Designing Effective Multimedia for Physics Education

A thesis submitted in fulfillment of the requirements for the degree of
Doctor of Philosophy

by

Derek Alexander Muller

School of Physics
University of Sydney
Australia

2008

Dedicated to all the great teachers who instilled in me the joy of learning,
my parents, sisters, and especially Jane.

Declaration of originality

To the best of my knowledge, this thesis contains no copy or paraphrase of work published by another person, except where duly acknowledged in the text. This thesis contains no material which has been previously presented for a degree at the University of Sydney or any other university.

Derek Alexander Muller

Included papers and attribution

The following refereed papers and presentations arose from work related to this thesis.

Chapter 1 **The future of multimedia learning: Essential issues for research**
D.A. Muller, J. Eklund, & M.D. Sharma
Paper presented at the Australian Association for Research in Education, Sydney, 2005.

Chapter 4 **Determining the factors affecting student perceptions of a popular science video**
D.A. Muller & M.D. Sharma
Published in the *Australasian Journal of Educational Technology*, 21(4), 2005, pp.491–509.

Video physics education: Falling cats and terminal velocity
D.A. Muller & M.D. Sharma
Paper presented at The Australian Institute of Physics, Canberra, 2005.

Chapter 5 **Inside the quantum mechanics lecture: Changing practices**
D.A. Muller
Paper presented at The Higher Education Research and Development Society of Australasia, Sydney, 2005.

Chapter 6 **Student conceptions of quantum tunneling**
D.A. Muller and M.D. Sharma
Paper presented at the International Conference on Physics Education, New Delhi, 2005.

Chapter 8 **Conceptual change through vicarious learning in an authentic physics setting**
D.A. Muller, M.D. Sharma, J. Eklund, & P. Reimann
Published in *Instructional Science*, 35(6), 2007, pp.519–533.
Starting a dialogue: The potential of vicarious learning
D.A. Muller, M.D. Sharma, P. Reimann, & J. Eklund
Presented at The American Association of Physics Teachers, Anchorage, 2006.

Chapter 9 **Changing conceptions in E-learning: Debates and dialogues**
D.A. Muller & J. Eklund
Paper presented at The World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, Honolulu, 2006.
Saying the wrong thing: Improving learning with multimedia by including misconceptions
D.A. Muller, J. Bewes, M.D. Sharma, & P. Reimann
Published in the *Journal of Computer Assisted Learning*, 2007.

Chapter 10 **Improving multimedia with alternative conceptions**
D.A. Muller & J. Eklund
Paper presented at The International Conference on Information Communication Technologies in Education, Crete, 2007.
Raising cognitive load with linear multimedia to promote conceptual change
D.A. Muller, M.D. Sharma, & P. Reimann
Published in *Science Education*, 2007.

Chapter 11 **Tackling misconceptions in introductory physics using multimedia presentations**

D.A. Muller and M.D. Sharma

Paper presented at the Uniserve Science Conference, Sydney, 2007.

Acknowledgements

This thesis would not have been possible without the efforts of my supervisor, Manju Sharma. Her unwavering enthusiasm kept my spirits up and encouraged me to develop and realize my research goals. She is a uniquely gifted researcher and educator because of her passion for physics and, most of all, her genuine concern for others.

I am greatly indebted to my co-supervisors in education and industry. Peter Reimann has provided critical insight and timely resources to propel my searches and analyses. John Eklund has brought a very worthwhile different perspective from the industry sector. His presentations of our work have also broadened the scope and accessibility of our research.

To the rest of the SUPER group: Ian Johnston, Ian Sefton, and Brian McInnes, thanks for your suggestions, time, and support. You represent a wealth of knowledge and one of the greatest assets to physics education in Australia.

This thesis would also have been impossible without the technical contributions from friends. Cyrus deserves more rewards for the countless hours he spent developing the web testing interface than I even promised. Thanks to Shelly for filming, and Jacqui, Emily, James and Kester for lending your acting talents to this project.

The students in the School of Physics have made this time very enjoyable. Thanks to Luke for being such a tolerant house and office mate when my forgetful habits bordered on the infuriating. Sacha has graciously endured many hours of tennis and debates on education and everything else. Jacqui, your support and keen editor's eye have been invaluable to my pursuit. To my office mates, Christine, Michelle, and Roberto, you will have to maintain our reputation as the hardest working office in physics. Indoor soccer has been a memorable experience with Alberto, Alex, Phil, Jong Won, Eddie, Iver, Dinko, Rob, and the rest of the Physicists crew.

I think one of my greatest fortunes is being blessed with a family that has supported me in all my endeavors even when that meant traveling thousands of miles to live on a different continent. Though we are often separated by great distances, I hold you very close to my heart and I appreciate that your love gives me the freedom to explore the world.

This work was supported by the Australian Postgraduate Award Scheme and the Australian Research Council.

Abstract

This thesis summarizes a series of investigations into how multimedia can be designed to promote the learning of physics. The 'design experiment' methodology was adopted for the study, incorporating different methods of data collection and iterated cycles of design, evaluation, and redesign.

Recently much research has been conducted on learning with multimedia, usually from a cognitive science perspective. Principles of design developed in this way have not often been tested in naturalistic settings, however.

Therefore in one preliminary investigation students' perceptions of a popular science video were investigated. Opinions aligned well with most principles though areas for further research were identified.

In order to understand the challenges and opportunities presented by physics teaching, a survey of all lecture courses on the topic of quantum mechanics was undertaken. The lectures were a sophisticated form of multimedia, however interactivity in all lectures was low.

The learning that results from this teaching was evaluated using a questionnaire on quantum tunneling, a key quantum mechanical phenomenon. The survey revealed that students had many alternative conceptions on the topic and that these could be grouped into a small number of alternative answers. This finding is similar to many of the findings from science education over the past three decades.

Using this background, two multimedia treatments were developed to teach the topic of quantum tunneling. One consisted of a lecture-style explanation with only correct information presented. The other took the form of a dialogue between a tutor and student, involving several of the common alternative conceptions. Students who saw the Dialogue performed significantly better on the post-test than those who saw

the Exposition.

In order to generalize the findings, four multimedia treatments on Newton's first and second laws were created and evaluated in a similar way. A refutationary treatment, in which alternative conceptions were stated and refuted by a single speaker, and an Extended Exposition treatment were evaluated in addition to the Dialogue and Exposition. The Dialogue and Refutation outperformed the two expository treatments, confirming the benefits of including alternative conceptions.

In a third iteration of the design experiment, four Newtonian mechanics treatments were evaluated with a new cohort of students. The Extended Exposition was replaced by a Worked Examples treatment in which important details were repeated to solve numerical problems. Cognitive load was directly measured in this experiment. Results showed that treatments containing alternative conceptions involved higher cognitive load and resulted in higher post-test scores than the other treatments.

Contents

List of Figures

List of Tables

Chapter 1

Framing the study

During the course of this doctoral thesis I was awarded a grant to develop multimedia for senior high school physics students. In the initial stages of the project, my team of teachers, academics, educational technologists, and I selected syllabus topics that we thought would benefit from additional multimedia resources. We planned to develop tools that could be distributed freely over the Internet for students to use in their own time or for teachers to show and discuss in class. Searching the physics education literature we honed in on topics that students often find confusing. We noted common misconceptions and the methods that have shown some success in achieving conceptual change in the classroom. Through web searches and discussions with educators we determined which physics concepts were thoroughly covered in textbooks and online and which had few or unclear resources.

The only question that remained at the conclusion of our preliminary analysis was: once the learning objectives are specified, how does one produce effective multimedia for physics education? This question had been the focus of my PhD research and now I was faced with it in a very real sense. The problem of designing multimedia to promote learning is a common one yet it has received uneven and, until recently, inadequate attention from academics. This thesis represents one attempt to understand the challenges and opportunities presented by multimedia for the learning of physics.

Although the term multimedia has many definitions and connotations, in this thesis, it refers to any presentation that combines words and pictures to form a coherent message. Illustrated books, lectures that include diagrams, and animations with narration all therefore constitute multimedia. This might seem like an incredibly broad definition, but there are good theoretical and empirical reasons for collecting such a diverse range of presentations under one banner (Clark 1994*a*, Mayer 1997), as I will explain in this chapter.

1.1 The rise of multimedia

Since the development of language only a few hundred thousand years ago, multimedia has grown in its sophistication and availability. The first recorded multimedia probably consisted of hieroglyphs and paintings on stone tablets. It was undoubtedly time-consuming to create, and comprehensible only to scribes and scholars. In the centuries that followed, multimedia was rare and accessible only to the educated upper classes.

Several inventions led to significant advancements in the development of multimedia. One was the printing press, which, after 1450, allowed large volumes of text to be readily copied and distributed. Lithography, a similar technique for printing images, was developed in 1796. Photography followed in the early nineteenth century. These inventions made text and accompanying pictures available to more of the population at less expense.

The development of the motion picture marked another important milestone at the end of the nineteenth century. The quick succession of static images created the illusion of motion with explanatory text interspersed at intervals throughout the movie. The early twentieth century saw the invention of 'talking pictures,' with speech and sounds synchronized to the action in the film. This meant that literacy was no longer a barrier to understanding multimedia.

Television and video permitted a similar experience to film, but at less expense and with greater flexibility. Again, words and pictures became more widely available. Digital video discs (DVDs) and interactive video discs provided higher quality sound and images, and added an element of interaction between user and multimedia. Finally, over the past two decades, the computer has been transformed from a calculator and word processor into a multi-faceted multimedia communication device.

It goes without saying that we are currently experiencing the greatest information-sharing explosion in human history. Multimedia availability is increasing at an extraordinary rate with new online repositories and distributors appearing daily. The Internet is maturing, bandwidths are expanding, new software is being created, and hard drive capacities are on the rise. It has never been easier to create or distribute multimedia. Video, one of the most common multimedia formats, has become so widely available that a popular website *YouTube.com* regularly handles over 100 million video downloads per day (BBC News 2006).

The objective of this thesis was to investigate how multimedia can be designed and used to promote the learning of physics. I focused mainly on the video form because it encompasses most attributes of other multimedia, however, this research should have implications for other approaches.

Because multimedia is being used at all levels of education, studying and improving its effectiveness is a significant and worthwhile challenge. It would be ideal if students could learn about science by working in groups, devising and performing experiments, and discussing their ideas with knowledgeable, experienced teachers. However, resources are limited and students must often learn by themselves with textbooks, videos, and online multimedia. Furthermore, after leaving formal education, learners must be able to build on their knowledge with different types of learning resources.

There are additional reasons for studying multimedia. Almost all learning experiences, whether interactive or not, consist of segments of linear multimedia; examining this building block can arguably provide insights into more complicated pedagogical methods. In addition, multimedia provides a confined arena in which to test different instructional strategies with large cohorts of students in real learning environments. Novel teaching implementations in science education have been criticised for varying several aspects of instruction simultaneously without attempting to understand the features essential to their success (e.g. Guzzetti, Snyder, Glass &

Gamas 1993). Multimedia offers a transparent and repeatable way to study specific aspects of the teaching and learning process.

1.2 The research that wasn't there

Given that people have been using multimedia in education for decades, it seems reasonable to expect a sizable body of research to exist on how it may best be designed. Unfortunately, this does not seem to be the case. At the outset of this research, Moore, Burton & Myers (2004) summarized in their review of the topic that "with few exceptions there is *NOT* a body of research on the design, use and value of multimedia systems" (p.997, emphasis in original). Although the exceptions referred to in the preceding quote form the theoretical foundations of this thesis (see Chapter 3), it is startling that a century of research and use of educational technology has yielded so few productive outcomes.

This lack of research is readily apparent in the literature. A study in the *American Journal of Physics* illustrates the types of questions that have been asked repeatedly in educational technology studies, with little success. Lewis (1995) explored the impact on students' grades and attitudes of replacing standard tutor introductions to experimental laboratories with video introductions. The videos did not include anything that was not part of the usual tutor presentations. Perhaps unsurprisingly, the researcher found that students' marks were the same with the videos as they were with tutor introductions. Accounting for this result, Lewis speculates, "it may be that the video medium is unsuitable for the purpose of laboratory introductions or that the particular videos used here were deficient in content or presentation" (p.469). He further supposes that it may be the 'passive' nature of video that limited its effectiveness. He does not consider, however, the possibility that one standard multimedia presentation may be as good as another. Why should one expect a video to outperform a tutor, presuming the tutors are knowledgeable and readily available during laboratory?

In another study, Rieber, Tzeng & Tribble (2004) measured the learning about Newtonian mechanics that resulted from several different instructional treatments.

All students interacted with a computer simulation in which the goal was to move a frictionless 'ball' to a specified target. Half of the students received graphical feedback while the other half received textual feedback. In addition, only half of the sample received brief multimedia explanations of the physics involved, interspersed throughout the simulation. The best-perfoming group on the post-test was the graphical feedback with multimedia explanation group. Without a multimedia-only control, the authors conceded "the issue of how much learning is taking place just by having participants view the explanations *without* participating in the simulation is open to question" (p.321, emphasis in original).

Kim, Yoon, Whang, Tversky & Morrison (2007) investigated student learning about bicycle pumps using multimedia materials. Still graphics were presented under four conditions: (1) all at once, (2) successively, (3) self-paced, or (4) animated. It was thought that the animated materials might have a superior effect because they could be seen as "more interesting, aesthetically appealing, and therefore more motivating" (p.261). Presentation mode did affect student perceptions of the materials, including interestingness, enjoyment, and motivation; however, comprehension test scores did not differ among the groups.

The three studies outlined above are symptoms of a body of research that has failed to establish answers to fundamental questions about learning with multimedia. Aspects of these studies typify the difficulties with educational technology research and help understand why a more relevant theoretical base hasn't been established. During the twentieth century, film, radio, television, video, and computers were all introduced into classrooms at different times. The patterns of their implementation, use, and supporting research bear striking similarities, with all technologies failing to live up to expectations.

Research on these educational technologies did not establish a general and robust theoretical foundation for designing multimedia for several reasons:

1. The advantages of new technologies were seemingly self-evident. So much hype accompanied each innovation that rigorous research was seen as unnecessary.
2. The questions asked by researchers were generally media-comparative and

disconnected from theoretical considerations.

3. The practical drive to introduce technology into schools limited the time in which research was carried out, transferring the burden of using and proving the efficacy of new inventions to designers.
4. Fundamental perspectives on how people learn shifted continually over the past century.

These four points are addressed in the sections below.

1.2.1 Technology: The obvious solution to our problems

The implementation of any new technology into education has typically begun with incredible rhetoric and expectations. Marketers and technology developers have focused on the ground-breaking abilities of the new technology to promote interest in its application to the educational domain. Thomas Edison's appraisal of the motion picture is an oft-cited example of the excitement that accompanies innovation. Promoting his invention, he proclaims "that the motion picture is destined to revolutionize our educational system and that in a few years it will supplant largely, if not entirely, the use of textbooks," (Edison 1922 as cited in Cuban 1986, p.9). Such claims are not restricted to a bygone era, however; for example Semrau & Boyer (1994, p.2) note "the use of videodiscs in classroom instruction is increasing every year and promises to revolutionize what will happen in the classroom of tomorrow." Clark & Estes (1999) attribute the ineffectiveness of past research programs, at least in part, "to a history of mindless and demonstrably wrong advocacy of popular electronic media to foster motivation and learning" (p.5).

Another common element of marketers' campaigning is a contrast between the promises of new technology and the existing state of education. Pessimistic claims about the school system have been routinely juxtaposed with the dramatic prophesies for future technologies. For example, Edison took aim at textbooks.

> I should say that on the average we get about two percent efficiency out of schoolbooks as they are written today. The education of the future, as I see it, will be conducted through the medium of the motion picture

> . . . where it should be possible to obtain one hundred percent efficiency. (Edison 1922 as cited in Cuban 1986, p.9)

Where Edison comes up with the figure of two percent for the efficiency of textbooks is unclear, as is the notion of one hundred percent efficiency, but his argument was understandably a persuasive one for politicians and the public alike. In his time, the technological revolution was in full swing and people were eager to consider the concept of efficiency as it related to agriculture, steam engines, and education.

The excitement surrounding new technologies diverted attention away from rigorous research. Researchers and the general public were intuitively convinced of the effectiveness of new inventions. "Their reasoning seems to suggest that if research does not find evidence for something that *seems* so powerful, then research as an inquiry strategy *must* be flawed" (Clark & Estes 1998, p.5).

1.2.2 Is this medium better than the other one?

Researchers adopted the perspective that educational efficiency could be measured and optimized, and began to investigate the intrinsic advantages of one medium over another (Russell 1985, p.47). The medium itself seemed the obvious variable for investigation, rather than the experience of the learner. McLuhan's (1964) refrain 'the medium is the message,' focused attention on new and exciting inventions, fuelling the technology-centred approach. Early studies compared the performance of students who watched an instructional film to those who received only traditional lecture instruction, in experiments similar to Lewis's (1995) study. The results showed increased motivation among students who watched films and either superior or equivalent academic performance compared to a control group (Cuban 1986). Excitement due to novelty, methodological confounds, or a Hawthorne effect likely account for much of the success of these studies (Clark 1983). Similar research on educational television showed impressive results, increasing math, science and reading scores on standardized tests. However, researchers did not ensure the socioeconomic statuses of different treatment groups were comparable and failed to report any differences in this measure (Cuban 1986, p.35). Even when comparative research showed negative or no significant difference results for new technology, its promoters used the enthusiasm, assumptions, and excitement surrounding the technology as effective counter-arguments.

When Clark (1983) concluded that no particular media had a unique impact on learning and that research seeking such an impact should be abandoned, he believed the point to be uncontentious and well-supported by the evidence. "Media are mere vehicles that deliver instruction but do not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition," he wrote (p.445). The paper kicked off debate in the research community because, stated explicitly or not, the notion that media inherently affects learning had been a presupposition of virtually all previous studies in educational technology. Many researchers debated the claim, although perhaps the best articulation of opposing viewpoints is contained in the writings of Clark and Kozma (Clark 1988, 1994*a*, 1994*b*, Clark & Salomon 1986, Kozma 1991, 1994*a*, 1994*b*, 2000, Kozma & Anderson 2002).

Kozma (1994*b*) argued that different media have particular capabilities, which enable different learning experiences. "A particular medium can be described in terms of its capability to present certain representations and perform certain operations in interaction with learners who are similarly engaged in internally structuring representations and operating on these" (p.11).

Clark (1994*a*) maintained that the learning experiences in any form of multimedia could be made almost identical to any other with adequate preparation. For example, the technique of zooming used in video to focus on a component of a larger system could be illustrated diagrammatically with a magnification bubble. Based on this interchangeability, he proposed the 'replaceability' challenge: "to find evidence, in a well designed study, of any instance of a medium or media attributes that are *not* replaceable by a different set of media and attributes to achieve similar learning results for any given student and learning task." The challenge was meant to demonstrate the equivalence of different platforms and highlight the methodological differences that actually impact on learning.

It cannot be overlooked that the history of educational technology, with new technologies repeatedly delivering much less than they had promised, seems to bear out Clark's argument. After a technology's initial implementation into schools, interest has waned and its use has declined. For example, instructional television and computers occupied at most four to eight percent of instructional time, even in well financed schools with professed interests in implementing technology (Cuban 1986). Thus, it is understandable when critics claim technology has failed to live up to expectations (Tyack & Cuban 1995). If there were even one instance where a particular media afforded a unique and profound benefit over competing technologies, would it not be widely adopted and repeatedly cited as evidence by media proponents?

1.2.3 Implementing technology in schools

Although the outcomes of early media research may have been dubious, they were sufficient to encourage educational administrators to implement new technologies in schools. This worsened the research deficit as technologies gained the appearance of maturity and academics became experts in a field with little theoretical or empirical basis.

