract
Online instructional videos are becoming increasingly commonplace within education for a wide variety of topics. This study adopts a quasi-experimental 2 x 2 crossover design with sixty eight level 5 university students (20.6 ± 2.8 years) to evaluate the efficacy of instructional videos as a method of teaching practical rehabilitation skills. The students were required to devise practical rehabilitation sessions for peripheral joints using traditional resources (control group) and instructional videos in addition to traditional resources (experimental group) to demonstrate their understanding of the rehabilitation process. The students then performed the practical sessions in class and were formatively assessed by their lecturers against specific criteria. The data was collected and analysed for groups and individuals. Effect size and magnitude based inferences are reported. The results of the study demonstrate that the instructional videos had a small to moderate beneficial effect on all criteria that the students were being assessed on.

Keywords: instructional videos; e-learning; psychomotor skills; constructivism.

Introduction
During the last decade the use of instructional videos has gained significant momentum within varying educational contexts. For example videos are being used in music instruction (Kruse & Veblen, 2012), as pre-teaching tools for English language (Seilstad, 2012), as aids for DIY hobbyists who are moving away from the traditional manual to instructional videos (Swarts, 2012) and by students themselves to visualise elements from chemistry manuals (Benedict & Pence, 2012).

Videos present a model which demonstrates how to perform a task and are increasingly used in educational settings as they have become simple to create and disseminate via e-learning environments (Van Gog, Verveer, & Verveer, 2014). A predominant problem associated with multimedia learning is whether it is possible to promote constructivist learning from passive media (Mayer, 2002). A problem that is still relevant today as video is fast becoming the primary medium for instructing users about procedures (van der Meij & van der Meij, 2014), however questions arise as to its design and effectiveness. Multimedia Learning Theory (MMLT) provids a well-evidenced instructional design approach (Moreno & Mayer, 1999) that encompassed the main pedagogical theories and is used as a foundation for the development of instructional design principles that are emerging such as those presented by Meij and Meij (2013). Meij and Meij (2013) state that their guidelines (e.g., ensure you provide easy access, use animation with narration, enable interactivity, preview the task, provide procedural information, use clear and simple tasks, use short videos and strengthen the demonstration with practice) summarise key notions of accepted thinking rather than advancing new theory for instructional video design. Although the instructional videos used in this study were created prior to the publication of Meij and Meij (2013) the guidelines have been adhered to as they are common themes amongst previous sources of guidance (e.g., Mayer, 2008). Therefore, the application and effectiveness of the specific e-learning resource used in the current study is the prime focus of this paper. In doing so we aim to advance previous work by Cooper and Higgins (2014) that suggested that instructional videos can be used successfully to aid the teaching of practical skills.

Study Background

The study was employed in a practical rehabilitation module were students are taught the underpinning fundamentals of the rehabilitation process. Rehabilitation is an integral skill for Sports Therapists that needs to be taught and founded upon an evidence-based approach. Whilst many sources exist that provide criteria for the safe and effective rehabilition of individuals with musculoskeletal injuries the current study has been based on the work of Smith (1998). Rehabilitation is a continuum that can be divided into stages that the patient has to progress through before they can return to sport successfully. The efficacy of rehabilitation is dependant upon the knowledge (cognitive domain) of the therapist, their ability to communicate and empathise appropriately with the ‘patient’ (affective domain) and demonstrate the required exercises with appropriate modifications, regressions or progressions (psychomotor domain) for the patient.

Pedagogical Approach

When assessing an intervention’s efficacy it is prudent to understand the fundamental theories of the approach to learning being employed. The basic learning theories are behaviourism, cognitivism and constructivism. Each have a role within instructional design of e-learning resources. The role of behaviourism within multimedia design has predominantly been associated with lower-level learning tasks such as the rapid acquisition of basic concepts and skills and is yet to be evidenced for higher-order learning tasks (Deubel, 2003). The rationale behind this argument is that behaviourist designs are passive and do not foster cognitive activity that is generally associated with higher-order learning (Samaras, Giouvanakis, Bousiou, & Tarabanis, 2006) and underlines the predominant problem associated with multimedia learning.