> No sooner had my colleagues and I begun exploring the potential use of the computer for teaching science than colleges began offering Master's degrees in computer education. Although no one had any knowledge or experience using computers to teach anything, experts were instantly trained, hired, and funded to bring computers into the public schools. (Cromer 1997, p.108)

Similar observations are made to this day of university educational technology programs that de-emphasize scientific research and work under the assumption that technology is inherently beneficial (Clark & Estes 1998).

Most recently school districts have invested incredible sums of money to bring computers into the classroom.

> The adoption of microcomputers by U.S. schools has been explosive, going from essentially zero in 1980 to better than one for every nineteen students by the early 1990's. The computer's educational roles change from year to year, as their functionality evolves, and today their purposes are as unclear as they are unquestioned. (Cromer 1997, p.121)

The rapid introduction of technology to education has been witnessed many times. In the 1930's, with the drop in price of radio receivers, governments invested heavily in educational radio broadcasting. A majority of schools bought into "the textbook of the air" and purchased at least one receiver set (Cuban 1986, p.19).

With a particular technology available in classrooms almost from the date of its invention, instructional designers have carried the burden of the outlandish promises made by technology salesmen. To this day, designers must make do with what little reliable literature is published and use intuition or industry rules of thumb to make the balance of decisions. These designers are also the only qualified sources to write textbooks on instructional design based on their experience. Blinn (1989) outlined a number of design criteria for educational animations, drawing on his experience as an animator for a physics education video series. Although likely very useful, these guidelines are a starting point for investigation rather than established principles of best practice. Yet these types of resources have been the only references for designers making costly decisions about how to create multimedia. Clark & Estes (Clark & Estes 1998, 1999, Estes & Clark 1998) have classified educational technology produced in this way as 'craft' solutions, uninformed by scientific research. They claim that it is the lack of concrete theoretical foundations and the subsequent proliferation of craft technologies that has led to the unreliability of technological solutions.

These craft solutions are the most common type of educational technology, and, since they are not developed or evaluated scientifically, are unable to directly inform the body of research on learning with technology. This perpetuates the cycle of craft educational technology, further inhibiting progress in the field.

1.2.4 How do learner's learn?

Another factor that has limited the development of multimedia research is the shifting perspectives in the educational literature of what constitutes learning and how it is achieved. Early behaviorist research rejected the notion of cognitive entities and focused instead on observable actions. Under this paradigm, punishment and rewards were used to modify behavior. In the 1960's this view was displaced by cognitivism. However, within this movement, different branches of research have formed with little overlap between groups. With social and radical constructivism, information processing models, connectionism and associationism, it has been incredibly difficult to form a coherent body of knowledge.

In media research, very few studies were based on theoretical frameworks that accounted for the effects of technological interventions (Clark 1994*a*). When learning theories were employed, they were dependent on a delivery model of education (Kozma 1994*b*, Clark & Estes 1999). Media were viewed as delivery vehicles, permitting the question 'does this medium deliver information more efficiently than other media?' In fact, it has only been in the last decade that researchers in the field have moved to a constructivist paradigm. "It is time to shift the focus of our research from media as conveyors of methods to media and methods as facilitators of knowledge-construction and meaning-making on the part of learners" (Kozma 1994*a*, p.13).

1.3 The equivalence principle

In sum, until recently volumes of educational technology research have yielded little theoretical or practical guidance for the design of multimedia. It is arguable that the most significant conclusion yielded by previous studies is that the learning experiences with and without a particular technology can be made equivalent with adequate forethought. On this matter, it is worth pursuing an analogy with an equivalence principle from physics that is central to the theory of General Relativity. When it dawned on him, Einstein called this principle his happiest thought.

For all of human history before the twentieth century, gravitational forces were perceived quite separately from the concept of acceleration. A person experiences a strong gravitational force any time he or she is in close proximity to a large mass, as is the case on the surface of the earth. Acceleration, on the other hand, occurs any time one's velocity is changing, during space shuttle takeoff, for example. Although forces are involved in both cases, the two phenomena appear entirely distinct from each other. One occurs due to the presence of a large mass, while the other occurs due to changes in motion.

Now consider an astronaut in a space shuttle with no windows. What could she conclude if she woke up to find a force pressing her into her seat? She might be at rest on the launchpad, in Earth's gravitational field, or she might be in deep space experiencing no gravitational force but accelerating at a constant rate. The two instances appear very different but the astronaut's experience of them is identical. This is the equivalence principle.

There seems to be an equivalence principle in learning with multimedia, albeit much less profound, which parallels the equivalence principle of General Relativity. Although others have expressed similar ideas about the interchangeability of multimedia, I apply the term equivalence principle to emphasize similarities with its physics counterpart.

Consider a student reading a book with words and pictures about Newtonian mechanics. Then, consider the same student watching a movie about Newtonian mechanics. The two experiences appear very different. One involves written text and static images while the other involves spoken text and dynamic images. If we found that following these two instructional treatments, our hypothetical student performed equally well on the same test, what could we conclude about the two different forms of multimedia? We might, like Lewis (1995), suspect that a movie may not be an appropriate medium for teaching Newtonian mechanics, or perhaps that the movie was deficient in content or presentation. The alternative is to conclude that both media encouraged similar cognitive processes in the student. The equivalence principle in multimedia then states that the relevant cognitive processes inspired by different formats of multimedia can be made indistinguishable, by choosing appropriate methods.

Both equivalence principles, like any new ways of looking at the world, unravel the alleged paradoxes of previous research. Although experiments in search of the aether were undoubtedly useful in developing our understanding of space-time, accepting that there is no privelaged reference frame made repeated precise measurements seem unnecessary. Similarly, experiments searching for 'media effects,' were always doomed to failure by the multimedia principle of equivalence.

Both equivalence principles change the way their related phenomena are viewed and illuminate critical areas for consideration. In General Relativity, matter, it was realized, warped space-time so the distribution of matter in the universe and the geometry of space became central concerns. With the multimedia equivalence principle, the cognitive processes necessary for learning and methods by which they can be triggered become central areas of investigation. Furthermore, the multimedia equivalence principle implies that teaching and learning techniques developed in different forms of multimedia learning can be applied with similar successes across platforms.

1.4 Conclusion

Since its invention, multimedia has become increasingly sophisticated and accessible. Most recently, computer technology has allowed for the creation and propagation of multimedia with increasing speed. The development of the technology itself has far outpaced efforts at understanding how people learn with multimedia (Rieber 1990). Excitement and intuition displaced research and critical thinking at the outset of each new educational technology. Comparative media studies sought but failed to find evidence of media effects. This line of reasoning obscured the need for theories that explain the interaction between learner and multimedia and how it gives rise to productive cognitive activities for learning. The urge to introduce technologies into schools limited the research that was done and lent an air of maturity to the technologies, discouraging further research. Finally, with shifting perspectives of the teaching and learning process, establishing a coherent base of theory was next to impossible.

Despite the lack of reliable supporting research, multimedia technologies have become commonplace in educational establishments. The costs of technology have dropped dramatically such that virtually every student has access to a computer (Cuban, Kirkpatrick & Peck 2001), where, during the introduction of film, schools had at most one projector. Sophisticated multimedia has also become a larger part of students' lives with films, television, and the Internet accounting for much of their entertainment and education. Students rely on computers to produce reports and on the Internet to access virtually limitless amounts of information instantly. Technology failed to live up to the promises of its promoters, but it has permeated the school system quite independently of the work of researchers. Without appropriate supporting research however, the successes of multimedia are bound to be unpredictable (Sweller 2004). We are doomed to invest significant amounts of money, time and effort in developing multimedia resources that fail to promote meaningful learning.

The problem can be refined in terms of the questions considered by multimedia developers. For example, when are interactive resources advantageous over non-interactive media? How does one handle, if at all, the topic of misconceptions? Should the material be presented by a single speaker as in a lecture? Should this speaker appear on-screen or provide narration only? Are different methods advantageous for novice and expert learners? Should interesting examples be included to keep the viewer's attention if only of tangential relevance? Although the possible questions of this type are endless, there are clearly a handful that are vital for understanding and developing effective multimedia.

The equivalence principle, that all forms of multimedia can be made equally effective, yields three major implications for this study. First, it warns against searching for differences in learning simply due to the use of different media, an enterprise that has a long history of failure. Second, and more informatively, it suggests that teaching and learning experiences that have proved effective in general educational studies can be recreated in multimedia. If these experiences are inherently rare or difficult to facilitate, multimedia can act as a substitute. Third, the equivalence principle underscores the strong link required between sound theory and experiment. In order for a learning experience to be uniquely beneficial, there must be strong theoretical support for the mechanisms proposed and the anticipated results. Similarly, the results must be capable of informing theory to move both design and research forward.

For media proponents, acceptance of the equivalence of learning experiences with different technologies, from textbooks to lectures to computers, would eliminate the need for further research; however in this thesis, it is this very equivalence that forms the jumping-off point. Decades of unproductive comparative media studies have left unanswered the questions relating to which methods can be employed in multimedia to achieve the greatest conceptual learning gains.

Although the focus of the research is on establishing principles for effective multimedia design, the implications of the research should be generalizable across a range of environments in which students learn from words and pictures. Virtually any presentation that can be created in a classroom or lecture setting can be recreated as multimedia. In addition, the stable and reusable nature of multimedia makes it an ideal tool for carefully investigating methodological differences in teaching.

1.5 Advance organizer

The main finding of my thesis research is that multimedia which involves explicit discussion of alternative conceptions is more effective for learning than more concise expository summaries. This was demonstrated three times in two different areas of physics with students with different levels of prior knowledge. Supporting data from an empirical study on quantum mechanics are located in Chapter 8 and data from two Newtonian mechanics studies are reported in Chapters 9 and 10.

Students were better able to learn with misconception-based multimedia, in which they also invested more mental effort. The construct of mental effort was initially introduced in Chapter 3 as it relates to cognitive load theory and it was measured directly in Chapter 10.

The three multimedia learning studies were informed by two main bodies of theory, constructivism (Chapter 7) and cognitive theories of learning (Chapter 3). Three preliminary studies also helped identify the research questions and methods. Multimedia learning theories were investigated in their applicability to authentic classroom practice (Chapter 4). Quantum mechanics teaching (Chapter 5) and resulting learning (Chapter 6) were explored to understand the challenges and opportunities presented by physics education in the local context.

Chapter 2

Methodology

As outlined in Chapter 1, theories of multimedia design and use are still developing and thus far it has been difficult to bridge the gap between research and practice. Therefore the 'design experiment' methodology was selected for this investigation. This research method is becoming increasingly accepted especially in the study of teaching interventions in authentic settings (Lagemann & Shulman 1999, Klahr & Li 2005). The methodology makes use of numerous different data collection and analysis techniques and iterative cycles of design, development, implementation, analysis, and redesign. The underlying goal of design experiment research is to build upon theory while developing effective interventions in authentic contexts.

In this chapter I outline the design experiment methodology, its origins, characteristics, strengths, and weaknesses. I discuss how the methodology was applied in this investigation including how challenges were addressed. In the process, I present the layout of the thesis, indicating the questions asked and conclusions drawn at each stage of the iterative research process.

2.1 Design experiments

Why use design experiments? A common criticism of educational research is that it fails to translate effectively into improved practice. Some fault the lack of rigorous scientific methods in educational research (National Research Council 2002), while others suggest the problem is research decontextualized from practice (Lagemann & Shulman 1999). Many researchers claim it is impossible and unproductive to attempt controlled experiments in education since the number of variables that may impact on learning in authentic education settings is so great.

Education research is not the only field, however, that attempts to achieve practical successes while isolating the variables essential to that success. Design sciences like aeronautical engineering and artificial intelligence also encounter complex environments with multiple confounding variables. In these areas, practical solutions are not only grounded in theory, they are also iteratively refined to achieve the best outcome within a given context. Contrary to the natural sciences, research in these areas occurs on flexible, pragmatic grounds. It is from these disciplines that the design experiment methodology was adopted for use in educational research.

What are design experiments? According to The Design-Based Research Collective (2003), there are five distinguishing characteristics of educational design experiments:

1. The activities of theory building and learning environment development proceed together, rather than in isolation.
2. Research and design both employ a cyclic structure with planning followed by implementation, evaluation and modification.
3. The theories developed through this procedure must be general enough that they are applicable to other instructional designers and educational researchers.
4. The design experiment must have applicable conclusions for authentic settings.
5. The entire enterprise must ensure the methods can connect implementation to outcomes of interest.

Although most design experiment researchers would agree on the points above, the methodology is conceived of and applied slightly differently by different researchers. Brown (1992), in her seminal work on the topic, recommends that controlled laboratory experiments be used to complement authentic classroom research. Sometimes observations in the classroom, she argues, yield fruitful research questions to be investigated in the more controlled laboratory environment. Other times, laboratory results steer the design and enactment of classroom interventions. Although undoubtedly useful, laboratory experiments are used to varying extents in design experiments by other researchers depending on logistic constraints (Kelly & Lesh 2002). Another feature Brown strongly emphasizes is the collection of rich and varied data. She advocates the use of clinical interviews, video and audio tapes, teacher's notes, and transfer tests, in addition to standard pre and post measures. In practice, different data collection and analysis techniques are employed by different researchers depending on their research questions and circumstances.

In her prototypical design experiment, Brown (1992) set out to investigate if and how young learners could develop and apply specific learning strategies during their lessons. In the laboratory, subjects were trained to use strategic cognitive activities and asked to apply them to the memorization of a list of words. Results showed that even limited training improved performance on the memorization tasks. However the observed benefits did not translate into improved learning outside the laboratory. Once in the classroom, students failed to practice the strategic techniques they had been taught.

In light of this, the research shifted its focus to understanding what essential features must be present in a standard classroom to encourage the development and application of these learning strategies. Based on existing theory, Brown came up with the idea of 'reciprocal teaching', in which reading comprehension is practiced in groups. Members of the group take turns leading the discussion about a text passage, providing an initial question before the reading and a summary statement when the passage has been read. The group rereads the passage as necessary, discussing any discrepancies in interpretation. This method was trialed in the classroom and evaluated through standard tests, teacher notes, laboratory interviews, and audio and visual recordings. Findings were used to both modify the practice of reciprocal teaching and to inform existing theory. Teachers and researchers flexibly made changes in the classroom to investigate which features were essential and to understand the enactment process. The design experiment went through a series of phases in which the reciprocal teaching method was modified, observed, evaluated, and revised. The process took place both in the laboratory and the classroom. At the conclusion of the experiment, researchers claimed both a workable classroom intervention and revised theory as the fruits of their investigation.

2.2 Thesis overview

The question of how one produces effective multimedia for physics education inspired this thesis work and guided subsequent research directions. Reviews of literature on the theory and practice of multimedia and physics education narrowed the area of inquiry. Studies of the local environment, the practices of physics teaching, and the achieved learning outcomes further refined the questions for investigation. The layout of the thesis is shown in Figure 2.1, and discussed in the following sections.

2.2.1 Multimedia in theory

Surveying the history of educational technology research, it seems that different media are incidental to the learning process (Chapter 1). Proponents of new technologies promise to solve a range of educational problems but a strong rationale for their proposed effectiveness is often lacking. Results show that uptake and impact of new technology reach only a fraction of expectations and imply that what is truly important for learning is the experiences of the learner (Clark 1994*a*). This is heavily influenced by the quality of the words and pictures and the method with which they are presented. Questions of method have not been well addressed in multimedia because research has focused on the media itself and perceived learners simply as recipients of knowledge (Mayer 1997).

What are the most relevant theories for multimedia learning?

Some areas of cognitive science have significant implications for multimedia learning (Chapter 3). The study of memory bears directly on theories of learning and it has a long and important history in psychological research. Significant developments in the 1960's introduced the ideas of short-term and long-term memory

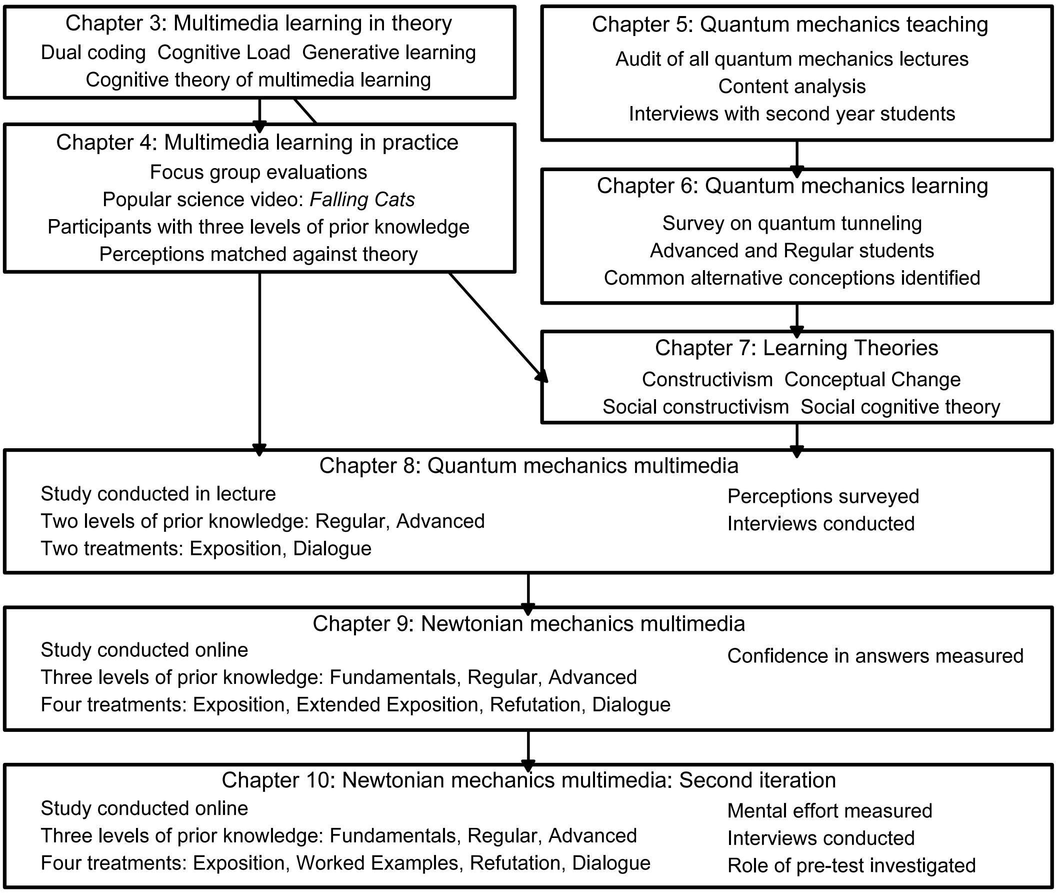


Figure 2.1: Overview of thesis chapters.

(Baddeley 1997). Working memory is limited in its ability to process information, handling around seven chunks of information, actively working with two to four at a time.

Cognitive load theory (Sweller 1988, Chandler & Sweller 1991, Sweller, van Merriënboer & Paas 1998) builds on these findings and is formulated in a way to guide decisions about the design of multimedia instruction. Most frequently, the guidance is to reduce the information presented wherever possible. This may take the form of cutting out extra words, pictures, or sounds; moving text closer to the picture to which it applies; presenting verbal information as narration rather than on-screen text; or eliminating redundant sources of information. New objectives of theorists are to accurately measure cognitive load, to focus on motivational factors influencing learning, and to manage intrinsic cognitive load in complex subject areas (van Merriënboer & Sweller 2005).

Paivio's (1986, 1991) dual coding theory is pertinent to multimedia learning because it suggests that the human mind has two separate processing channels, one for verbal information and the other for non-verbal information. It is supported by experiments that show combined word and picture presentations result in improved recall. The theory is also in line with the two modal subsystems proposed for working memory and studies of brain-damaged patients.

The cognitive theory of multimedia learning (CTML, Mayer 2001, 2005) combines cognitive load and dual coding theories with a view of learners as active participants in the learning process. Like cognitive load theory, the CTML assumes that processing in each channel is restricted by inherent biological constraints. And like dual coding theory, the CTML asserts the formation of relational links between verbal and non-verbal systems is essential for deep understanding. From the generative learning model (Osborne & Wittrock 1983) the CTML draws on a set of processes proposed to occur during learning. This theory now underlies a vast number of studies conducted on multimedia learning in a range of disciplines (Mayer 2005).

Although many of these studies have yielded impressive results and clear principles for multimedia design, they have often been conducted in inauthentic learning settings. Participants have typically been psychology students with little experience in the domains of instruction. They have also been tested in well-controlled laboratory settings. Therefore, it was important for the present investigation to explore how learners in more authentic settings interact with multimedia.

2.2.2 Multimedia in practice

How well does theory account for student perceptions of multimedia in an authentic learning setting?
Do students focus on aspects of multimedia that are not well addressed by theory?
Does prior knowledge affect the way in which students view the multimedia?

To investigate the applicability of multimedia learning theories in authentic contexts, a series of focus groups was conducted on a popular science video, *Falling Cats* (Chapter 4). Every year in lectures this five minute animation is shown to first year physics students at the University of Sydney. The video would definitely be considered a 'craft' solution aimed at getting students interested in physics.

The focus groups aimed to assess any learning that occurred as a result of the intervention and to explore how student perceptions of the multimedia mapped onto relevant learning theories. Student perceptions that did not clearly link onto existing theories provided avenues for further exploration. As prior knowledge was expected to play a significant role, three separate focus groups were conducted with students from low, moderate, and high prior formal physics backgrounds.

Although the study was not conducted in an actual lecture, volunteer students watched the video in groups with similar backgrounds. Two of the three groups were expected to understand the concept of terminal velocity as it was part of their first year curriculum. In this way, the study straddled the laboratory-classroom border.

Transcripts of the focus groups were analyzed with a grounded theory approach (Strauss & Corbin 1998). The data were first broken down in an open coding step in which comments were coded based on content. In the axial coding phase similar codes were drawn together to form major categories. Comments among the different categories were iteratively compared and contrasted to ensure the categories reflected the underlying data. In addition, existing theoretical frameworks helped guide the coding and grouping processes.

Results showed that the video achieved a measure of learning in all focus groups. To varying degrees, students in each focus group could explain the concept of terminal velocity. As might be expected, the groups with lower prior knowledge were more uncertain when discussing the physics and expressed more alternative conceptions. Students in all groups showed heightened interest, asking questions about physics without being prompted.

Student perceptions of aspects of the video were remarkably similar in all focus groups and they appeared to align in many cases with existing theory. Students appreciated, for example, explanations that involved both clear narration and vivid corresponding illustration in line with the dual coding hypothesis. The on-screen pedagogic agents and casual tone were also selected by students as effective teaching devices. These findings correspond both with the personalization principle: learning is enhanced when personal rather than formal language is used (Mayer, Fennell, Farmer & Campbell 2004), and with social agency theory: on-screen pedagogic agents can encourage learning (Moreno, Mayer, Spires & Lester 2001). Cognitive load implications were also apparent, with lower prior knowledge groups perceiving the colours and sounds as distracting.

2.2.3 Quantum mechanics teaching

How is physics currently taught?

An overview of current teaching practices in physics provided a locally relevant perspective on the challenges and opportunities presented by multimedia learning in physics (Chapter 5).

In order to narrow the field of inquiry, the area of quantum mechanics was selected from a vast number of possible alternatives. Quantum mechanics is a subject that until recently had attracted little pedagogic attention. It typifies the complex knowledge domain characteristic of physics. Concepts in quantum mechanics are especially counter-intuitive. They involve phenomena that cannot readily be visualized and often can only be managed through computational approximations. Multimedia is therefore particularly valuable in this area to demonstrate phenomena more directly.

Although the general structure of lecture courses is well known, few independent accounts of physics lecture teaching exist. A survey was therefore undertaken to document all quantum mechanics teaching within the School of Physics. Lectures were broken down into segments of teaching time, and each segment was rated on twelve dimensions pertaining to the teaching methods used and specific predetermined content areas, identified by previous research (Fletcher 2004).

Results showed that lecturers were making extensive use of teaching technologies with PowerPoint used at all levels except honors. In addition, it was clear that in conjunction with mathematical explanations, lecturers were employing imagery to help overcome the abstractness of the subject. A rare occurrence in all lectures however was discussion among students or between students and lecturers. This finding had important implications for the development of the thesis. Specifically, it motivated an investigation into the role of social interactions in the learning of physics. Recent innovations in physics teaching and a literature review of social learning theories revealed a strong emphasis on class discussions.

2.2.4 Quantum mechanics learning

What learning outcomes are students achieving through existing educational practices?

A survey was carried out to assess the conceptual learning outcomes students achieved in the intermediate quantum mechanics class through the documented teaching methods (Chapter 6).

The field of quantum mechanics was too broad to evaluate comprehensively, so quantum tunneling was selected as a problem representative of the domain. Quantum tunneling incorporates a range of important ideas like wave-particle duality, the Heisenberg uncertainty principle, and the wave-function in a macroscopically observable and important phenomenon. Existing research has identified this area as fruitful for pedagogic exploration (Falk 2004, Morgan, Wittmann & Thompson 2004, Fletcher 2004, Domert, Linder & Ingerman 2005).

After completing the second year course on quantum mechanics, which included explicit instruction on quantum tunneling, students were asked to complete a brief paper-based questionnaire on the topic. The survey was conducted informally in the experimental laboratory and participation was voluntary. Students were required to demonstrate their understanding both with diagrams and short written answers.

Student difficulties were readily apparent in the analysis. Conceptions of energy, including diagrams of potential, were commonly misunderstood. Ideas about the wave-function and probability density also seemed to lack coherence. Interestingly, non-scientific conceptions for each question were easily grouped into a small number of categories, a common feature of physics education studies (diSessa 2006). From these categories, possible causes of the alternative conceptions were hypothesized.

2.2.5 Learning theories

What theories underly the learning of physics?

The observations of local teaching practices and measurements of student learning in quantum mechanics prompted a second look at the learning process and constructivism (Chapter 7). Constructivism is an epistemology that emphasizes the active role of the learner in learning and the importance of prior knowledge. This is particularly significant for physics education because in this domain student prior knowledge is often at odds with scientifically accepted conceptions.

The realization that learners may have preconceptions or misconceptions that interfere with learning led researchers to document student ideas before, during, and after instruction. Catalogues of student conceptions were developed for a wide range of topics in physics and other disciplines. Translating this body of knowledge into effective teaching practices, however, was a more substantial challenge.

Researchers developed different theories to explain why alternative conceptions are so resistant to change. Vosniadou (1994) argues that learners possess coherent theories of the world that shape their perceptions of new experiences. Conceptions that contradict existing theories are therefore extremely difficult to understand, especially when understanding requires a change in fundamental world views. DiSessa (1996), on the other hand, suggests that learners don't have structured coherent theories, but rather little bits of knowledge called 'p-prims' that explain everyday experiences. According to diSessa, the challenge of changing conceptions involves organizing p-prims into scientifically acceptable concepts. Chi (1992) discounts the significance of the structured-fragmented nature of conceptions and instead focuses on ontological categories. Conceptions that are miscategorized ontologically, she argues, are the most resistant to change.

Whatever the structure of misconceptions, physics education researchers have tackled the problem in classrooms with many innovative methods. These methods all share a constructivist perspective and aim to actively involve students in hands-on experiments and discussion. Examples are 'Peer Instruction' (Mazur 1997), 'Tutorials in Introductory Physics' (McDermott & Shaffer 2001), 'Interactive Lecture Demonstrations' (Sokoloff & Thornton 1997), and 'Workshop Tutorials' (Sharma, Millar & Seth 1999, Sharma, Mendez & O'Byrne 2005).

The focus on discussion in these methods builds on the theory of social constructivism. This branch of constructivism holds that the interactions among students and between students and tutors are central to the learning process. According to this view, learners engage in meaning-making activities not only individually but in social groups. The learning process is recursive and relational, allowing learners to co-construct meaning with their peers.

Although multimedia cannot act as a competent dialogue partner for students, it can model the discussions students would have in social environments. Observing a learning dialogue might seem like a poor substitute for being an active participant but there are solid theoretical reasons why it may be equally effective or even superior. 'Social cognitive theory' (Bandura 1986), like social constructivism, emphasizes the role of social interaction in the learning process. The theory asserts that an important aspect of this interaction is observation. Observational learning limits the mistakes and faulty effort that accompany learning by trial and error.

Similar to cognitive load theory, it contends that learning in new situations often occurs through a borrowing or mimicking process. This helps to develop mental representations until they are sufficiently formed to attempt the observed skill.

2.2.6 Quantum mechanics multimedia

The next step in the design experiment was to develop and implement a multimedia intervention to address the pedagogic issues raised by the preliminary studies and literature reviews (Chapter 8).

The following specific research questions were formulated:

1. Can the observation of dialogue involving misconceptions be as effective as didactic modes of instruction?
2. Can the alternative conceptions, dialogue, and representation of a student on-screen encourage learners to consider their prior knowledge and reflect upon their learning?
3. Can vicarious learning provide affective benefits, improving self-efficacy or validating students concerns?
4. Do students perceive this strategy as potentially helpful for fostering a question-asking environment in lectures?

To investigate these questions, two multimedia treatments were created on the topic of quantum tunneling. Both treatments used diagrams, animations, equations, and live action video to address the central aspects of tunneling. The multimedia adhered as closely as possible to established multimedia design principles. The significant difference between the two treatments was that one included only correct scientific information while the other featured common misconceptions uncovered with the tunneling survey. Misconceptions were expressed by a modeled student in the multimedia and subsequently discussed by the student with a tutor to reach a scientifically consistent explanation.

After the topic of tunneling had been addressed in lectures, volunteer students from the Regular and Advanced classes were randomly assigned to two lecture rooms. Students were tested before and after viewing a multimedia treatment with a modified version of the tunneling questionnaire.

Results showed that students who watched the misconception-based dialogue outperformed those who viewed the more straight-forward presentation on the post-test. This outcome supported the notion that students could alter their existing conceptions by observing, rather than being direct participants in, a misconception-based learning dialogue. Furthermore, it demonstrated that existing practices that did not explicitly incorporate misconceptions were largely ineffective at promoting conceptual learning.

The study raised additional questions for the use of misconception-based multimedia in physics instruction. For example: would a single speaker presenting and refuting common misconceptions have the same effect on student learning? Would multimedia be as effective at changing misconceptions in domains like Newtonian mechanics where they may be more deeply ingrained?

2.2.7 Newtonian mechanics multimedia

Newtonian mechanics was selected as the subject matter for the second major iteration of the design experiment (Chapter 9). Student misconceptions in this area have been studied for decades and are characterized as internally consistent and robust. Even extensive dialogues with students over the course of an introductory physics course can fail to change the way they think about mechanics (diSessa 1996). Student difficulties are so well researched in these areas that simple tests have been developed and refined to evaluate the state of student conceptions. As a central component of most introductory physics courses, Newtonian mechanics is also studied by a large student population, allowing for a large sample size.

Can conceptual change be facilitated by the observation of a dialogue involving misconceptions even when alternative ideas are firmly held?
Is a misconception-based single speaker presentation equally effective?
Does adding extra material to a concise presentation to slow its pace help or hinder learning?

Four multimedia treatments, called the Exposition, Extended Exposition, Refutation, and Dialogue, were created to explain Newton's First and Second Laws of Motion. The Exposition contained only correct physics information explained with diagrams, graphs, animations, and live action demonstrations. The Extended Exposition contained additional interesting information, beyond the learning outcomes measured on the post-test. The Refutation consisted of the Exposition plus common alternative conceptions stated and refuted. The Dialogue was different in format to the other three treatments in that it was presented by two speakers who took the roles of tutor and student. The student raised the same alternative conceptions as in the Refutation and, through discussion, arrived at scientifically accurate conceptions.

All first year students were asked to participate in the study by accessing a website as part of an assignment. Participants were tested before and after viewing a randomly assigned multimedia treatment using the same multiple-choice pre/post test. The test contained one original question and items from the Force and Motion Conceptual Evaluation (FMCE, Thornton & Sokoloff 1998) and the Force Concept Inventory (FCI, Hestenes, Wells & Swackhamer 1992). Students were also asked to rate their confidence in their answers on a seven-point Likert scale on the pre- and post-tests.

Results showed that students who watched the Refutation or Dialogue achieve significantly greater gains than those who watched the Exposition or Extended Exposition. However, confidence scores improved by a similar amount regardless of which multimedia treatment was watched.

2.2.8 Newtonian mechanics multimedia: Second iteration

The results of the first Newtonian mechanics study suggested that the presentation of alternative conceptions in multimedia could improve student understanding of physics, especially when compared with more traditional approaches. However the study raised several new questions:
Did the inclusion of alternative conceptions raise the cognitive load on students who watched the Dialogue or Refutation multimedia?
If additional useful information were presented in the Exposition to make it equal

in length with the longest treatment, would it be as effective as the misconception-based methods?
What was the effect of the pre-test on learning with different multimedia?
With larger sample sizes, might there be a difference between the Refutation and Dialogue learning for Advanced students?

These questions were addressed by testing separate hypotheses with the three different streams of first year physics students. At the University of Sydney, there are three different streams of first year physics: Fundamentals, for students with no or very little prior formal physics instruction; Regular, for students with high school physics backgrounds; and Advanced, for students with high school physics backgrounds who excelled in a majority of high school subjects.

Fundamentals students followed a similar procedure to the first Newtonian mechanics study, however the Extended Exposition was replaced with a Worked Examples treatment. Instead of including material beyond the learning outcomes, this multimedia contained relevant worked examples, which involved the repetition of key principles and diagrams. Another important difference was students were asked to rate the mental effort they invested during the multimedia on a nine-point Likert scale, to obtain a direct measure of cognitive load.

Regular students were assigned to either the Exposition or the Dialogue. However, students were also randomly assigned to pre-test or no pre-test conditions, to investigate the role of the pre-test. Again mental effort scores were recorded.

Advanced students were assigned to either the Dialogue or Refutation treatment. It was supposed that for novice learners, the presence of alternative conceptions in the Dialogue and Refutation multimedia would be sufficient to promote conceptual change, considering that pre-test scores were initially very low. However for higher prior knowledge students, it was hypothesized that the format of misconception presentation, whether Dialogue or Refutation, might impact on the way in which students viewed the multimedia and therefore the learning that occurred from it.

Fundamentals students provided a replication of the first Newtonian mechanics study. Treatments involving alternative conceptions led to higher post-test scores than the other multimedia. An important novel finding in this study was that students who watched a misconception-based treatment reported investing higher mental effort in the treatment, suggesting the method induced a germane cognitive load.

Considering the results from the pre-tested half of the Regular class in conjunction with those of the Fundamentals revealed that students who watched the Dialogue achieved higher post-test scores than their Exposition counterparts. In addition, mental effort scores were higher for the Dialogue than the Exposition. However, for those students who were not pre-tested, post-test scores were non-significantly higher for Exposition students than Dialogue students, an unexpected finding that requires further investigation.

In the Advanced class, there was no significant difference between the gain scores achieved by students in the Dialogue and Refutation groups. Reported investments of mental effort were also independent of the multimedia treatments.

Interviews were conducted with Fundamentals students to help interpret the quantitative data collected over the two iterations of Newtonian mechanics studies. Students who watched a misconception based treatment were more likely to recall accurately the information presented in the multimedia, including the alternative conceptions. Those who watched the Exposition or Worked Examples treatments felt the material was review and therefore didn't pay full attention to the multimedia.

2.3 Application of the design experiment methodology

In the relatively short history of educational design experiments, the methodology has been applied to a diverse range of questions using disparate methods. It has been characterized as much messier than many other methodologies (Gorard, Roberts & Taylor 2004), yet it is this disorganization that allows for flexibility and customization of inquiry techniques to meet research objectives. This investigation adhered to the central tenets of design experiments, but details of the methods applied were moulded to suit the research questions and local circumstances.

As with all design experiments, this investigation was rooted in theory. The most applicable theories for multimedia learning were used as starting points for the development of interventions and research questions (Chapter 3). As the study developed, new theories were incorporated into the theoretical base. For example, social learning theories became important after preliminary investigations into authentic multimedia, current teaching practices, and learning outcomes were completed (Chapter 7).

A key outcome of the study was the elaboration and modification of theory. Theories like the cognitive theory of multimedia learning were investigated for their applicability in authentic contexts. The division in cognitive load theory between extraneous and germane cognitive load was explored in relation to misconceptions. Developments in theory proceeded in parallel with developments of multimedia interventions. In fact, the development of multimedia theory on the topic of conceptual change was a primary aim and outcome of the study.

The relative contributions of laboratory and authentic classroom research are variable in design experiments, but relevance for classroom practice is always essential. This was achieved through observations of physics lectures, presentation of multimedia in lecture environments, and surveying of existing syllabus content. Student participants were enrolled in courses that addressed similar material to that presented in multimedia treatments. Experiments were also conducted online, in participants' own time, for course credit. The balance with laboratory research was established by random allocation of participants to multimedia treatments and follow-up interviews and focus groups.

The developed interventions were used in authentic contexts after their completion. For example, the quantum tunneling Dialogue video that was tested in 2005 was used in lectures in 2006 and is now an available lecture demonstration. Multimedia developed as part of the grant I have worked on also employ the methods developed in this design experiment.

Overall, the multimedia materials developed during the course of this research reflect the central principle of design experiments:

> Interventions embody specific theoretical claims about teaching and learning, and reflect a commitment to understanding the relationships among theory, designed artifacts, annd practice. At the same time, re-

> search on specific interventions can contribute to theories of learning and teaching. (The Design-Based Research Collective 2003, p.6).

How does this study differ from other design experiments? Many design experiments focus on long-term interventions in classrooms. This study, although extending over a period of years investigated interventions from seven to fifteen minutes in length. Many design experiments also place less importance on standardized conceptual measures. These featured heavily, however, in this study.

Chapter 3

Multimedia in theory

Despite the growing availability of multimedia and some students' and teachers' preference for it, it is pertinent to ask what theoretical advantages multimedia affords. What is the rationale behind learning with multimedia? Of what limitations must instructional designers be aware?

In this chapter I first consider the limitations on human processing due to the structure of human cognitive architecture. Then I outline the theoretical supports of learning from multi-modal instruction. Constructivism and the generative learning model are further considerations that highlight the active role of the learner in meaning-making. This set of theories has been combined into a general 'cognitive theory of multimedia learning' (CTML, Mayer 2001, Mayer 2005) and investigated in numerous multimedia experiments.

3.1 The structure of human memory

A simplified model of memory is shown in Figure 3.1. The schematic shows the ways in which new sensory information can become stored in long-term memory, through attention and rehearsal, or forgotten through decay, displacement and interference. The model for short-term memory has been elaborated to include three subsystems: a phonological loop for verbal and auditory processes; a visio-spatial sketchpad for non-verbal processes; and an episodic buffer, for the integration of

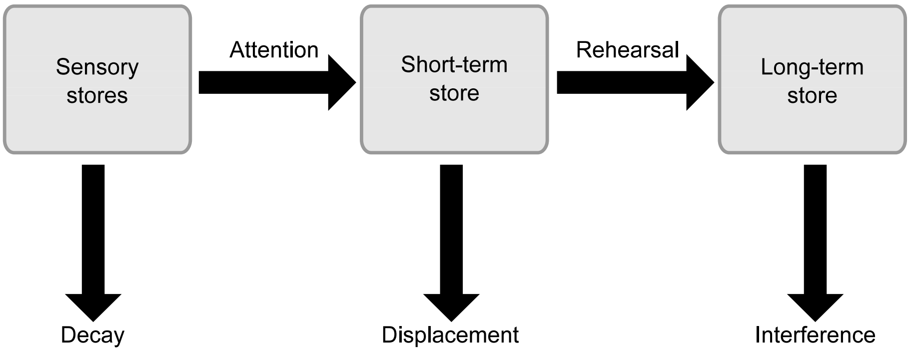


Figure 3.1: Schematic of the multi-store memory model, adapted from Eysenck & Keane (2005).

information from long-term memory and the other two subsystems. Together with a central executive, these components are known as working memory.