Cognitive theory views learning as the acquisition or reorganisation of cognitive structures to process and store information (Alonso, Lopez, Manrique, & Vines, 2008). The social cognitive model of sequential skill acquisition summarised by Wouters, Tabbers, and Paas (2007) describes how learners initially start with observing a model, but then begin practicing and gradually learning how to self-regulate their own performance which is accompanied by different cognitive processes such as problem solving and cognitive behaviour modification.

Constructivism is less well defined and has varying approaches but each of them has a common philosophy; learning is understood, knowledge stored and applied most effectively when learners develop their own mental representations from presented information (Mayer, Moreno, Boire, & Vagge, 1999; Vogel-Walcutt, Gebrim, Bowers, Carper, & Nicholson, 2011). Constructivism builds upon behaviourism and cognitivism as it accepts multiple perspectives and maintains that learning is a personal interpretation of the world (Alonso, López, Manrique, & Viñes, 2005). In line with constructivism at present, educationalists are emphasising the need for a student-centred learning approach and the transition from teacher-centred to student-centred (Kala, Isaramalai, & Pohthong, 2010). Although it has been acknowledged that constructivism is complex and the detail rife with controversy, a constructivist approach in education can be reduced to the concept that learners actively construct their own personal learning experience rather than passively receive it (Kim & Reeves, 2007; Vogel-Walcutt et al., 2011). Although the media is passive as the viewer only observers rather than interacts, constructivism can be fostered when cognitive activity is present within the learner (Mayer, 2002). As a result Mayer (2002) concludes that it is not necessary to have a hands-on activity or social collaboration but in essence a well-designed multimedia explanation embedded within an appropriately interactive learning environment in order to provide conditions that will promote constructivist learning.

Group vs Individual Learning

In addition to the general learning theories another dimension to consider is that of group as compared to individual learning (see Hill, 1982; Laughlin, Hatch, Silver, & Boh, 2006; Lou, Abrami, & d'Apollonia, 2001). This has several implications from a performance perspective as if the adage 'two heads are better than one' (Hill, 1982) is true then a group based activity may not be the most conducive learning task for individuals. However Laughlin et al., (2006) showed that cooperative groups of three to six members performed better than independent individuals on a wide range of problems. Therefore previous research suggests that as the group size increases the performance also improves up to a threshold of six. When students worked in two-person groups there was a positive effect on both group and individual achievement (Laughlin et al., 2006; Lou et al., 2001). Therefore controlling the group size may be of critical importance, however none of the identified studies relate to the teaching of practical skills through online videos and are purely focused on problem solving tasks. Therefore this study aims to contribute to this field of enquiry as the use of instructional videos is steadily increasing.

In sum, the aim of this study is to evaluate the effectiveness of online instructional videos as a complementary method of teaching practical rehabilitation skills to students in groups and individually.

Methodology

Ethics

Ethical clearance was granted by Durham University School of Education, Ethics Committee and Teesside University Research Ethics Committee, the former as the lead researcher's supervising institution and the latter as the institution where the research would be conducted.

Participants

The participants were 68 level 5 students (mean age 20.6 ± 2.8 years, 53% female). Students were enrolled onto the Rehabilitation and Remedial Therapy module at a University in the North East of England (2012-13).

Study Design

A quasi-experimental design was adopted utilising existing seminar groups as control and experimental conditions for the duration of the study. The experimental condition had access to instructional videos for the first half of the experimental period (9 weeks) whilst the control condition only had access to traditional teaching resources. The groups then swapped conditions for a further 9 weeks as it would be unethical for them to not have access to a resource that could potentially improve their knowledge, skills and summative assessment (2 x 2 repeated measures crossover design). The design is similar to that used in Cooper and Higgins (2014) with one fundamental difference; the previous 3x3 crossover design has been replaced with a 2x2 crossover design. The purpose of the modification was to increase the duration of the intervention and thereby increase the sample size for each aspect of the criteria the students were assessed against.