Cognitive psychologists first conceived of two distinct types of memory in the 1960's after behaviorism was supplanted by cognitivism as the dominant ideology. Baddeley (1997) provides a summary of the studies that led to the established two-component framework. Proponents of the working memory model believed it elegantly explained the phenomenon of rapid forgetfulness when participants were prevented from rehearsing recently acquired information. Opponents suggested that forgetting was not due to a memory decay but to interference (Melton 1963). Strong evidence for the model of working memory came from studies of patients who had suffered significant brain trauma. Case studies showed that individual patients with lesions in different areas of the brain had either unimpaired short-term or long-term memory. For example, one individual could repeat back phone numbers immediately after they were presented but could not remember if he had met the researcher before. The other could repeat back only a few digits but displayed normal long-term learning.

3.2 Learning without being overwhelmed

Humans do not have an unlimited capacity to assimilate new information. Over the past half-century, cognitive scientists have attempted to understand the limitations of human cognition. Cognitive load theory (CLT) is one outcome of these efforts, appropriately formulated for instructional design.

The theory is based on a set of simple, empirically verified assumptions (Pollock, Chandler & Sweller 2002, p.62):

1. Humans possess a limited working memory, able to process around seven bits of information at a time.
2. These limitations are mitigated by interaction with a virtually limitless long-term memory.
3. Schemas stored in long-term memory structure information into 'chunks,' in effect increasing the amount of information that can be processed in working memory.
4. Repeated processes become automated over time, reducing the load on working memory.

Cognitive load theory has recently been formulated in analogy to biological evolution (Sweller 2004, van Merriënboer & Sweller 2005). Both evolution and learning, it is argued, are similar information processing systems. Each organism has, as its base, an incredible amount of genetic information stored in a molecule of DNA. This genetic information is expressed in an organism that has met the challenges of its environment up to a particular point in time. In order to survive in a changing environment, a species must have a mechanism for adaptation. This mechanism is the random mutation of a tiny fraction of the DNA molecule and the trial and error process that ensues in the interaction between the organism and the environment. If a slight genetic change adapts the organism to the environment without reducing its ability to overcome previously surmountable obstacles, then the mutation becomes established in more and more individuals in the population with each successive generation. If, however, the adaptation is detrimental, it will be repressed in future populations.

In this way, evolution can be seen as a trial and error process that is enabled by tiny random mutations. If the mutations were too great in magnitude, maladaptive changes would be common, endangering the future of the species.

Pursuing this analogy with respect to cognition, the equivalent of the DNA molecule is the large amount of information stored in long-term memory. It is this knowledge that has allowed a person to cope with his or her environment up to a given point in time. Again, in a capricious environment, there must be a mechanism for adaptation. In cognition, this is provided by working memory. Small changes to long term memory can be attempted through conscious thought processes. If these alterations prove fruitful, they will be rehearsed and so further cemented into long-term memory. If the ideas are maladaptive, on the other hand, they will be superseded by structures that better allow a person to function. As in biological evolution, drastic changes to long-term memory are unlikely to be adaptive. This may explain in part the limitations on working memory.

In a classic paper, Miller (1956) summarized findings quantifying humans' limited capacity for processing information in working memory. In many different cognitive tasks, the number of entities that could be held in memory was seven (plus or minus two). Cognitive load theory has built on this conclusion, contending that the fundamental concern of instructional design should be the allocation of cognitive resources within this limit.

Cognitive limitations can be alleviated with the use of structures stored in long-term memory called schemas. This term was first used by Bartlett (1932) to explain why people remember certain parts of stories and forget or distort other parts. Although the term is used for a range of different concepts, according to Bartlett the word schema refers to

> an active organization of past reactions, or experiences, which must always be supposed to be operating in any well-adapted organic response. That is, whenever there is an order or regularity to behaviours, a particular response is possible only because it is related to other similar responses, which have been serially organized, yet which operate not

as individual members coming one after another but as a unitary mass. (p.201)

Three types of cognitive load, called intrinsic, extraneous, and germane, are proposed to break down the expenditure of cognitive effort. Intrinsic cognitive load is a measure of the inherent difficulty of a subject area due to the number of interacting bits of information involved. Learning vocabulary in a foreign language is a standard example of a task that demands low intrinsic load. Such a task involves only one unknown quantity, a foreign word, which is mapped directly onto a known word. Learning a foreign grammar, on the other hand, involves higher intrinsic load because it requires the consideration of several unfamiliar words as an ensemble. Physics generally involves a high cognitive load as natural processes and the equations that describe them often involve multiple interacting entities. Furthermore, physics is cummulative in that fundamental material is repeatedly drawn upon and augmented in the higher years of physics teaching. Extraneous cognitive load refers to invested mental effort that does not result in learning. Searching for information in a poorly laid out document, for example, would constitute extraneous load. Germane cognitive load refers to the mental effort used to form schemas and actively integrate new information with prior knowledge. The effort required to execute learning strategies like self-explaining (Chi, Bassok, Lewis, Reimann & Glaser 1989, Chi, Slotta & De Leeuw 1994), analogical reasoning (Holyoak, 2005), and some aspects of metacognition would fall into this category. Finding ways to increase germane load and minimize extraneous load has been a central pursuit of researchers under this paradigm.

Early studies on cognitive load evaluated the efficacy of solving novel problems as a method of acquiring problem solving expertise (Sweller 1988). It was suggested that the processes required to solve novel problems were actually fairly independent of the schema building activities required for learning. Learners working within problem solving frameworks focus on the information given and the answer required, executing a ‘means-ends analysis.’ This process involves reducing the differences between the present problem state and the goal state with each sequential step through permissible operations. Often students using this strategy work backwards from the answer before working forwards from the beginning. This contrasts with the method of experts, who classify a problem first based on their extensive schemata and then work forwards towards the solution.

Studies showed that when means-ends strategies were used, little comprehension of underlying problem structure was achieved (Sweller, Mawer & Howe 1983). In contrast, when learners were prevented from using means-ends analysis, they quickly learned essential problem characteristics (Sweller & Levine 1982). Furthermore, learners who calculated all possible variables committed less errors on a secondary task than did learners who were given a specific goal. Sweller (1988) argues the substantial cognitive load demands of the means-ends analysis limit the cognitive effort that can be devoted to schema acquisition.

To reduce the cognitive load related to attaining problem solving expertise, Sweller (1988) advocates a number of changes to problem solving exercises. Asking learners to solve for as many unknowns as they can instead of one particular unknown eliminates the means-ends analysis freeing up cognitive resources for schema building. Providing worked examples (Sweller & Cooper 1985) or partially worked examples also reduces the load on working memory. These techniques have been demonstrated to improve learning and problem solving performance.

From reducing the cognitive load involved in problem solving, researchers applied the theory to standard text and diagrammatic instructional materials (Chandler & Sweller 1991). In conventional learning resources, it is common for mutually referring text and diagrams to occur in different locations on a page. When these sources of information cannot be understood independently of each other, they place substantial demands on working memory. To reduce cognitive load, text can be closely integrated with diagrams, increasing the cognitive resources available for schema building. In addition, seemingly useful but non-essential text was found to inhibit learning even when presented in an integrated format.

Prominent cognitive load effects relevant to instrucion are listed below (Sweller et al. 1998).

- Goal-free effect

Learning is enhanced when learners are presented with non-specific goals, eliminating extraneous cognitive load caused by means-ends analysis.

- Worked example effect

  Learning is enhanced when learners carefully study worked examples rather than attempt the problems themselves.[1]

- Completion problem effect

  Learning is enhanced when learners complete a partial solution rather than attempt the problem themselves.[1]

- Split attention effect

  Learning is enhanced when different sources of information are integrated, reducing the extraneous load involved in mental integration.

- Modality effect

  Learning is enhanced when verbal material is presented orally rather than as text when accompanying visual material.

- Redundancy effect

  Learning is enhanced when multiple sources of redundant information are condensed into one.

Recently, cognitive load theorists have broadened the scope of their investigations, evaluating methods for managing the intrinsic cognitive load of instruction, something previously thought to be beyond the influence of instructional design (Sweller et al. 1998). Pollock et al. (2002) found that complex material was better presented in two phases. In the first phase, only isolated elements were presented allowing learners to process the information sequentially rather than simultaneously. In the second phase, learners put these pieces together after being presented with a complete description involving all interacting components.

In addition to managing intrinsic cognitive load, researchers are aiming to understand the role of motivation and the development of expertise that are involved

[1]These effects are particularly relevant for students with limited prior knowledge. They are less important for advanced learners.

in authentic settings over longer periods. Another prominent direction for cognitive load theory is in the measurement of cognitive load and in the development of adaptive systems.

3.2.1 Measuring cognitive load

Cognitive load can be measured in a variety of ways, using rating scales, psychophysiological measures, or secondary task techniques (Paas et al., 2003). Rating scales assume that learners can accurately judge and report the mental effort they expend during a learning experience. Psychophysiological measures are based on the principle that mental activity predictably influences physiology. Using this reasoning, researchers have gauged small changes in heart rate, brain activity, and eye activity to measure cognitive load (e.g. Van Gerven, Paas, Van Merriënboer & Schmidt 2004, Paas, Van Merriënboer & Adam 1994). Researchers using secondary task techniques require learners to perform a task concurrently with the primary learning activity. Such a task might involve detecting a visual or audio cue. Reaction times and error rates are then used to infer the cognitive load of the primary task. Self-reported rating scales might seem like the least reliable cognitive load measures, but they have shown better sensitivity and reliability than some physiological methods (Paas et al. 1994). In addition, they are non-invasive, easy to administer, and do not interfere with the cognitive load of the learning task, a common concern with secondary task techniques.

The most common criticisms of cognitive load theory are that the three proposed types of cognitive load depend on the learner and that they cannot be measured independently of one another. To overcome the restrictions on working memory, humans use prior knowledge, in the form of schemas, to interpret incoming information. Because each individual possesses a unique set of schemas, the intrinsic cognitive load he or she experiences during a given lesson is also unique (Cook, 2006). Experts, for instance, perceive information that pertains to their area of expertise often differently from novices, usually in much larger (meaningful) 'chunks' (Gobet, 2005).

Something similar can be said of extraneous and germane cognitive load. The variability of the types of cognitive load becomes problematic when one attempts to estimate their relative abundance during a particular learning episode. Total cognitive load can be measured during instruction and during the assessment following instruction, as outlined above; however, it is difficult to reliably estimate the proportions of different types of cognitive load that make up the total. This problem can be somewhat overcome by comparing instructional treatments in the same subject area that differ only in one specific way. If post-test performance and cognitive load of treatment B are greater than those of treatment A, then the method in B increased germane cognitive load. If the cognitive load is greater in treatment B but the post-test performance is lower, then an extraneous cognitive load was induced.

3.3 Learning from words and pictures

Although the addition of pictures to instruction is acknowledged to improve aesthetics and sometimes student motivation (Rieber 1994), there is theoretical support for the assertion that images serve an important cognitive function in learning. Dual coding theory (DCT, Paivio 1971, Paivio 1986, Paivio 1991) is the dominant theory that addresses the role of imagery in human cognition. A diagram of the model is shown in Figure 3.2.

The theory proposes that "there are two classes of phenomena handled cognitively by separate subsystems, one specialized for the representation and processing of information concerning nonverbal objects and events, the other specialized for dealing with language" (Paivio 1986, p.53). Although the nonverbal system is concerned with all experiences independent of language, it is commonly referred to as the visual or 'imaginal' system. Details of these two subsystems, or channels, have been developed through a systematic series of experiments.

In the dual coding model, the verbal and visual channels possess distinct representational units called 'logogens' and 'imagens,' respectively. Logogens are linked together through associative connections, building a hierarchical structure within the verbal channel. Similarly, imagens relate to each other through associative links

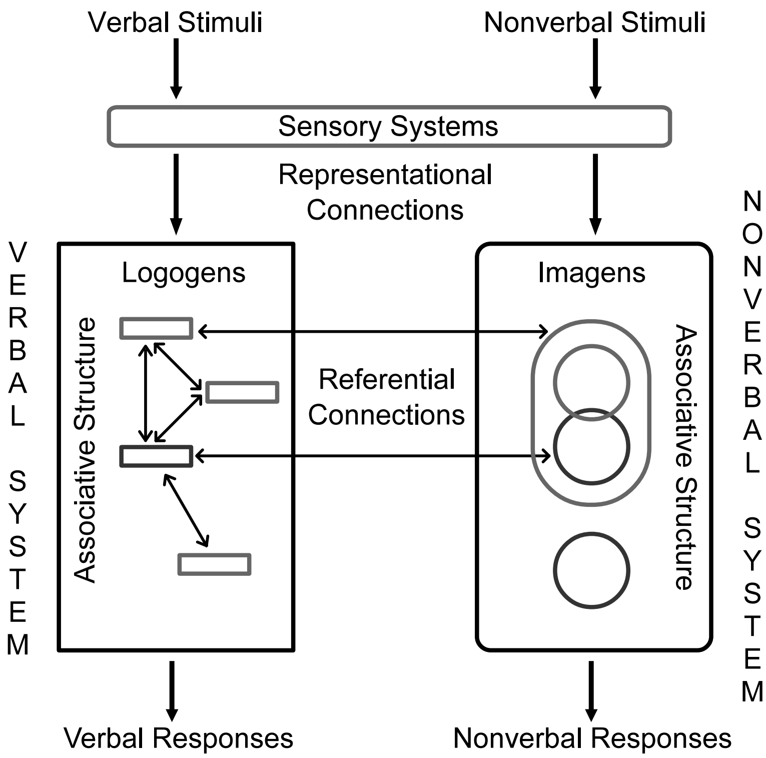


Figure 3.2: Schematic of dual coding theory adapted from Paivio (1986).

in the non-verbal system. Thus processing related to associative links can occur in one channel independent of the other. An important difference between the two systems is that processes are said to be sequential in the verbal channel and synchronous or parallel in the nonverbal channel. For example, recalling the layout of one's office is a simple task with items to the left or right of one's desk easily accessible. In contrast, recalling a line from a speech or the national anthem typically requires going through words in sequence to reach the target phrase (Rieber 1994).

Activity in one system can also trigger activity in the other system in what is called referential processing. The word 'car,' for example, might bring to mind the image of a car. It has been found experimentally that images are much more likely to trigger referential processing and therefore be coded in the verbal system than vice versa. This leads to what some have called the 'picture superiority effect,' (Nelson, Reed & Walling 1976). "The evidence suggests that imaginal and verbal codes are unequal in mnemonic value, perhaps by a 2:1 ratio favoring the image code" (Paivio 1986, p.77).

This theory has important implications for the design of multimedia instruction. Memories coded in two channels are more likely to be recalled than those that exist only in one. "The evidence to date suggests that imaginal and verbal codes are functionally independent in the strong sense that activation of both can have additive effects on recall" (Paivio 1986, p.77). In addition, the nonverbal system allows memories to be stored and processed synchronously, rather than sequentially as in the verbal system.

The theory was inspired by mnemonic techniques like those used to memorize a numbered list of objects. Paivio (1986) found that images were especially well suited as 'conceptual pegs' (i.e. retrieval cues) for other items. Subjects attempting to memorize pairs of adjectives and nouns demonstrated better recall when the noun was presented first and when it was concrete rather than abstract. Further studies dispelled a competing hypothesis, that the meaningfulness of nouns increased their efficacy as mnemonic aids. In fact, meaningfulness was found to have a small or even detrimental effect on recall. Qualitative data from participants confirmed that they used imagery to perform the recall tasks. Later studies uncovered that pictures were more easily recalled than abstract or even concrete nouns, supporting the central role of imagery in memory.

Neuropsychological studies provide further support for dual coding theory. It has been known for some time that the left and right hemispheres of the brain are asymmetric in their functions (Hellige 1993). Although some believe this asymmetry has been exaggerated and distorted in popularized accounts (Efron 1990), different parts of the brain are implicated in different cognitive tasks. The left hemisphere seems to play an important role in speech, while the right is more proficient in select nonverbal tasks. Abstract words are better recognized by the left-hemisphere, while concrete words are recognized equally well by both sides of the brain.

The most common criticism of dual coding theory is that there is no need to propose two distinct representational systems when one would do (Baddeley 1997, Paivio 1986). Critics argue that all stimuli, verbal and non-verbal, are processed in working memory and converted to abstract propositions (or 'mentalese') before being stored in long-term memory. Proponents of this theory do not deny the existence of images in working memory, they simply suggest these are constructions made from abstract propositions stored in long-term memory. They maintain that people naturally rehearse images more readily than words leading to the superiority of images in recall. Evidence for the propositional storage of images comes from subjects' inabilities to correctly recall details of commonly viewed items like coins (Baddeley 1997).

The results of decades of research indicate that images have a unique and beneficial impact on learning and memory. Whether one accepts dual coding theory or the propositional hypothesis, the advantage of images in instruction is unchanged. Due to extensive empirical support for dual coding theory and its straightforward model, it is particularly applicable to the theory of multimedia learning. Mayer (2001) conducted a series of experiments in which students received either multimedia or verbal only explanations of pumps, brakes, generators, and thunder storms. In six out of nine studies, students who received the multimedia presentation demonstrated better recall than their single-mode counterparts and in all nine experiments the multimedia group outperformed the single-mode group on transfer tests (i.e. tests

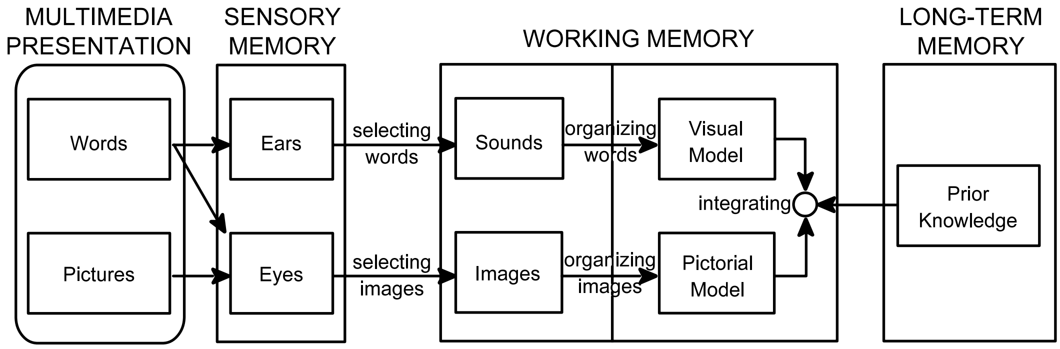


Figure 3.3: Schematic of the cognitive theory of multimedia learning adapted from Mayer (2001).

that required students to use the learned principles in novel contexts).

3.3.1 The cognitive theory of multimedia learning

The cognitive theory of multimedia learning (CTML, Mayer 2001, 2005) is composed of a combination of dual coding, cognitive load, and constructivist theories, shown diagrammatically in Figure 3.3. It builds upon previous research on learning with technologies (e.g. Salomon 1994, Hegarty, Carpenter & Just 1991, Taylor 1980).