Condition Allocation

The Rehabilitation and Remedial Therapy module in 2012-13 contained four seminar groups. As two distinct groups were required for the study two seminar groups were combined to form the control condition and two were combined to form the experimental condition. This was deemed to be the most feasible option for condition allocation, it also ensured that the entirety of each seminar group was undertaking one condition or the other minimising the potential for treatment diffusion.

Academic Setting

For the majority of the Rehabilitation and Remedial Therapy module, the students were requested to perform practical sessions to demonstrate their understanding of the rehabilitation processes for peripheral joints based upon the work by Smith (1998). Smith (1998) devised a rehabilitation continuum ranging from early to late stage with specific criteria as to when a 'patient' should progress to ensure their safety throughout their rehabilitation. In the practical sessions the students were using peers as 'patients' and conducted the session as though their peer had a particular injury.

Instructional Videos

The current study used a group of instructional videos (twenty-five short duration videos ranging from 55 to 118 seconds) that were designed in accordance with the principles for instructional design set out by Mayer (2008). The short duration videos were selected over the long duration videos as there was only a small difference between the long and short interventions when groups were analysed (see Cooper & Higgins, (2014) for further details). Participants in the experimental group had access to the instructional videos through a virtual learning environment (VLE, Blackboard). The participants who had access to the resource could access it at will throughout the entire duration of their experimental period via any means that supported the VLE.

Procedure

All of the students regardless of group were informed prior to the lesson as to what they would be required to demonstrate in class that week (for example; late rehabilitation for the ankle) enabling them to prepare by accessing either the traditional teaching resources (control group) or the traditional teaching resources and the online instructional videos (experimental group). Each week the students were asked to devise a series of exercises to form a programme up to twenty minutes in duration that would be suitable for the particular stage and joint. Due to the size of the classes the practical sessions were devised and delivered by a small group of students (three to four), multiple sessions (three to four) were delivered within a specific practical lesson.

The students were formatively assessed throughout each of their individual practical sessions by their lecturers against an instrument based upon their summative assessment (detailed below). There were six peripheral joints and three distinct rehabilitation stages which yielded nine sampling opportunites per condition. As three to four groups performed sessions during each lesson a maximum of 36 aggregate scores could be obtained for each of the two conditions (control and experimental). Although the sessions were delivered in a group format the groups were continually changing therefore it was possible to create an individual score as a sum of the scores they have gained throughout the group sessions they have contributed to.

Students were requested not to share elements of the instructional videos with friends outside their seminar group to minimise treatment diffusion which might have threatened the internal validity of the study. Students did not receive scores during the period to prevent resentfulness or demoralisation between groups, particularly from those who initially had access to the instructional videos and then experience the control condition subsequently. The lecturers responsible for the academic delivery and formative assessment of the module remained consistent throughout both of the research periods.

Instruments

The students were formatively assessed on five criteria (Motivation, Safety, Exercise, Timing and Progression) that each used a 10cm analogue scale (from 0 to 10 with no marked incements). This assessment form was also used for student assessment in Cooper and Higgins, (2014). The five criteria where directly derived from the end of module summative assessment and mapped to one, or a combination of the cognitive, affective or psychomotor learning domains from descriptors provided by Rovai, Wighting, Baker, and Grooms, (2009) during the development of their self-reporting instrument. The aim was to maintain the structure of the assessment form the lecturers were familiar with to enhance the validity of the assessment and the reliability to which it is completed. The assessment form was deemed to be accurate as the formative assessment form positively correlated with summative results (Cooper and Higgins, 2014). Totals were summed to provide an aggregate score and a score for each criteria, averages were then calculated accordingly.

Although the scores were gained within groups, they were assigned to each individual (within the group) to provide an individual score for each participant throughout the duration of the study. The students worked in different groups each week ensuring no individual score would be the same. In essence, each participant had the potential for nine separate scores within each condition (experimental or control), yielding an overall total for each phase.

Motivation

This scale is used to identify components of the affective domain and was assessed according to the level of enthusiasm, motivation, professional demeanour and efficacy and clarity of the instructions given by the student to the group throughout the rehabilitation session.