From dual coding theory, the CTML employs the idea that people have two independent but related channels for processing verbal and non-verbal information. The amount of processing that can occur in each channel is limited by cognitive load theory. By using two channels rather than one for instruction, the theory asserts that learning can occur more effectively and efficiently with multimedia. The CTML also identifies a set of active processes that a learner must undertake in order to learn. From the stimulus material a learner must: select relevant words and relevant pictures; organize these words and pictures separately into coherent mental models; form relational links between verbal and mental models; and integrate these models with prior knowledge. This set of tasks is a simplification of the cyclic, iterative mental processes proposed by Osborne & Wittrock (1983) in the generative learning model. As shown in Figure 3.4, a key aspect of this model is that long-term memory

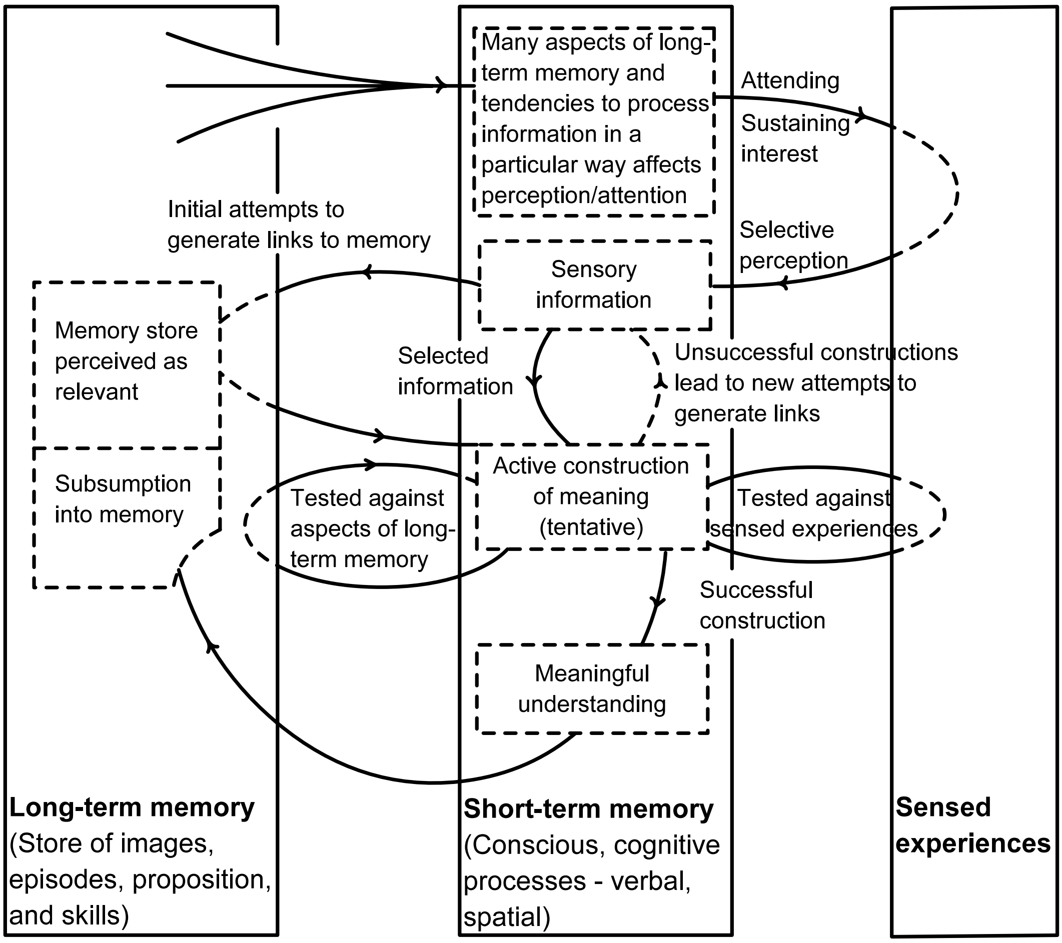


Figure 3.4: Schematic representation of the generative learning model, adapted from Osborne & Wittrock (1983).

plays a central role in perception and attention. Processes in working memory loop internally and in relation to sensed experiences and long-term memory in order to achieve meaningful learning. This is one of the most complex and clearly articulated constructivist models of learning.

Using this theoretical framework, a set of multimedia design principles has been established and empirically verified (Mayer 2001). Some significant examples are summarized below:

- Multimedia principle

Learning is enhanced when instruction is presented as words and images rather than words alone.

- Spatial contiguity principle

  Learning is enhanced when corresponding words and images are presented in close proximity.

- Temporal contiguity principle

  Learning is enhanced when corresponding words and images are presented simultaneously.

- Coherence principle

  Learning is enhanced when material extraneous to the learning outcomes is excluded.

- Modality principle

  Learning is enhanced when words are presented as narration rather than as on-screen text.

- Redundancy principle

  Learning is enhanced when narration is not duplicated as on-screen text when competing with dynamic visuals.

This list has now been extended to include multimedia design principles relating to navigation, collaboration, and diverse discipline areas (Mayer 2005).

The experiments conducted to establish these principles usually follow similar methodologies. Two multimedia treatments that differ only in the variable to be investigated are created. Participants are often psychology students with little formal experience in the domain of instruction. In a computer laboratory, they are randomly assigned to a multimedia treatment. The sample size is typically in the range of 20–30 participants per treatment. After the multimedia, participants complete a post-test that consists of short-answer retention and transfer questions. Retention questions require students only to recall information directly presented in the multimedia. Transfer questions, on the other hand, require students to apply material they have learned to new situations. These two different types of questions are considered separately in analysis.

This chapter provides a theoretical foundation for learning with multimedia. Most of the models discussed above conceptualize learning as something that happens exclusively in the brain. External representations, indexical knowledge and social resources are less significant in these views of learning, though it should be noted that other perspectives give them considerable attention (e.g. Clark 1997). Constructivism and related learning theories are discussed in Chapter 7.

Chapter 4

Multimedia in practice

As described in Chapter 3, many principles of multimedia design have recently been proposed and empirically verified. Often, however, the supporting experiments have been conducted in well-controlled laboratories. Participants have been psychology students with little domain knowledge, for whom the relevance of instruction has therefore been limited. For most of these principles, it remains an open question if or how they generalize to real settings with learners who have both domain knowledge and an interest in the subject matter.

Tabbers, Martens & van Merriënboer (2004) reported a study in which they investigated the modality and cueing effects in an authentic learning setting. The modality effect implies that verbal information should be presented as narration rather than on-screen text to facilitate learning (Sweller et al. 1998). The cueing effect suggests that learning is enhanced when visual cues in an animation help link images to their associated narration (Kalyuga 1999). Although these effects had been well established in previous studies, Tabbers et al. failed to replicate the results in a classroom environment. Their study found only a slight cueing effect and even a reverse modality effect. This suggests that the translation of multimedia principles into effective practice is not trivial.

Another body of knowledge exists on the design and use of multimedia systems and this is the existing multimedia itself. This knowledge has the added legitimacy that comes with routine classroom use, but the drawback that it must be disentangled from the artistry in which it is imbedded. According to Clark & Estes (Clark & Estes 1998, 1999,Estes & Clark 1999), the majority of educational technologies can be categorized as 'craft'—"limited, contextualized, non-transferable . . . solutions to educational problems" (p.5). However, successful 'craft' solutions, it is argued, often contain the most advanced information available on a topic.

To move from 'craft' technologies to more scientifically grounded 'authentic educational technologies,' Clark & Estes emphasize the need for evaluation, in the form of Kirkpatrick's (1994) four principles. In brief, these are: (1) reaction, gauging participants' perceptions of an intervention; (2) learning, measuring changes in participant knowledge, skills, or attitudes; (3) behaviour, ensuring learning is put into practice; and (4) results, the bottom line outcomes of training, most applicable in business settings. Clark & Estes suggest this evaluation must be used both to investigate successful craft technologies and to validate future multimedia interventions.

In this chapter I describe an evaluation of a multimedia intervention currently used in first year physics lectures. Goals of this study were: (1) to assess student learning with multimedia in a semi-authentic setting; (2) to explore how student perceptions of multimedia map onto established multimedia design principles; and (3) to uncover potential avenues for future research, where student opinions diverge from existing research. In this investigation, only the first two levels of Kirkpatrick's evaluation were conducted because the latter two were beyond the scope of the research objectives.

4.1 Method

Three focus groups of university students were assembled on the basis of prior experience and overall interest in physics into low ($n = 5$), moderate ($n = 8$), and high ($n = 18$) level groups. Students who self-reported having a lower interest in physics had less prior physics instruction and vice versa. Interest in physics was self-reported on a five point scale from very low (1) to very high (5). Students in the low, moderate, and high knowledge groups averaged scores of 2.2, 3.3, and

4.3 respectively. Students in the moderate and high level groups were first year students who had taken physics in first semester. At the University of Sydney, there are three different streams of first year physics: Fundamentals, for students with no or very little prior formal physics instruction; Regular, for students with high school physics backgrounds; and Advanced, for students with high school physics backgrounds who excelled in a majority of high school subjects. The low level group consisted of students from various university disciplines and levels with little prior physics instruction besides compulsory high school science.

The focus group characteristics are summarised in Table 4.1, below. A similar number of students was expected to participate in each group. However due to the nature of the study, students with greater interest and experience in physics volunteered more readily. Participants were all between the ages of 17-27 and were enrolled in undergraduate studies at the University of Sydney.

Prior knowledge	Low	Moderate	High
Sample size (n)	5	8	18
Experience in physics	High school	Fundamentals	Advanced
Average self-reported interest (SEM)	2.2 (.6)	3.3 (.3)	4.3 (.2)

Table 4.1: Summary of focus group attributes.

The qualitative method of focus group research was selected because it allowed for a rich and detailed exploration of student perceptions of the animation. During the discussion, students built on the ideas of others. The opinions of individual participants were examined by other members of the group in the process of reaching a consensus. These features make focus groups ideal tools for commercial evaluations of products (Greenbaum 1998). The focus group setting also allowed students to watch the video in an environment similar to a lecture.

Students were shown a five minute animated video by popular scientist, Dr. Karl Kruszelnicki, on terminal velocity. Every year this multimedia is shown to first year physics students during lectures. The film outlines the basics of termi-

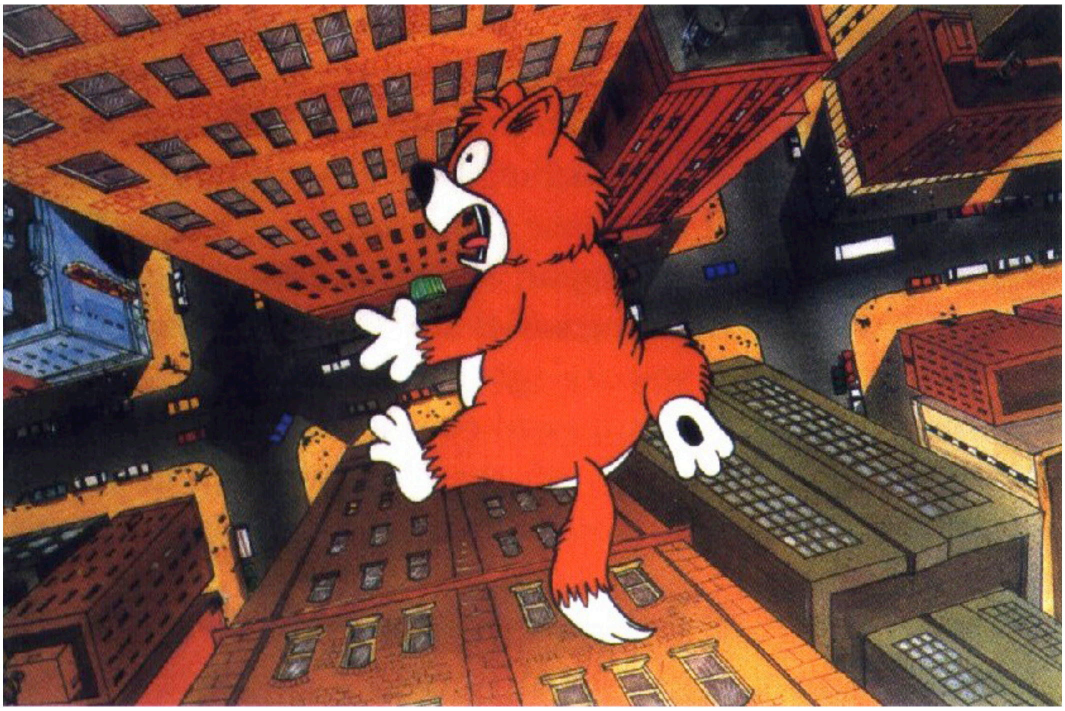


Figure 4.1: Still frame from the animation *Falling Cats*.

nal velocity in answering the question: 'why is it sometimes safer for cats to fall from taller buildings than shorter ones?' The film is based on data collected by New York City veterinarians on the injuries sustained by cats after accidental falls (Diamond 1988, Diamond 1989). Narrated by an animated Dr. Karl and starring an expressive orange tabby cat (Figure 4.1), the film explores the physics of terminal velocity. Concepts like gravity and wind resistance are clearly illustrated to explain why falling bodies have a speed limit. The film has been shown on TV and has achieved international recognition (Embassy of France in Australia 1997).

All of the focus groups lasted approximately one hour. This included watching the video, independently filling out a short questionnaire (Appendix B.1), and discussing the results. The questionnaire contained six broad, opinion-based questions about the video and one question that asked for a definition of terminal velocity. The goal of the questionnaire was to get participants thinking about the issues to be discussed in the focus group and to obtain independent opinions of the video prior to discussion.

Determining what students learned from the multimedia was a complex task. First year lecturers discussed their expectations for the film in casual conversations before the study. They believed showing the video promoted interest in the subject matter and captured the attention of students, carrying over to the remainder of the lecture. This view has previously been expressed with regards to the effect of interesting information on learning (Mitchell 1993, Schraw & Lehman 2001). Lecturers also felt watching the video was genuinely enjoyable; an important aspect of higher education that they thought was often overlooked. These views around teaching with video are common among tertiary lecturers (Oliver, Grant & Younger 1994, p.29).

With lecturer objectives in mind, learning was evaluated using a framework established for science communication. Learning was assessed along the dimensions of Awareness, Enjoyment, Interest, Opinion forming, and Understanding (the AEIOU analogy) (Burns, O'Conner & Stocklmayer 2003). The depth of student understanding and the willingness to discuss conceptions about terminal velocity were examined as the most significant indicators of student learning. Willingness to speak about the physics was taken as a demonstration of heightened awareness and interest, whereas depth of understanding was evident in the breadth of accurate conceptions discussed, the forming of opinions, and in the confidence with which physics issues were raised.

4.2 Results and analysis

Discussions from the focus groups were transcribed and coded using open and axial coding as described by Strauss & Corbin (1998). Although this grounded theory technique is typically used for theory generation, the selective coding and theory generation steps were omitted as they were beyond the scope of this study. In open coding, all statements in the transcripts were broken down into components using 'in situ' codes—phrases taken verbatim from the data, for example 'humour was efficient'. In the axial coding step, the data broken down in open coding were synthesised by combining relevant data into categories (e.g. 'humour was efficient' was combined with 'humorous expressions made learning enjoyable'). A single comment was included in all categories to which it pertained. These groupings were then arranged into higher order categories (e.g. 'humour for learning'), and iterative comparisons between the data and higher categories were made to ensure fidelity to the observations. Where quotes are given below to support generalisations from the data, participants are referred to by an arbitrary code. Codes beginning with L, M, and H refer to participants from the low, moderate, and high level groups respectively. The letter 'I' is used for the investigator.

4.2.1 Learning

Discussion of physics

One feature common to all three groups was participants' willingness to spontaneously ask questions regarding terminal velocity. This showed that the video got students thinking critically about the physics. For example, in the low interest group, one student expressed her confusion without being prompted: (L1) "I would have thought that it would be, um, different numbers actually. Because a cat is not, like, half the weight of a human. Why would it have a terminal velocity that is half that of a human?" Despite her uncertainty, this student was willing to explore the conflict in her mental model raised by the numbers from the video. While discussing this among the group, participants attempted to explore the factors affecting terminal velocity but were unsure of their suggestions.

L2	Well, why are we saying that it should be directly proportional to, um, weight? Presumably, it's a curve, not a straight line.
I	Why wouldn't the curve be directly proportional to weight? Anyone?
L1	Depends on how aerodynamically shaped your object is.
I	Anyone else? Do you like that idea of L1? [laughter]
L3	We don't know.

Despite the lack of confidence characteristic of this group, physics issues were raised and discussed indicating increased interest and awareness (I and A from the AIEOU analogy). In the moderate and high interest groups, students were more confident with their questions and actively responded to each other. Again, they brought up issues relating to terminal velocity without being asked specifically about the physics content of the video.

Range of physics concepts

In order to examine the depth of understanding qualitatively, the method of small scale quantitative analysis was used Greenbaum (1998). The number of statements from each group pertaining to terminal velocity were compared. These statements were made in explaining why the cat's terminal velocity was half that of a human, not in discussing factors affecting terminal velocity in general. Statements were categorized as confident or unconfident depending on the tone in which they were made. For example, if the student phrased the remark as a question, it was considered unconfident. Statements were also classified as correct or alternative. The results are summarised in Table 4.2.

The table clearly shows that the higher interest groups discussed more correct conceptions more confidently, indicating a deeper understanding of terminal velocity. There was also a distinct separation between the high and moderate interest groups. The moderate interest group was comfortable with the physics ideas

Statements	Correct			Alternative		
Focus Group	Low	Moderate	High	Low	Moderate	High
Confident	3	7	8	1	2	1
Unconfident	4	1	0	1	1	0

Table 4.2: Number of statements relating to terminal velocity, categorized by group, correctness, and confidence

and spoke confidently even about alternative conceptions. They readily jumped to conclusions, often suggesting mathematical models before considering the physical factors involved. They also spent more time discussing alternative conceptions. This shows less critical thinking and a shallower understanding than the high interest group, which did not seek to explain the numbers but rather the factors affecting terminal velocity.

The difference in understanding between the moderate and high level groups is highlighted in the following exchange.

I	That the cat falls at 100 km/hr and humans fall at a terminal velocity of 200 km/hr. What do you make of that?
H1	That we're denser, so we have less air resistance.
H2	We don't have fur all over our skin.
H3	Maybe just yeah the fur and like the shape of the cat.
H4	We don't actually do this when we fall [makes a spreading out motion]. Cats do that.

These students were confident in providing qualitative answers and did not resort to mathematical explanations like the other two groups did. This demonstrates both comfort with the subject matter and an implicit understanding that the factors affecting terminal velocity are so numerous and complex that the question did not require a mathematical explanation.

Opinion forming

Another measure of depth of understanding was the ability of participants to form opinions about the physics. The nature of the *Falling Cats* story allowed students to construct opinions as evidenced in the moderate and high level focus groups. For example, in the moderate interest group, a student, M1, asked "Why would a cat relax more the longer it falls?" The video gave a simple explanation which M2 reiterated, "Because it's realized that it's not accelerating anymore. Did you listen to the video?" M1 takes the explanation a step further by introducing his own opinion. "I don't think it realizes it is not accelerating, it realizes—it thinks it's not falling anymore." M1 thinks it's unlikely that the cat truly realizes its state of acceleration; instead, he supposes the state of constant velocity is inconsistent with the cat's experience of falling. He views the problem from the cat's perspective, believing the inconsistency results in the cat thinking that it is no longer falling, and that is what causes it to relax. This demonstrates that this student understands the concepts to a depth where he can apply his knowledge in forming coherent opinions that extend beyond those expressed in the video. The video gave students a wider range of ideas on which to base their opinions due to the grounding of the ideas in a real world setting and the use of a narrative context.