Safety

This scale identifies how safe the programme and how controlled the environment was that the rehabilitation participants were conducting their rehabilitation exercises in. The criterion is a core component of any sports therapist's practice and straddles the cognitive and affective domain.

Exercise

The purpose of this criterion is to determine the appropriateness of the exercises the students have chosen and will straddle the cognitive and affective domain.

Timing
This scale was used to quantify how the student was able to manage their time and implement the planned rehabilitation programme logically; this scale straddles the cognitive and psychomotor domain.

Progression
This scale is used to assess the ability of the students to progress the exercises that they have chosen as their patients perform them; as a result this scale straddles the cognitive and psychomotor domains.

Data Analysis
The principal outcome was the difference between the mean aggregate scores and the means of the specific criteria obtained throughout each condition by the student groups. The secondary outcome was the difference between the individual students mean aggregate score and each of the criteria assessed in the study across the two conditions. Cohen's effect size statistics were calculated (Hopkins, 2008) and evaluated using the descriptors and magnitude-based inferences (determined by referring to a three-level scale: beneficial, trivial and harmful) provided by Batterham & Hopkins, (2006).

Individual effect sizes were calculated using the same calculations as above with the addition of Hedges and Olkin's correction for *d* statistic bias that is recommended when there are less than 10 in each sample (Hedges & Olkin, 1985).

Due to the design of the study having to conform to the academic calendar there was no distinct opportunity for a washout period. As a result the potential for a carryover effect was analsyed using the procedure and calculations from Welleck and Blettner (2012).

Results
Group-level analyses
The mean difference in aggregate formative scores for student groups between control and experimental conditions was 2.33 (95% confidence interval 0.57 to 4.08). The magnitude of this effect was moderate (effect size 0.68; 0.04 to 1.31) and the probability that the true population effect is beneficial/trivial/harmful is 93.9/5.4/0.6. The difference between student groups is illustrated in Figure 1. During weeks 1 to 9 there is a significant observable difference in mean aggregate formative scores between the control and video condition with the trough at week 4 being the only time the score of the video group B is lower than the score of the control group A. Additionally the scores of the video group B are generally more consistent than those of the control group A, the latter of which vary considerably almost each week. At week 10 when the student groups crossed over an interesting pattern begins to emerge. There is no distinct

difference as observed in weeks 1-9 for weeks 10-18 between the control group B and the video group A. The control group B maintain their performance throughout this period as the video group A increase their total formative scores typically week on week resulting in the video group A outscoring the control group B in the final week.

Insert Figure 1 here.

Table 1 details the mean difference (with 95% Confidence Intervals) and effect size (with magnitude, 95% Confidence Intervals and the beneficial/trivial/harmful probabilities) for the five individual criterion scores for the student groups between the experimental and control conditions. The most notable of the five criteria are Timing and Progressions as they have a moderate effect size. Despite this all of the criteria have a probability of more being more than 80% beneficial with the most harmful probability only being 3.2%.

Table 1: Individual criterion scores for student groups.

Insert Table 1 here.

Crossover

Figure 1 demonstrates how the videos effectively increased mean assessment scores within both experimental conditions throughout the experimental phase, but it also demonstrates how Group B experienced a carryover effect due to a minimal/no washout period. Using the calculations from Welleck and Blettner (2012) on the data set (n=68, df=134) the pre-test (to check the assumption of negligible carryover effects) T value is -0.019 and is significant ($p < 0.000$) demonstrating that there is a carry over effect and the test for differences yields a T value of -34.76 that is again significant $p < 0.000$.

Individual-level analyses

The mean difference in total formative assessment for individuals between control and experimental conditions was 3.92 (95% confidence interval -4.10 to 11.93). The magnitude of this effect was small (effect size 0.40; -0.40 to 1.21) and the probabilities that the true population effect is beneficial/trivial/harmful are 40.9/56.0/3.2.

Table 2 details the mean difference (with 95% Confidence Intervals) and effect size (with magnitude, 95% Confidence Intervals and the beneficial/trivial/harmful probabilities) for the five individual criterion scores for individuals between experimental and control conditions. The mean differences range from 0.38-0.78 yet despite this the magnitude of each effect size was small with each ranging from 0.32-0.39. As a result the probabilities of each criterion being Beneficial/Trivial/Harmful are comparable throughout.

Table 2: Individual criterion scores for individuals (n=68)

Insert Table 2 here.

Discussion

Group-level analyses

The results from the current study clearly demonstrate the benefit of online instructional videos when teaching rehabilitation skills. The aggregate formative scores between the student groups during the first 9 weeks of the study exemplify the difference that the instructional videos made overall with an effect size of 0.68 and a probable beneficial effect of 93.9% with only a possibility to harm of 0.6%. In every week during the first nine (except one) weeks the video group B scored far higher on the formative assessment than the control group A. When the groups swapped conditions a different trend emerged due to a crossover effect. Normally a crossover contamination were the results are not analysed independently by sequence group are of limited, if any, scientific value (Welleck & Blettner, 2012). In this case however, the crossover effect theorectically illustrates the efficacy of the instructional video as the information was retained by those in the first experimental group and then used by them when they were in the control condition. It is impossible to establish if there was any treatment diffusion between groups however during the first phase of the study, the differences between mean aggregate scores is generally equal. Treatment diffusion may have been present throughout the second phase as the students with prior knowledge of the instructional videos would be able to communicate effectively with precise information and/or details about the resources with the other students potentially gaining reinforcement of prior knowledge. During the second phase of the study Group B, now the control group, continued to outscore the video group A in each week except week 15 (marginally) and in the final week. During this period it is also apparent from Figure 1 that the video group A steadily began to increase their aggregate formative score, which resulted in a more predictable pattern between the two groups as the relative ease or difficulty of each joint and stage followed a similar path.

The study is able to demonstrate how Group A (video) when using the instructional videos consistently outperformed their counterpart Group B, which then enabled them to continually perform at a comparable level even when the resource had been withdrawn, an aspect that has been directly cited as a component of constructivism (Kim & Reeves, 2007; Vogel-Walcutt et al., 2011). This is in agreement with Mayer (2002) who stated that constructivism can be fostered when cognitive activity is present within the learner despite the media being predominantly passive. This also has the potential to demonstrate that higher-order learning can be facilitated with the instructional videos as exemplified by both groups, an aspect that was previously unsupported (Deubel, 2003).

When the criteria were analysed separately the largest effects are measured in Timing and Progressions, criteria that straddle both the cognitive and psychomotor domains. This finding is not unexpected as for the students to foster a constructivist approach cognitive activity needs to be present and as the students perform an increasing number of rehabilitation sessions they gain external and internal feedback relating to the

Timing of exercise and knowledge about how patients can be progressed. Overall each of the criterions demonstrates a positive effect when the students were working in groups with the lowest effect size being 0.46 and the lowest beneficial probability being 80.6%.

Additionally it is also worth noting that the study demonstrates that the 2x2 crossover design is also an ethically sound model as Group B was able to match and exceed the scores of Group A by the end of the study.

Individuals
In addition to the student group data each individual was attributed with a cumulative score from the different groups they participated within throughout the study. Although the scores were obtained within a group setting, it is possible with some assumptions and limitations to track an individual using this method. The mean difference for individuals was 1.59 greater than the mean difference for student groups however this was not all positive as the confidence intervals demonstrate clearly. Greater variance between individuals is not unexpected and the magnitude based inference is most revealing regarding this, especially when compared to the student groups. The student groups magnitude based inference was highly beneficial, however for individuals it is predominantly trivial (56.0%) rather than beneficial (40.9%), it is however worth noting that the potential for harm is only slightly greater than that from the student groups at 3.2%. When these findings are compared to those of Laughlin et al., (2006) it is clear that some of the individuals will have stood out from the student groups and in theory outperformed them as demonstrated by the mean difference, but this is the exception rather than a consistent finding. The student groupings within this study were of three or more therefore the findings identified within pairs (Lou et al., 2001) are not directly applicable, however it is mentioned as this demonstrates that smaller groups and individuals are capable of outperforming groups of three or more.