4.2.2 Perceptions of multimedia design

The second objective of this study was to explore student perceptions of the authentic multimedia design and compare them to established principles. As described above, comments from the focus groups were broken down and coded into factor categories. These were iteratively refined to ensure fidelity to the data. Figure 4.2 shows the final categories and the relative frequencies with which they emerged in each focus group. Each of these categories is addressed below, indicating the way in which the factor appeared and how this compares with relevant research.

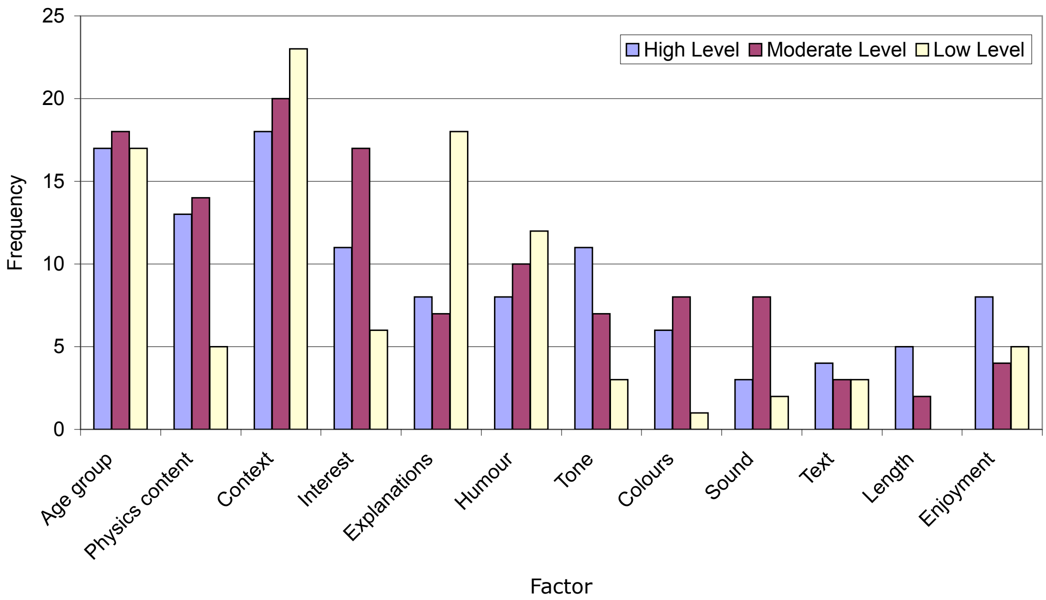


Figure 4.2: Number of comments pertaining to factor categories.

Age group

In all of the focus groups, the target age group of the video emerged as an important issue. Specifically, all participants felt that *Falling Cats* was appropriate for younger viewers, between the ages of twelve and sixteen. This opinion had important implications for the ensuing discussions as participants adopted different views of their roles in the focus groups. Some students saw themselves as authorities on learning, based on their experience at school, and commented on how younger students would view the video. Others viewed themselves as the target audience and responded to questions as though they'd seen the video in a lecture. Sometimes these two perceptions were discussed explicitly.

M3 They didn't really explain what it was all about. I think there was more on cats than on actual terminal velocity.

M4 There's books, there's lectures...

M3 I know, but if you're going to make it about [terminal velocity]...

M4 Depends what age group it's for, though. For us we want to know more about it. Whereas for younger audiences... It introduces them to it, so it just depends on the age group.

The focus on the target age group of the video was unanticipated, as *Falling Cats* is commonly shown to first year students. However, this issue often resulted in contrasting opinions:

H5 I suppose maybe it's because it's probably for a younger audience, but [it's] just that [the video] kind of talked down to the audience.

Later, a participant in the same focus group remarked,

H6 I know for younger audiences it would help a lot. I don't know if it's just that I'm really immature but I'd prefer to learn from that than listening to a lecturer at the front of a lecture theatre.

To prevent these two viewpoints from skewing results, comments specifically pertaining to a younger age group were classified in the age group category and were not included in the other categories. In research literature, animation and humour are two aspects of videos that have been shown to be more successful for younger audiences (Wetzel, Radtke & Stern 1994). This could have contributed to participants' perceptions that the video would be more appropriate for younger viewers. The consensus reached by the vast majority of participants was that the video was targeted at younger viewers but could be used as a starting point for university students.

Physics content

This desire for more physics to be included in the video was expressed by participants in all focus groups. As shown in Figure 4.2, this theme, not surprisingly, emerged twice as strongly in the moderate and high level groups than in the low level group. Although it was dependent on the target age group, in general participants felt the video "should be more information dense." This was not necessarily a negative reflection on the film, however, since participants believed they could learn more from it.

> *M5* I think it should have gone more in depth about terminal velocity.
> *M6* Yeah, could've gone a bit longer.
> *M5* I think, cause it was interesting...
> *M7* You're just getting into it and then it's done.

Because the film caught their interest, participants felt they could pay attention and learn more if the video explored the physics to a greater extent. However, some participants in the high interest group, doubted that the physics could be explained in enough depth with only the use of the video.

> *H7* Well, it was good in the fact that it tells you what the topic's about, but it's bad that you don't have firm maths behind it as well. So it would be a good introduction, but whether it would work during the lecture...

Context

Two different ideas of context arose strongly in the focus groups: (1) the use of an animated video for teaching, and (2) the use of the *Falling Cats* narrative to situate ideas about terminal velocity.

Animated video. Participants uniformly agreed that the animation provided a unique and effective context for learning about terminal velocity.

M8 I think it's a different approach. It's a different approach to...

M9 Other physics teachers.

H8 If you can take that and put in university subject matter, then it would be a pretty good way to learn.

As is evident from the quotes above, participants related to and enjoyed learning in the animated video context. They often contrasted this type of learning with what they experienced in lectures, viewing this approach as more innovative and interesting. The use of animation in illustrating physics concepts is in line with the studies reviewed by Wetzel et al. (1994) and the principles suggested by Rieber (1990). Wetzel et al. found that animations were most effective for younger students, possibly contributing to the age group issue discussed above. Rieber argued that animation should only be used when it helped visualize motion or trajectory. These attributes are particularly relevant to the explanation of falling cats and terminal velocity.

Falling Cats narrative Participants also enjoyed many aspects of the storyline and narration. They thought the use of a cat to demonstrate the physics allowed them to better relate to the subject matter. They believed the details of the cat narrative and interwoven real world facts were intriguing and beneficial to their learning.

M10 I thought the fact that they used falling cats instead of something boring like a ball was a good idea. I think using something in real life was a good idea, because you can relate to it.

M11 It's an indirect way of teaching people, because it's not based on terminal velocity, it's based on cats.

Not only did the context situate the physics ideas of terminal velocity in a quasi-realistic setting, it added an affective component that caught students' attention.

M12 I liked it how Dr. Karl's cartoon face popped up occasionally as the narrator and the cat would sort of be like 'huh?' [laughter] Yeah, that was funny, engaging.

Clark & Mayer (2003) warn against using attention grabbing devices because they can bombard learners with extraneous information, making it more difficult to select the salient points from an intervention. *Falling Cats* is different from the examples described by these researchers, however, as the physics is integrated into the interesting context. Therefore, the narrative gives learners a structure on which to organise their knowledge, which has a positive impact on learning as shown by Mayer (2001).

Clark & Mayer (2003) have also demonstrated the utility of onscreen coaches or pedagogical agents; these are tutors that appear in multimedia presentations to guide learners. They argue that use of these onscreen personalities promotes student interactions with the technology in a more natural, social way that is more conducive to learning. Their research has shown that students perform better on tests of retention and transfer with the use of these agents than without.

Interest

As the lecturers predicted, a key aspect of the video recognized by students was its ability to inspire interest. This interest was expressed both towards the video and towards the subject matter in general. Students cited several reasons why they were interested in the video, ranging from the animation techniques to the narrative context.

> *M13* With the sounds and colours it was more interesting than just a boring video tape – you'd fall asleep. It was really jazzed up so you want to see what's going to happen.
>
> *M14* You're sort of interested because you want to know how many cats died...

Participants often voiced the belief that the video could make the physics more attractive. One student commented, "It would just be a quick introduction, to get [students] interested." (H9) From this introduction, participants believed the lecture would be more effective.

H10 It gives you a basic idea of the concepts, and then maybe you'd be better able to understand [the lecture] than without this intro.

Not all participants believed that the increased interest resulted in more learning, however, as the following quote demonstrates:

L4 The video would hold my interest more, but I'd probably in the long run learn more in the lecture, because though I wouldn't pay a lot of attention at the time, I'd write down a lot of notes, so. . .

This student doubts the informational content of the video or believes she'd better be able to recall the lecture by note taking. Studies have shown that students view televised instruction as effortless and ineffective for learning due to its close relationship to entertainment (Krendl 1986).

Explanations

The explanations of the physics in *Falling Cats* were regarded by the moderate and high level groups as very effective. Participants commonly identified the clear, simple language employed in the narration, supported by explicit visuals as the reason the physics material was easy to follow. They also felt that presenting the material in both verbal and visual streams was beneficial.

H11 It used video explanations in conjunction with talking so you were learning the same thing in two different ways.

Mayer (1997, 2001) has repeatedly shown that learning with text and pictures or animation and narration is more successful than with any one mode alone. This finding is in line with Paivio's dual coding theory (1971, 1986), which says that the brain processes verbal and non-verbal information separately, as discussed in Section 3.3. When the information in these two streams is part of the same message, linkage formation between verbal and non-verbal models significantly increases the probability of recall and transfer.

In the low level session, the explanations received a mixed response. Again, some respondents felt the physics was presented clearly.

> *L5* They wouldn't just say that gravity is this thing and give you a formula, but they explain it to you in a way that you know what they mean, and show examples.

Others, on the other hand, felt the details given in the video lacked the rigor necessary to support the physics assertions presented.

> *L6* I thought some of the concepts that I wasn't familiar with, I would have liked to see explained. I don't really understand dynamic tension. I also felt the details were anecdotal, like I don't necessarily feel I've learned some facts I can use or rely on.

This suggests that the video explanations were less accessible to the low level group due to these participants' lack of familiarity with physics material. The difficulty that novices encounter in learning from presentations that are extremely useful to experts has been documented in many different contexts (e.g. Lowe 2003, Lowe 2004).

Humour

Participants in all focus groups addressed the humour in Falling Cats before discussing most other aspects of the film. Despite some commenting that the jokes were not targeted at the university age group, humour was appreciated by the majority of students, who also felt it helped them learn.

> *L7* I just thought that the humour was very efficient.
> *I* Efficient?
> *L7* Yeah, as a means of teaching.

Humour was mentioned by participants in all groups as one of their favourite aspects of the video. The only contentious point surrounding humour was to do with the use of puns. Some students thoroughly enjoyed them, while others thought they were tacky and inappropriate. Whether liked or disliked, however, the puns elicited a strong reaction from the focus groups.

In reviews of related literature, the consensus reached is that humour is most effective for younger viewers with decreasing impact as age increases (Zillmann, Masland, Weaver, Lacey, Jacobs, Dow, Klein & Banker 1984). "Only humour that is well integrated with educational materials is likely to enhance student-teacher rapport and does not produce the negative effect on learning" (p.802). These findings may not be applicable to the current cohort of students, however, as the culture of entertainment and learning has changed dramatically in the past twenty years. Adult cartoons are much more common in entertainment and students increasingly expect to be entertained in lectures (Altschuler 1999). Since there are ever expanding avenues through which students can find information outside the lecture (McInnis, James & Hartley 2000), lecturers must find innovative ways of presenting material or risk declining attendances (Stevenson 2005).

Tone

Participants appreciated the tone of the video, especially the lively, colloquial narration by Dr. Karl. They felt this attracted them to the video.

M15	The tone of voice. It wasn't monotone, like it was–
M16	It was engaging...
H12	It was pretty easy-going and it didn't take itself too seriously, [that] kind of made it fun.

This topic also lead to comparisons with physics lecturers as students felt the tone contrasted strongly with what they experienced in class. Of Dr. Karl, one student remarked, "he's not some droning lecturer" (H13).

Clark & Mayer (2003) found that using personal rather than formal language promoted learning. They believe the use of casual language in the first or second person encourages more natural social interactions with technology, analogous to the effect of using a pedagogical agent. Learners achieved higher scores on transfer tests after watching personalised interventions than after watching more formal ones.

Colour

The colours used in the video were another aspect of the animation mentioned in all focus groups. This is an area in which the video design departs from research-based recommendations and adopts a more intuitive approach. It is one of the ways in which the video is a 'craft' solution. The video used an extensive palette of bright colours to create complicated and colourful scenes and backgrounds. Most participants agreed this added to the visual appeal.

> *M17* I thought it was really colourful. Just the colours, really—it comes at you, but in a good way.

Only one student in the low interest group commented on the colours, finding them distracting:

> *L8* I think at times it was a bit hectic, because there were too many colours and it changed all the time.

Despite a small increase in learning with colour as opposed to black and white media, the main recommendation from the literature is that colours should be limited to allow for clear discerning of objects (Wetzel et al. 1994). Practitioners like Blinn (1989) agree, arguing the essential objects in the animation should stand out and background details must be minimised. It is possible that some negative student perceptions around colour use in this video were due to a loss of clarity in the video image due to VHS dubbing. One student in the high interest group may have picked up on this:

H14 I felt that since it is a visual medium for educating people, the colours used were all reds and light purples and all that, so sort of confusing that . . . images were blurred so you couldn't really bring out where the cat was in between the sky.

Some studies have shown that a loss in pictorial clarity can result in detrimental effects on students' perceptions and learning from a medium (Wetzel et al. 1994).

Sound

The sound in *Falling Cats* is layered throughout with several musical backgrounds, narration, cartoon sound effects, the cat's vocalisation, and ambient street noise. This choice again clearly illustrates the 'craft' nature of the intervention as it goes against research findings that suggest sounds should be kept to a minimum to promote learning (Clark & Mayer 2003, Mayer 2001). Previous findings also showed adding sounds "as background noise or to add realism to narrative" have no effect on learning (Barrington, 1970, 1972, as cited in Wetzel et al. 1994, p.143).

Participants in the low level group reacted negatively to the sound in the video. They felt the sounds were sometimes distracting or that the music didn't seem appropriate.

L9 I didn't like the music in the beginning because it made too much of an effort like 'OK, let's do something that young people like.'

In the moderate level group, reactions were mixed.

M18 I said the music, the sound effects, and also the fast cartoon, I thought it might distract kids a bit from Karl's speech.

Others thought that the sounds attracted their attention and made the video more enjoyable.

M19 I thought the music and all that, was like—I actually had that as one of my likes, because it kept me interested.

This was a thought echoed by the high level group, who felt the sounds and music added to the humour of the film. The moderate and high interest groups may have felt more positively about the sounds than the low level group because they had more experience with the physics and were therefore less likely to feel overwhelmed or distracted by extraneous sounds and colours.

Text

One aspect that was appreciated by participants in all three groups was the use of large bold animated text that appeared on screen every time a new topic area was introduced or when some concepts were explained.

M20 One thing I did like was that they reinforce all the main points with big bold writing.

H15 Topic headings like how he goes 'this is the realm of fact but now we're moving into the land of theory.'

Mayer (2001) showed that use of on screen text in conjunction with pictures and narration had a detrimental effect on learning, as it split the attention of learners between the visuals and text. He suspected, however, that use of on screen text when not in competition with other visuals might yield different results. This is a point supported by Borzyskowski (2004), who argues that literacy is more a basic skill than the ability to comprehend complicated visuals. Therefore, animated text may be the intermediate step between text and visual representations necessary to scaffold novices and learners from one representation to another.

Length

In the moderate and high level groups, participants agreed that the video was short and concise.

M21 It was quick. Like it just didn't drag on.

M22 Yeah, it was straight to the point.

This opinion had both positive and negative implications, however, as some found the video succinct, while others found it lacking in substance. This led to suggestions that more physics content be included, as outlined above.

Since the majority of teachers use videos "as a supplementary part of their teaching strategies and not as an alternative to them" (Barford & Weston 1997, p.46) the video is probably an appropriate length for viewing in lectures. In addition, the finding that students wanted more physics from the video suggests that lecturer reviews and additions after the film would be an effective method of teaching. Introductions and follow up discussions have been shown to increase learning, sometimes more than a second viewing of the film (Wetzel et al. 1994, Russell 1985).

4.3 Discussion

In this study, the first two levels of Kirkpatrick's evaluation were performed on the popular science video, *Falling Cats*, using three focus groups of students with varying physics backgrounds. From the transcripts, a series of categories were developed, which were then used to compare student preferences to multimedia design principles. These preferences can be applied to the development of future research questions and educational interventions since "many successful technologies have resulted from descriptive reasoning about why the craft solution worked" (Clark & Estes 1999, p.10).

The implication for teaching is that showing *Falling Cats* in class should stimulate thinking about the physics ideas in students of all levels of interest and prior knowledge. In a discussion setting, students showed a willingness to ask questions and respond to peers. Although this may not be easy to replicate in lectures, it may be possible to duplicate in tutorials. This makes use of the video more effective, as shown in many previous studies (Wetzel et al. 1994). It is possible that the results of this study were achieved due, at least in part, to the focus group setting and the nature of the study.

Not surprisingly, students with higher interest and experience in physics were more confident in talking about the concepts and responding to each other. These students showed a deeper understanding of terminal velocity, measured by the range of concepts suggested, the confidence with which these ideas were presented, and the ability to form coherent opinions. Irrespective of interest or prior teaching in physics, students thought about the physics after watching Dr. Karl's *Falling Cats*, and were open to talking about their conceptions and concerns. In fact, even in the low interest group students chose to discuss the physics concepts without prompting in this area.

In general, students' preferences matched up surprisingly well with prominent multimedia design principles. Participants were positive about aspects of the video, such as the casual tone, use of an onscreen pedagogical agent, and context on which to organise knowledge, supported by research findings. Reactions were mixed, however, to the colour and sound in the video, which were more complicated than studies have shown to be effective. Humour was one aspect that has been shown to have little effect on learning for older students but was strongly liked by the majority of participants in all focus groups. This may be due to the way in which humour was well integrated into the subject matter or due to the changes in student culture over the last twenty years. In recent decades humorous animations for all age groups have risen in popularity.

The results of this study were complicated by two factors: (1) the emergence of the target age group as an important issue in all focus groups, and (2) the overwhelming of novice learners by sounds, colours, and anecdotal evidence. Due to conflicting opinions on the target age group of the video, participants either adopted an expert stance or viewed themselves as the target group of the intervention. This led to contrasting opinions of various aspects of the film. The low interest group was more negative about the context, explanations, colours, and sounds in the film, likely due to their unfamiliarity with the subject matter.

Students, regardless of background or intrinsic interest in physics, wanted to know more about the physics after watching the video. They also wanted more physics content to be included in the video style. This suggests that *Falling Cats* is effective in the manner perceived by lecturers.

The categories outlined in this study cover a broad range of aspects that were important to students and would be vital to consider when designing future educational videos. This study also highlights the need to evaluate educational interventions, not only when they are first developed and introduced, as is often the case, but when they have been selected by lecturers or used routinely.

Chapter 5

Quantum mechanics teaching

Designing multimedia for physics education not only requires understanding learning with multimedia but also current teaching and learning practices in physics. Therefore the second major focus of preliminary research was on current physics teaching and the learning that results from it. In this chapter I describe an audit of physics lecture courses across junior, intermediate, senior, and honors levels at the University of Sydney. In Chapter 6, I evaluate the learning that takes place in the intermediate course.

Physics is a large and diverse field and, consequently, undertaking a comprehensive review of teaching practices was impossible. It was important, however, to characterize physics teaching across all levels of tertiary education. This cross-sectional view arguably better represents the diversity in teaching than a survey of different topic areas at one level.