The individual-level analysis of the criteria demonstrates that although the range of mean differences and the distribution is similar to that of the groups the effect sizes are more comparable and only range from 0.32-0.39 and are all small in nature. Therefore no particular criterion or domain emerges as being most beneficial to an individual over a group.

A limitation of the study is the lack of a washout period, however those undertaking research within an academic timetable may not have such an opportunity due to calander constraints. It is therefore recommended that a 2x2 crossover study is utilised ideally with the winter holiday as a washout period if possible. Another potential limitation of the study relates to the participants and their allocation, whilst it was an effective mechanism for condition allocation it was not a random allocation and did not fully ensure that the groups were balanced academically. That said however, the studies design and the minimisation of treatment diffusion were of paramount importance and the final phase of the study demonstrated that the groups were able to produce comparable standards.

The current study has identified a far greater effect size for the use of instructional videos for student groups and individuals alike than reported in Cooper and Higgins, (2014). In addition, the current study has revealed the specific criteria were instructional videos improved group and individual performances.

Conclusion

In summary the instructional videos have consistently proved to be beneficial at best or trivial at worst for students when learning and performing rehabilitation skills in both a group and individual context. Whilst there was a greater variance within the individual analysis the probability that the instructional videos would have a harmful effect was never greater than 5%. In addition the instructional design and dissemination of the videos in theory has led to a student centred learning approach that fostered constructivism allowing students to retain acquired knowledge leading to a consistent performance without the use of the resource.

Statements on open data and conflict of interest

Open Data

The raw data is available to interested parties via request to the corresponding author. The corresponding author will attempt to place the data onto the University of Worcester's Research and Publication repository (https://eprints.worc.ac.uk/).

Conflict of Interest Declaration

The authors have no conflicts of interest relating to the study.

References

Alonso, F., Lopez, G., Manrique, D., & Vines, J. M. (2008). Learning objects, learning objectives and learning design. *Innovations in Education and Teaching International*, *45*(4), 389–400. doi:10.1080/14703290802377265

Alonso, F., López, G., Manrique, D., & Viñes, J. M. (2005). An instructional model for web-based e-learning education with a blended learning process approach. *British Journal of Educational Technology*, *36*(2), 217–235.

Batterham, A. M., & Hopkins, W. G. (2006). Making meaningful inferences about magnitudes. *International Journal of Sports Physiology and Performance*, *1*(1), 50–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19114737

Benedict, L., & Pence, H. E. (2012). Teaching Chemistry Using Student-Created Videos and Photo Blogs Accessed with Smartphones and Two-Dimensional Barcodes. *Journal of Chemical Education*, *89*(4), 492–496. doi:10.1021/ed2005399

Cooper, D., & Higgins, S. (2014). The effectiveness of online instructional videos in the acquisition and demonstration of cognitive, affective and psychomotor rehabilitation skills. *British Journal of Educational Technology*, *46*(4), 768–779. doi:10.1111/bjet.12166

Deubel, P. (2003). An Investigation of Behaviorist and Cognitive Approaches to Instructional Multimedia Design. *Journal of Educational Multimedia and Hypermedia*, *12*(1), 63–90.

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

Hill, G. (1982). Group versus individual performance: Are N+ 1 heads better than one? *Psychological Bulletin*, *91*(3), 517–539. Retrieved from http://psycnet.apa.org/journals/bul/91/3/517/

Hopkins, W. G. (2008). Estimating Sample Size for Magnitude-Based Inferences. *Health (San Francisco)*, *2006*, 63–70.

Kala, S., Isaramalai, S.-A., & Pohthong, A. (2010). Electronic learning and constructivism: a model for nursing education. *Nurse Education Today*, *30*(1), 61–66. doi:10.1016/j.nedt.2009.06.002

Kim, B., & Reeves, T. C. (2007). Reframing research on learning with technology: in search of the meaning of cognitive tools. *Instructional Science*, *35*(3), 207–256. doi:10.1007/s11251-006-9005-2

Kruse, N. B., & Veblen, K. K. (2012). Music teaching and learning online: Considering YouTube instructional videos. *Journal of Music, Technology and Education*, *5*(1), 77–87. doi:10.1386/jmte.5.1.77_1

Laughlin, P. R., Hatch, E. C., Silver, J. S., & Boh, L. (2006). Groups perform better than the best individuals on letters-to-numbers problems: effects of group size. *Journal of Personality and Social Psychology*, *90*(4), 644–51. doi:10.1037/0022-3514.90.4.644

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small Group and Individual Learning with Technology: A Meta-Analysis. *Review of Educational Research*, *71*(3), 449–521. doi:10.3102/00346543071003449

Mayer, R. E. (2002). Cognitive Theory and the Design of Multimedia Instruction: An Example of the Two-Way Street between Cognition and Instruction. *New Directions for Teaching and Learning*, (89), 55–71. doi:10.1002/tl.47

Mayer, R. E. (2008). Applying the science of learning: evidence-based principles for the design of multimedia instruction. *The American Psychologist*, *63*(8), 760–9. doi:10.1037/0003-066X.63.8.760

Mayer, R. E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, *91*(4), 638–643. doi:10.1037/0022-0663.91.4.638

Mayer, R., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*(1), 107–119. doi:10.1016/S0959-4752(01)00018-4

Meij, H. Van Der, & Meij, J. Van Der. (2013). Eight Guidelines for the Design of Instructional Videos for Software Training, *60*(3), 205–228.

Moreno, R., & Mayer, R. (1999). Visual presentations in multimedia learning: Conditions that overload visual working memory. *Visual Information and Information Systems*, *1614*, 798–805. Retrieved from http://www.springerlink.com/index/01117565685641m8.pdf

Rovai, A. P., Wighting, M. J., Baker, J. D., & Grooms, L. D. (2009). Development of an instrument to measure perceived cognitive, affective, and psychomotor learning in traditional and virtual classroom higher education settings. *The Internet and Higher Education*, *12*(1), 7–13. doi:10.1016/j.iheduc.2008.10.002

Samaras, H., Giouvanakis, T., Bousiou, D., & Tarabanis, K. (2006). Towards a new generation of multimedia learning research. *AACE Journal*, *14*(1), 3–30. Retrieved from http://www.editlib.org/index.cfm?fuseaction=Reader.ViewFullText&paper_id=5858

Seilstad, B. (2012). Using tailor-made YouTube videos as a preteaching strategy for English language learners in Morocco: Towards a hybrid language learning course. *Teaching English with Technology*, *12*(4), 31–49. Retrieved from http://www.ceeol.com/aspx/getdocument.aspx?logid=5&id=6db4a4a0db514a8d87cdaafaf21d03bf

Smith, G. (1998). Return to fitness. In *Orthopaedic Physiotherapy* (Vol. 270, pp. 255–270). Mosby. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20048467

Swarts, J. (2012). New Modes of Help : Best Practices for Instructional Video, *59*(3), 195–206.

Van der Meij, H., & van der Meij, J. (2014). A comparison of paper-based and video tutorials for software learning. *Computers & Education*, *78*, 150–159. doi:10.1016/j.compedu.2014.06.003

Van Gog, T., Verveer, I., & Verveer, L. (2014). Learning from video modeling examples: Effects of seeing the human model's face. *Computers and Education*, *72*, 323–327. doi:10.1016/j.compedu.2013.12.004

Vogel-Walcutt, J. J., Gebrim, J. B., Bowers, C., Carper, T. M., & Nicholson, D. (2011). Cognitive load theory vs. constructivist approaches: which best leads to efficient, deep learning? *Journal of Computer Assisted Learning*, *27*(2), 133–145. doi:10.1111/j.1365-2729.2010.00381.x

Welleck, S., & Blettner, M. (2012). On the proper use of the crossover design in clinical trials: part 18 of a series on evaluation of scientific publications. *Deutsches Ärzteblatt International*, *109*(15), 276–281.