Quantum mechanics was selected as the area of physics to investigate in each year for several reasons. It is a topic with which students have little experience in high school. It is a fundamental aspect of physics, dealt with in each of four years of an honors degree. Quantum mechanics is also growing in importance in research and technology development. Considered by many to be one of the most difficult and counter-intuitive areas of physics, it has only recently begun to attract the attention of educational researchers.

This investigation sought to achieve a number of objectives. Primarily, the goal was to provide a snapshot of teaching strategies employed by different lecturers at different levels of a physics degree. It was important to document teaching practices in conjunction with resulting learning to identify strengths and weaknesses of existing approaches. As suggested in Chapter 1, weaknesses in current teaching present potential opportunities for multimedia. At the very least, characterizing the way physics is taught and learnt in a tertiary institution yields insight into the unique challenges and possibilities presented by physics education.

5.1 Introduction

Quantum mechanics was selected as the focus of the lecture audit for numerous reasons. Importantly, the subject is representative of physics more generally. Widely regarded as one of the most difficult subjects to learn or teach, quantum mechanics is counterintuitive, relies on high-level mathematics, and involves abstract subject matter. It typifies the 'complex knowledge domain' characteristic of physics (Sharma, Millar, Smith & Sefton 2004). Students of quantum mechanics must learn to reject classical conceptions of nature like the wave/particle dichotomy and develop entirely new understandings of matter and energy. Even experts in the field debate the interpretations of quantum mechanics because traditional notions of reality are inadequate in modelling phenomena on small scales. In addition, only a limited range of quantum problems can be attempted analytically as computers and approximate methods are required to solve even the most rudimentary problems. Furthermore, students cannot readily see nor do they have personal experience with quantum effects. Quantum mechanics is an intricate subject, demanding of its students a wide array of skills from advanced mathematics to conceptual abstraction.

Quantum mechanics is also fundamental science. It is a relatively new area of physics, developed over the last century, yet it has far-reaching implications for our understanding of the universe. Because it is so fundamental, it is taught at all levels of an undergraduate degree.

In addition to its scientific significance, quantum mechanics is becoming more and more relevant to technological developments and our world view. Recent discoveries demonstrate the central role quantum mechanics will play in a myriad of fields and in future scientific research. From nuclear medicine and radioactivity to nanotechnology and quantum information theory, quantum mechanics promises to influence not only technological developments, but fundamental philosophical perspectives as well. Therefore it is essential that future generations of students, not just physics majors, learn and understand quantum mechanics on a deep, conceptual level.

Compared to many other areas of physics, however, quantum mechanics has attracted fairly little attention from educational researchers (*Research on teaching and learning quantum mechanics* 1999). Debates have erupted among educational researchers and practitioners over how to approach the teaching of quantum mechanics. Due to long-held views of some physics professors, "introductory courses are still taught in much the same manner as they have been for the past seventy years" (Fletcher 2004, p.1). Some support this practice believing "the mathematics should be covered 'step by step and then tie in the physics,'" (p.104) while others promote the opposite viewpoint. They suggest that traditional teaching methods may lead to misconceptions and advocate a qualitative conceptual approach without the historical underpinnings of the subject (Muller & Wiesner 2002). There is no clear resolution to this debate, however, because the development of improved teaching methods is in its infancy. In fact, most of the research on the teaching and learning of quantum mechanics has occured only in the last ten years. This is likely due to the high level of the subject matter and the small number of students taking quantum courses.

What research does exist on teaching and learning quantum mechanics?

Studies thus far have focused on misconceptions (Styer 1996), quantum curricula (Wittmann, Steinberg & Redish 2002), student conceptions and understanding (Fletcher 2004, Johnston, Crawford & Fletcher 1998), experimental teaching practices (Muller & Wiesner 2002, Zollman, Rebello & Hogg 2002, Lawrence 1996), and specific quantum topics (Wittmann et al. 2002, Olsen 2002). Recently, multiple-choice tests called the Quantum Concept Inventory (Falk 2004) and the Quantum

Mechanics Conceptual Survey (McKagan & Wieman 2005) have been developed to evaluate student learning of quantum mechanics. Efforts have also begun to teach the subject earlier in schools and in innovative ways to facilitate conceptual learning (Olsen 2002, Zollman et al. 2002, Muller & Wiesner 2002). Most of the research on teaching practices, however, has been performed by practitioners on interventions they have developed. Searches of educational databases yield no literature independently documenting the current practices of quantum mechanics teaching in a tertiary institution.

5.2 Method

Over the course of this study, nine lecture classes were surveyed: three first year classes, second and third year quantum mechanics—each of which was divided into Regular and Advanced streams, honors level relativistic quantum mechanics (RQM), and Advanced quantum mechanics (AQM). The number of lectures I attended, the total number of lectures, enrolment, and average attendance are summarized in Table 5.1.

Year	1	1	1	2	3	3	4	4
Designation	Env	Tech	Adv	Reg/Adv	Reg	Adv	AQM	RQM
No. attended	10	7	2	13	3	7	11	18
Total lectures	14	12	12	19	19	19	20	21
Enrollment	83	115	160	121	29	64	20	N/A
Average attendance	47	69	118	94	16	54	19	14

Table 5.1: Details of quantum mechanics classes surveyed.

For first year students, quantum mechanics is offered in second semester to the three different streams called Environmental (Env), Technological (Tech), and Advanced (Adv). These streams evolve out of the first semester designations of Fundamentals, Regular, and Advanced, though some students switch streams between semesters. The first year Environmental class is mostly made up of students from the Fundamentals stream. Therefore few have prior formal physics experience beyond first semester physics. The Environmental course places special emphasis on radiation and its interaction with matter. The Technological class draws the majority of its students from the Regular stream so most students have high school backgrounds in physics. The Advanced first year class is for students with strong backgrounds in high school physics who are interested in developing a deep, mathematical knowledge of the subject. Most physics majors and graduate students come from this stream. The second year class attempts to give students a strong conceptual foundation in quantum mechanics. Although this course is divided into Regular and Advanced streams, most lectures are common to both groups. A few times during the semester, the lecture streams split with the Advanced class working through enrichment material and the Regular class reviewing core topics. In third year the Advanced and Regular streams are entirely separate, taught by different lecturers with different focuses. At this level, students are introduced to Dirac notation, with the Advanced students working extensively in this formalism. Regular students explore the implications of quantum mechanics in the context of spectroscopy. AQM further develops quantum ideas and introduces topics of current research. RQM is the highest-level course on the subject, enabling students to understand and solve problems pertaining to the combined theories of relativity and quantum mechanics.

I collected observations as an observer-participant immersed in lectures as a student. The 'student perspective,' as it appears in this chapter, is a construct based on conversations with students, observations of the class as a whole, and personal experience. Relevant works on the student perspectives of learning in higher education include those by Prosser & Trigwell (1999) and McInnis et al. (2000).

To perform a content analysis, lectures were broken down into segments of teaching time, determined by topic or teaching method changes. These ranged from one to twenty-five minutes in length. Each teaching segment was rated on several dimensions pertaining to the subject matter addressed and the way in which it was presented. Fletcher (2004) identified areas of difficulty in quantum mechanics education by accumulating and thematically coding interviews with students and lecturers. These were used as a starting point for the dimensions used to rate each teaching segment. Selected quantum mechanics topics of interest were:

- wave-particle duality
- the uncertainty principle
- tunneling.

Eight teaching approaches were identified and used to characterize the mode of teaching in each lecture segment:

- demonstrations or visuals
- analogies
- real world examples
- mathematics
- discussions
- predictions
- history
- lecturer explanations.

A six-point rating scale was used to give a rough measure of the significance of the above items in each lecture segment. The criteria evolved over the course of the study and are shown in Table 5.2.

A sample teaching segment rating is given in Table 5.3. This particular lecture segment involved an explanation of the Heisenberg uncertainty principle with an analogy to uncertainty in measurement on a macroscopic scale.

In addition to these ratings for each teaching segment, the subject matter and any other relevant details were noted. These data were compiled in a database and analysed for the most significant trends.

5.3 Results

Data from all nine courses were grouped together and plotted as shown in Figure 5.1. This plot reveals the most and least common uses of lecture time. From the data, four key themes emerged in the areas of interactivity, visuals and demonstrations, mathematics, and course content. The theme of interactivity encompasses ob-

Score	Criteria	Student perspective
0	Item was not associated with the segment explicitly or implicitly over the entire duration	No correlation between the item and segment
1	Item was implicitly associated with part of the segment	Only the keenest of students would recognize the relevance of the item
2	Item was implicitly or explicitly associated with part of the segment but was not presented in a meaningful way	Advanced students may see the relevance of the item in the segment and gain something from it
3	Item was implicitly or more likely explicitly associated with part or all of the segment and was sufficient in its presentation	The majority of the class would see the relevance of the item in the segment and may gain something from it.
4	Item was explicitly associated with most if not all of the segment and was well presented	Almost all of the students would see the relevance of the item in the segment and should learn something from it.
5	Item was explicitly associated with all of the segment and was presented in a way that inspired interest	Even those students not paying attention would be attracted to the segment and therefore see the relevance of the item and engage with the content

Table 5.2: Overall teaching segment rating scheme.

Time (min):	12	Real world:	0
Visuals:	4	History:	0
Math:	2	Explanation:	4
Analogies:	3	Wave-particle duality:	3
Predictions:	0	Uncertainty:	5
Discussions:	0	Tunneling	0

Table 5.3: Sample teaching segment rating from the first year Environmental class.

servations of discussions, predictions, and lecturer explanations. These themes are discussed in detail below. Three observations that arose independently of the quantitative data, but which had strong bearing on the results were: the influence of the lecturer, students' attitudes towards learning and the unique attributes of relativistic quantum mechanics (RQM). These are addressed first, before the key themes.

Influence of the lecturer

The tone of the class, the methods of teaching, and the type of learning taking place were, not surprisingly, heavily lecturer dependent. This deserves emphasizing, as the lecturer's preparedness and presence in class were the most influential factor in the data set. While content was very similar in the first year Environmental and Technological courses, the lectures were very different in emphasis, length, depth of explanations, visual aids used, and mathematical descriptions. This is the most obvious cause of variation among the courses surveyed. From discussions with the students, lecturers who were perceived to be the most interesting were not always perceived to be the best to learn from, and vice versa.

Students' attitudes towards learning

Students in the first and, to a lesser extent, second year classes tended to view the lecture in a very social way. Noise levels in the Technological and Advanced first year streams were consistently high. Students were often late to class with many arriving up to twenty minutes into the lecture. This did not appear to bother the lecturers. It was evident that for the vast majority of students, intrinsic interest in the presented material was limited. This is in line with the study conducted by McInnis, et al. (2000, p.20) which established the trend that considerably fewer students are finding lectures a valuable source of learning. The findings may indicate that lectures are becoming a less important learning environment for students because they now have more access to other forms of information, especially via the Internet.

Interviews performed as part of the study described in Chapter 8 shed light on student perspectives of lecture learning. In small focus groups, students from the intermediate quantum mechanics course discussed their opinions of current physics teaching practices. In the following quotes, the letters 'R' and 'A' are used to refer to refer to students from the Regular and Advanced streams, respectively. The letter 'I' is used for the investigator.

An interesting theme common to all interviews was the limited interactivity in lectures and the declining importance of lecture learning. Students in most interviews regarded lectures as guides to the curriculum rather than as significant learning environments. This reflects the trend towards flexible learning in higher education. With growing employment commitments and increasing availability of diverse resources, students are focusing more on learning outside the traditional lecture (Prosser & Trigwell 1999).

A1 [The lecturer] is a really good lecturer and he's really charismatic, and you're really engaged when you listen to him but sometimes you leave the lecture and you're like ahh I understood it while he was saying it and now I've kind of lost the idea, and I find the Internet really useful. You find lots of different ways of explaining.

A2 I use the lectures as a guide for what I need to know and just work with the textbook so it's just like a…syllabus I follow—just to help direct the study.

R1 Some textbooks are too complex, like they just have the equations all over the place. Um, some textbooks are alright. I don't even mind using them. I prefer the web probably, because I can probably find a lot more resources on the web than just in one textbook.

These students felt the lectures had some value but they believed most of their learning occurred outside of the classroom. They cited the course notes and computational labs as important components of their learning. All interviewed students marginalized the role of question asking and discussion in lectures, concordant with the observations of quantum physics courses.

Relativistic quantum mechanics (RQM)

Of the nine courses surveyed, RQM was the obvious exception. All lectures in this course used the traditional 'chalk and talk' style. The vast majority of information covered was mathematical in nature, limiting the number of diagrams and pictures presented. The class was also the smallest surveyed and the attrition rate was the highest at 45%. For these reasons, RQM could not be analysed in the same manner as the other eight courses and is therefore omitted from the following figures and discussions unless otherwise noted.

5.3.1 Key themes

Figure 5.1 illustrates the breakdown of teaching time in all lectures. It is important to note that these data have not been normalized to give equal weighting to each class surveyed. The features of this chart were used to determine the key themes related to current teaching practices. The most common use of teaching time was the lecturer explanation, whereas discussions and predictions were among the least common. These aspects are considered together under the heading of interactivity.

The next most common teaching method involved the use of visuals or demonstrations. Mathematics was the third most significant aspect of teaching and was approached in very different ways by different lecturers. Finally, the historical emphasis of the curricula was a common theme in the majority of lecture courses.

Interactivity

In terms of the lecturer speaking for the vast majority of class time, almost all lectures adhered to the traditional format. The few exceptions were from second year classes where two lectures were explicitly question and answer format. However, the questions posed by students lacked preparation, were often only tangentially related to class material, and were commonly of interest to only a few in the class. This raises the issue of how to better structure or plan these sessions to make the best use of class time. In the Advanced first year classes, the lecturer often posed ques-

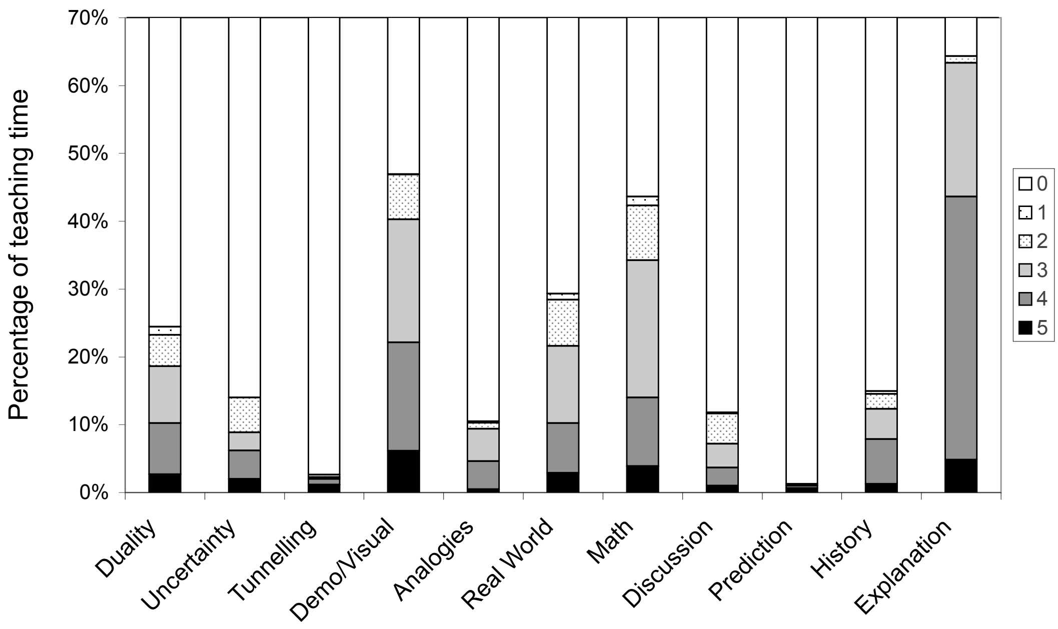


Figure 5.1: Breakdown of teaching time for all quantum mechanics lectures

tions but only a handful of students responded. In general, questions were rarely asked by the lecturers or by the students.

In all focus groups, students talked about the fear they had of showing their ignorance in large lecture classes. They also felt that questions asked by other students rarely added value to the lecture.

R2	You don't mind being wrong to the tutor but when you're wrong to a hundred others it becomes a bit of a stumbling block. [laughter]
R3	Yeah.
R4	You feel like you could be slowing them down as well.
I	Mm.
R3	Some of the questions that sometimes people ask, you don't even know what they're talking about.
R4	Yeah, the questions that people ask in lectures are always beyond the material rather than in the material.

Regular students felt they might be looked down upon by the Advanced students or that they might be slowing their peers down. "Normal people think they're probably going to look stupid if they ask a normal question like just something really simple. . . The Advanced people will think these guys are dumb." However, Advanced students had similar concerns.

A4	I think one of the common fears in lectures is that the question you ask uh– [laughter]
A3	Makes you look like an idiot.
A4	That's right.
A5	That's why hardly anybody ever speaks up in a lecture usually I think.
I	And is that just physics or is that all lectures?
A5	All lectures.
A3	All lectures, yes.
A4	All lectures.
A5	There's usually one or two people down in the front who might ask a few questions.
A3	They can whisper–
A6	They are the very smart people.
A5	'Cause that's right.
A4	They are very profound questions. It's hard to bring up real questions.
A5	They clearly have been following the lecture to the second. And then everybody else is just sitting there going 'somebody should ask this question' but they're not going to be the one to do it.

The topic of discussions in lectures elicited a strong reaction from all focus groups. Virtually all students agreed that question-asking is important for learning but that it rarely occurs in lectures. They believed the intimidation of the large lecture setting and a general fear of being wrong are the fundamental reasons so few questions are asked. Of the few questions that are raised in lecture, most felt they were either tangential to the subject matter or so sophisticated that they were incomprehensible.

Predictions. Although lecture demonstrations and computer simulations were common in lectures, students were rarely required to make predictions beforehand. One exception was the second year course, which involved the CUPS (Consortium for Upper-level Physics Software) simulations (Hiller, Johnston & Styer 1995, Johnston & The Consortium for Upper Level Physics Software 1996). Students were asked to predict what would happen to a wave packet at a boundary or what effect decreasing the depth of a potential well would have on the bound energy states. According to a study by Crouch, Fagen, Callan & Mazur (2004) students who witness demonstrations without being asked to make a prediction perform as well on follow-up tests as those who don't see the demonstration at all. Therefore, to make the most efficient use of these presentations, students should be asked to predict. Sometimes this is not possible due to the nature of the activity, but often with a bit of ingenuity a prediction could be included.

Demonstrations and visuals

Innovative use of technology in the classroom was one of the most significant findings of this study as it exemplifies the departure from traditional lecture methods and demonstrates the extent to which lectures are a sophisticated form of multimedia.

Lecture slides. PowerPoint presentations were the basis for all lectures attended (besides RQM) and students were given slide handouts in lieu of taking notes. Lecturers commonly expressed the belief that students should not be required to take notes to promote lecture learning. Support for this notion exists in literature. Badger, White, Sutherland & Haggis (2001) found that students often view note-taking as distracting. One student in the study remarked, "I have to concentrate on what he [the lecturer] says. I don't have time to take notes" (p.6). The most common purpose students cite for note-taking is to be able to recall the lecture later, but if the slide handouts fulfil this function, note-taking may be superfluous. This philosophy was beneficial for students who paid more attention to the lecture but it often allowed disinterested students to 'tune out,' relying on the handouts to contain the important information. Literature also exists supporting the practice of note-taking (Ryan 2001). The main advantage of PowerPoint was that it allowed lecturers to integrate complicated graphics into the lecture and display important information in large, legible writing.