Wouters, P., Tabbers, H. K., & Paas, F. (2007). Interactivity in Video-based Models. *Educational Psychology Review*, *19*(3), 327–342. doi:10.1007/s10648-007-9045-4

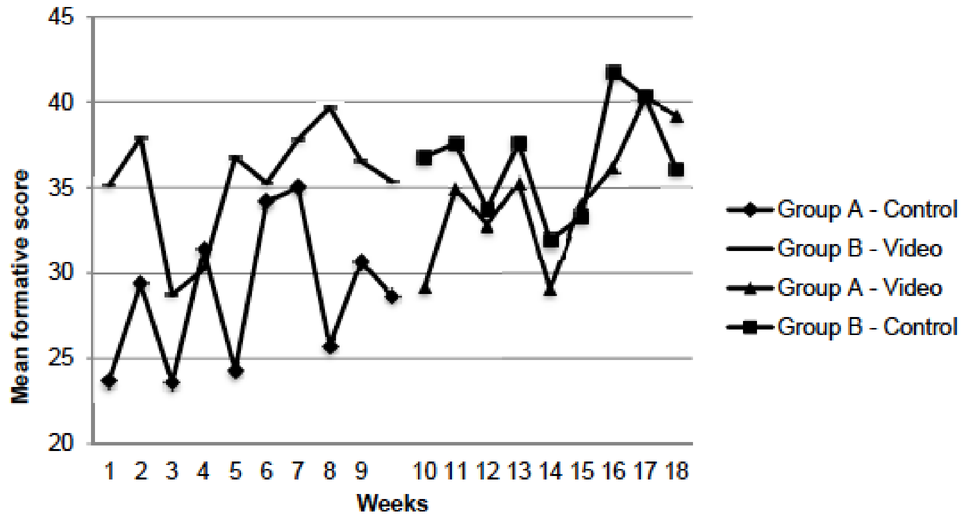

Figure 1: Mean aggregate formative scores

British Journal of Educational Technology submitted article

Table 1: Individual criterion scores for student groups.

Formative Component	Mean Difference	95% Confidence Interval		Effect Size	Magnitude	95% Confidence Interval		Beneficial	Trivial	Harmful
Motivation	0.27	0.02	0.52	0.50	Small	-0.07	1.08	87.0	11.8	1.1
Safety	0.41	-0.04	0.87	0.46	Small	-0.16	1.09	81.9	16.1	2.0
Exercise	0.42	-0.08	0.92	0.48	Small	-0.22	1.17	80.6	16.6	2.8
Timing	0.71	-0.08	1.50	0.72	Moderate	-0.27	1.70	86.9	9.9	3.2
Progressions	0.50	0.05	0.96	0.62	Moderate	-0.07	1.31	90.2	8.6	1.3

Table 2: Individual criterion scores for individuals (n=68)

Formative Component	Mean Difference	95% Confidence Interval		Effect Size	Magnitude	95% Confidence Interval		Beneficial	Trivial	Harmful
Motivation	0.38	-1.14	1.90	0.34	Small	-0.46	1.15	35.8	60.9	3.3
Safety	0.42	-1.01	1.86	0.35	Small	-0.41	1.10	35.9	57.7	3.3
Exercise	0.78	-1.08	2.68	0.35	Small	-0.45	1.16	37.3	59.1	3.5
Timing	0.73	-1.70	3.16	0.32	Small	-0.47	1.10	34.3	60.8	4.9
Progressions	0.74	-1.24	2.73	0.39	Small	-0.41	1.20	39.6	57.6	2.8

British Journal of Educational Technology submitted article

Practitioner Notes

What is already known about this topic

- The use of instructional videos is becoming increasingly commonplace
- Instructional videos are becoming easier to develop and disseminate via e-learning
- The potential for instructional videos is worthy of research to determine their efficacy

What this paper adds

- Instructional videos can be a beneficial resource for groups and individuals alike
- A crossover contamination effect demonstrates the potential of higher order learning
- Revising a methodology can be beneficial for the research findings of a study

Implications for practice and/or policy

- Instructional videos have proved to be beneficial the teaching of practical skills
- Individuals are capable of outperforming small groups
- The use of instructional videos led to a student centred approach that fostered constructivism enabling students to retain and apply knowledge after the resource had been withdrawn