Lecture demonstrations. Despite the traditional difficulty of demonstrating quantum mechanics concepts in class, several experiments were conducted in each of the introductory courses, often with the use of real-time technology. In the Technological lectures, use of a spectrometer and real-time software allowed for the illustration of the emission spectrum of a blackbody with comparisons to a hydrogen lamp and fluorescent lights. In the Environmental class, Planck's constant was determined in an Excel spreadsheet using data collected in lecture from a desktop photoelectric experiment. Simple but effective demonstrations were also performed, notably Bragg electron scattering and superconducting levitation. The latter inspired great interest from the students both during and after class.

Simulations. Where desktop demonstrations weren't available to illustrate quantum behaviour, visualization software was used. In the second year lectures, the CUPS program depicted the behaviour of wave functions at barriers and in potential wells. This program was frequently used to explain the counterintuitive concepts of interference and tunneling, and decreased the dependence on math to describe quantum systems. With the use of simulations, parameters could be changed and their impact on the outcome determined.

Interest agents. Not all of the technologies used in class were of a technical nature; in fact, some of the most interesting, from a student point of view, had little physics content. In the second year class, a clip from an old James Bond film was shown during a lesson on lasers. The actual characteristics and physical structure of a laser were compared against the Hollywood depiction. Students were certainly captivated by this teaching segment. In the Environmental lectures, Nobel laureate Richard Feynman's song "Orange Juice," was played to give students an idea of the scientist's eccentric nature.

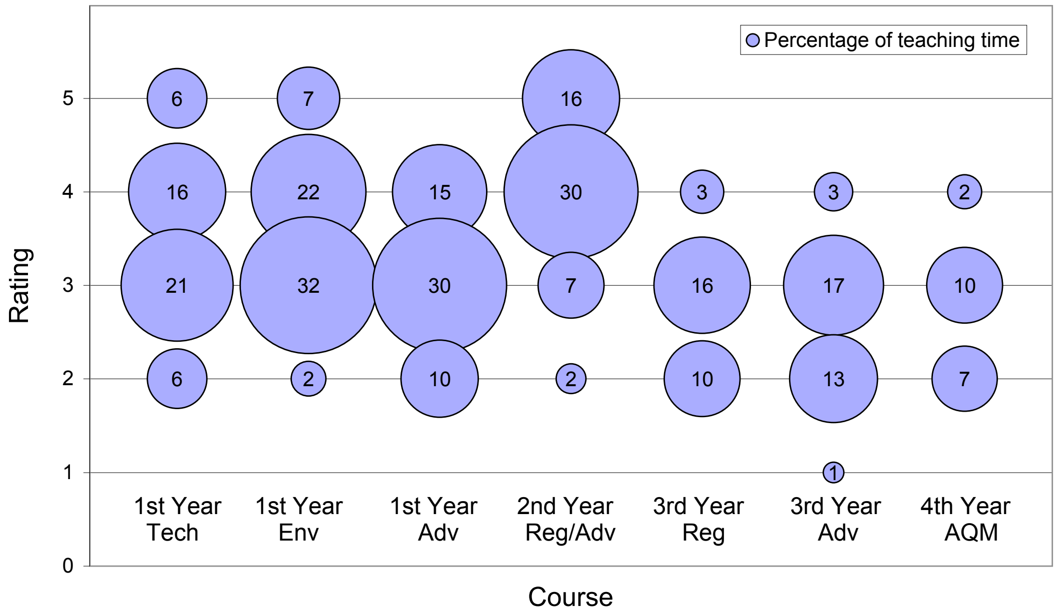


Figure 5.2: Use of visual aids; numbers in bubbles indicate the percentage of total teaching time.

Visual aids. Overall, the use of technology in the first and second year classes helped create learning environments employing visual aids for over half the total teaching time, as shown in Figure 5.2. It was reported by Fletcher (2004) that both students and teachers believe the lack of visualization in quantum mechanics to be at the heart of its educational challenge.

The general trend revealed in Figure 5.2 is that the number and quality of visuals and demonstrations decrease over the years of quantum mechanics lectures. This reflects the increasingly mathematical nature of the subject and the difficulty in displaying quantum phenomena visually. A notable exception is the second year class, which showed the highest frequency of effective visuals. This is likely because the lecturer for the course was one of the developers of the CUPS simulation software, which he used extensively in his lectures.

Mathematics

Each lecturer handled the intrinsically mathematical nature of quantum mechanics in a unique way. In the Environmental first year class, assignment-type questions were solved in their entirety on PowerPoint slides. In the Technological class, some extensive math problems were solved in detail by the lecturer on the chalkboard but were not included in the lecture handouts. Students in the Advanced class were asked to solve problems on their own during class time. The lecturer would give suggestions, but leave the students to perform mathematical operations individually. He would then ask students to reveal how they proceeded from one step to the next. Mathematics was limited in second year lectures as there was a much stronger emphasis on concepts than problem solving. Computational methods were explained in class and the numerical solver was used in many lecture problems. In AQM, the lecturer solved some mathematical questions using a tablet PC, displayed to the class through a projector. Other questions were assigned for students to solve in groups during the lesson. Mathematics was explicitly the language of RQM and derivations took up the vast majority of lecture time.

Course content: history

All classes approached quantum mechanics in the same way, with a historical narrative detailing the failings of classical mechanics and outlining the ways in which quantum mechanics solves its problems. This history was covered, to greater or lesser degrees, in all first year streams, and again briefly in second year. This is useful in that it introduces all students to the problems of quantum in a fairly intriguing way and primes the prior knowledge of students in these areas. Across the first year streams and even into second year the curricula are fairly similar. In second year, a wider range of topics are investigated and in greater depth with the use of computational methods.

5.4 Discussion

Although all classes audited followed the form of the traditional lecture, the assertion that content and teaching methods have remained unchanged for seventy years is patently false in the School of Physics. With the exception of RQM, all courses employ visual aids for a large fraction of class time, ranging in complexity from pictures and diagrams to dynamic simulations and excerpts from commercial films. In this way, quantum mechanics lectures are an intricate form of multimedia.

Integrated technologies are not being exploited to their full potential, however, as interactivity remains low in all classes. More predictions and discussions should be elicited from students to increase the impact of demonstrations and to engage students with the physics content. There are obvious barriers to more discussions in lectures, however. Students are concerned about revealing their ignorance in front of the lecturer and hundreds of their peers. Some feel that the problems they are experiencing are unique to them and therefore they don't want to take up class time. Social interactions, though recognized as important by both lecturers and students, are difficult to facilitate in the large lecture setting.

Some new technologies could help improve interactivity. Classroom communication systems (CCS's, also known as 'clickers') allow students to respond anonymously to multiple-choice questions and have their answers tallied in real time. Bar graphs made by accompanying software reveals the range and frequency of student responses. This immediate feedback is useful for both lecturers and students. The lecturer can see at a glance if the majority of the class has understood key ideas in the lesson and students can see whether their answer was common or not. Often clickers are used in conjunction with a strategy like Peer Instruction, which requires students to discuss their answers with a neighbor after viewing the bar graph (Mazur 1997). Classroom communication systems can be difficult to implement for practical reasons. Initial setup costs can be significant and training staff requires time and effort. Clickers are apt to get lost so either students or universities must take on replacement costs. Using low-tech flash cards has been shown to be as effective as using clickers (Lasry 2007). This may be a cheaper and easier way for lecturers to inject interactivity into their lectures. In accordance with the equivalence principle, it indicates that the selection of an answer and discussion inspires the cognitive activities required for learning regardless of the technology involved.

Chapter 6

Quantum mechanics learning

Evaluating the learning that takes place as a result of the teaching practices documented in Chapter 5 represents the other half of the physics education equation. The scope of the investigation was narrowed because quantum mechanics in itself is a vast area taught at different levels of sophistication over the four years of a physics degree. For a review of quantum mechanics education research, see Falk, Linder & Kung (in press). The intermediate quantum class was selected as the sample for this investigation because it has a reasonably high enrollment but it covers the fundamental aspects of the subject in more detail than the first year classes. Quantum mechanics is too diverse a topic to evaluate comprehensively so a problem representative of the domain, quantum tunneling, was used to assess student thinking.

Quantum tunneling has been selected by several researchers as an area for investigation because it highlights some of the key issues of quantum mechanics (Redish, Wittmann & Steinberg 2000, Morgan et al. 2004, Ambrose 1999, Domert et al. 2005, Zollman et al. 2002). Quantum tunneling is relevant to a range of emerging fields from solid state and nuclear physics to medicine. Redish et al. (2000) see tunneling not as a peripheral example of quantum mechanics but as a 'conceptual touchstone.' They highlight the important role played by quantum tunneling in molecular bonding and formation of band structures. Furthermore, they argue, "understanding quantum tunneling requires understanding a number of fundamental but difficult conceptual issues" (p.7). Given the abstract and often theoretical nature of quantum mechanics, it is also important to note that quantum tunneling is applicable to a number of real world processes including fusion in the sun and the workings of the scanning tunneling microscope.

Quantum tunneling refers to any phenomenon where a particle is able to penetrate a potential barrier that it wouldn't classically have enough energy to overcome. This is one of the counter-intuitive aspects of quantum behavior. At very small time scales, uncertainties in energy become large, allowing particles to pass through otherwise insurmountable barriers. Tunneling is responsible for alpha particle decay in which a helium nucleus 'leaks out' of a larger nucleus bound by the strong force. Tunneling also plays a role in allowing protons to overcome their electric repulsion during fusion. Condon (1978) provides a historical account of the origins of tunneling theory.

6.1 Previous research on quantum tunneling

The specific ways in which students' difficulties manifest themselves in interviews and assessments on the topic of quantum tunneling have been documented in a series of mainly qualitative studies. These studies have uncovered misconceptions pertaining to three general areas: energy, probability, and graphical representations.

Energy

By far the most widely reported misconception students hold about tunneling is that energy is lost as the particles pass through the barrier. More precisely, students believe that particles that have tunneled through a barrier have lower energies than those in the incident beam. This has been documented in numerous studies (Wittmann, Morgan & Bao 2005, Ambrose 1999, Wittmann 2003, Morgan et al. 2004, Redish et al. 2000) and has been termed the 'incorrect' energy response. Most would agree with Ambrose (1999) that this misconception is rooted in students' classical notions of particles penetrating physical barriers. In order to break through such barriers, energy must be lost.

Another common misconception is that total energy is lost in the barrier, but regained when the particle emerges on the other side. This is called the 'inconsistent' energy view (Redish et al. 2000).

Asking students what effect changing the barrier width or height would have yields further insight into students' mental models. Morgan et al. (2004) found that more students thought a wider barrier would induce a drop in energy than a taller barrier. "This suggests an analogy to macroscopic tunneling; it does take more energy to tunnel through a wider mountain, but does not take more energy to tunnel through a higher mountain" (p.4).

The concept of energy as it relates to quantum tunneling is much subtler than in any context students have previously encountered. "Students must separately track total, potential, and local kinetic energies," (Redish et al. 2000, p.7) often leading to confusion among these facets of energy (Ambrose 1999). It is therefore not surprising that students have consistently reverted to classical mental models when their new quantum mechanical models are both complicated and underdeveloped.

Probability

It has been demonstrated that students have less difficulty producing the scientifically correct answer in regards to the probability of incident particles tunneling through a barrier, than in regards to their energy (Morgan et al. 2004). This is not necessarily indicative, however, of a deeper understanding of the probabilistic mechanisms involved in quantum tunneling. As noted by Morgan et al., the relationship between transmission probability and barrier parameters is more similar to the macroscopic case than that between energy and barrier parameters; the presence of a barrier reduces the transmission probability, as does an increase in either barrier dimension.

Students have been shown to consider only a limited number of aspects of probability when describing tunneling (Domert et al. 2005). Four categories of description have been proposed for understanding probability in this context:

1. probability in terms of reflection and transmission

2. probability in terms of having a threshold energy
3. probability in terms of finding a particle at a particular location
4. probability in terms of an ensemble of systems

Only two of these categories (1 and 3) were commonly used by any one student to explain tunneling (Domert et al. 2005).

Bao & Redish (2002) also argue students' lack of familiarity and proficiency with probability hinders their learning of quantum mechanics. They found that students demonstrated common misconceptions about probability in classical contexts and hence were unable to transfer useful probabilistic ideas to the novel setting of quantum mechanics. In order to help students learn quantum mechanics, they suggest first addressing probability in more detail in classical contexts before 'bridging' this knowledge to the quantum realm.

Graphical representations

Pictorial representations have proved both particularly difficult for students of quantum mechanics and an important, if complex, diagnostic tool for probing understanding. Potential energy, wave function, and probability density plots bring together the ideas of energy and probability in an entirely novel way for students. Given their lack of familiarity with these representations, some diagrams may in fact be a source of misconceptions. One must exercise caution when analysing students' drawings of wave functions and probability densities since it is difficult to differentiate genuine conceptions from uninformed guesses and graphing difficulties.

Misconceptions shown in graphical representations. Drawings are an important diagnostic tool as they explicitly represent students' conceptions in the standard representations used by physicists. If these sketches are made with care, they can be analysed to determine where student reasoning is coherent and where it is either weak or inexpressible. For example, Ambrose (1999) found that students who believed that a particle cannot exist inside a classically forbidden region drew no wave function in this area.

Misconceptions arising from graphical representations. Some misconceptions that students display are likely the result of common diagrammatic representations. Students often have difficulty clearly distinguishing between energy and probability of quantum particles (Domert et al. 2005). The source of this difficulty has been identified (at least in part) as the depiction of the wave function and potential energy diagram on the same axes used in most textbooks and often drawn in lectures (Morgan et al. 2004). With the vertical axis representing two things simultaneously (energy and wave function amplitude or probability density) students are more likely to see the potential barrier as a physical obstacle and mix up the symbols corresponding to energy and probability. As a result, students often superimpose the barrier on their wave function drawings when not asked to do so (Morgan et al. 2004). Some student wave functions seem to depict velocity or kinetic energy (Ambrose 1999). One student used conservation of energy arguments to explain why the area below the probability density curve should be constant (Domert et al. 2005). Another purported manifestation of this confusion is the 'axis shift' that often occurs when students draw a wave function for a tunneling beam of particles (see Figure 6.2, conception 4, Wittmann et al. 2005, Redish et al. 2000, Morgan et al. 2004, Wittmann 2003, Wittmann & Morgan 2004). The sinusoidal oscillations of the incident beam are sketched systematically higher than those of the transmitted beam, implying a correspondence between average wave function height and energy (these students typically also exhibit the 'incorrect' energy loss conception.

The pictorial representations of the wave function and probability density seem to cause students difficulties because these depictions are much less visually representative than those of classical particles (Johnston et al. 1998). Students often draw the wave function representing constant particle flux as a horizontal line offset from the x-axis instead of a sinusoidal curve (Ambrose 1999) because the latter is incongruous with traditional notions of 'constant flux.' Given students' mathematical background it is easy to see why it is troublesome that performing a simple mathematical operation, say taking the square modulus of the wave function, can yield a horizontal line from a sinusoid. Furthermore, the representations are complicated by interference effects; the probability density appears sinusoidal when there are interfering beams of particles, but the wave function is always sinusoidal when describing scattering states. When drawing the wave function or probability density, students sometimes conflate the two pictures (Wittmann & Morgan 2004). Interestingly, the picture that results from a 'correct' drawing of the wave function mixed with aspects of the probability density closely resembles the 'axis shift' picture mentioned above.

A common theme that arises in quantum mechanics education papers is the disconnectedness of student knowledge. Concepts are understood only within a very fragile framework and are poorly linked with other related concepts. This was most eloquently expressed by Johnston et al. (1998), building on a constructivist metaphor from Pines & West (1983). The metaphor relates knowledge constructions to plants growing up a trellis.

> Shoots growing upwards represent the knowledge that students construct for themselves from their own experience. The parent vine growing down towards them represents the agreed corpus of knowledge they aspire to learn. Mature learning occurs when the two intertwine. In that metaphor, the quantum mechanical mental models of the present students are slender tendrils indeed, completely unsupported by neighbours or the parent vine. (p.443)

Misconceptions arising from concept disconnectedness and graphing difficulties. Students seem to lack integration of the representations of the wave function and probability density with other fundamental concepts. This underlies a series of errors that commonly appear in student sketches. For example, although the misconception that energy is lost during tunneling is common, it is rarely expressed as an increase in wavelength. The amplitude to wavelength ratio is typically maintained when compared to the incident beam, rather than the wavelength proper (Ambrose 1999). This mistake is so easy to make that it exists in some textbooks. Whether this is evidence of students' genuine misconceptions (only Wittmann et al. 2005 has thus far hinted at some students' belief of an energy increase through tunneling) or simply graphical carelessness, it is an indicator that students lack an appreciation of the significance of aspects of their drawings. The fact that students conflate the wave function and probability density or relate the average position of the wave function to particle energy to arrive at the 'axis shift' picture also signifies the poor interrelations of these concepts in students' mental models.

Previous studies have sought to uncover how students conceptualise quantum tunneling and how learning can be facilitated through innovative teaching practices. These studies have typically employed interviews, quizzes, pre-post testing, and questionnaires, all with relatively small sample sizes ($n < 30$). Investigations have been carried out in America (Wittmann et al. 2005, Ambrose 1999, Morgan et al. 2004, Wittmann & Morgan 2004, Redish et al. 2000), and Sweden (Domert et al. 2005) and recent studies have focused on probability (Domert et al. 2005) and energy (Wittmann et al. 2005) conceptions. In this chapter, I seek to develop a comprehensive picture of students' conceptions of probability, energy and graphical representations as they pertain to quantum tunneling. The ultimate goal is to understand the most crucial and troublesome areas for students so that instructional design can evolve to meet students' needs.

To guide the investigation, two questions were posed:

1. What are students' conceptions of probability and energy as they relate to quantum tunneling?
2. Do there exist simple, explainable misconceptions that can be traced back to teaching methods or prior knowledge?

6.2 Method

6.2.1 Questionnaire

An instrument to test students' understanding of quantum tunneling was devised based on research by Redish et al. (2000), Morgan et al. (2004), Ambrose (1999), Fletcher (2004), and my observations of the intermediate quantum mechanics lectures (Muller 2005). The questionnaire consisted of six questions and is included as Appendix B.2.

The first five questions of the questionnaire were based on the standard arrangement of a beam of mono-energetic electrons incident on a square potential barrier. This question has been used in many studies and frequently appears on undergraduate quantum exams (Redish et al. 2000, Wittmann 2003, Wittmann & Morgan 2004, Ambrose 1999, Morgan et al. 2004). Variations on this theme are common; for example, Domert et al. (2005) interviewed students during and after interacting with a computer simulation in which a single wave packet was incident on a barrier.

In question a) of this questionnaire, students were asked to comment on the kinetic energy of the electrons in the three regions. The wording of this question differs from previous studies in that it asked specifically about kinetic energy and it required students to consider the energy in Region II. Kinetic energy was chosen in place of total energy for several reasons: kinetic energy relates directly to the wavelength of the wave function to be drawn in the next question (although students often have difficulty making this connection), and students are more familiar with kinetic energy than with total energy. Contexts with which students are more comfortable are more likely to elicit their genuine conceptions (Yeo & Zadnik 2001). Students were asked to explicitly rank the kinetic energies in all regions rather than simply compare the energies of the incident and transmitted beam. This gives a more complete picture of students' conceptions of energy during the tunneling process.

Questions b) and c) asked students to sketch the wave function and probability density of the electron beam. Similar questions have been asked in interviews (Wittmann & Morgan 2004, Wittmann 2003, Morgan et al. 2004), on surveys and pretests (Ambrose 1999, Redish et al. 2000), and are commonly used as examples in lectures. During the lecture audit, I observed the solution to these questions worked out on the board approximately four weeks before the survey was carried out.

The ability to understand energy diagrams, plots of wave functions and probability densities and the relationships among them has been identified as an important and difficult skill essential to constructing scientific conceptions (Bao & Redish 2002). Students' sketches of the wave function and probability density allowed for comparisons among the first five questions to determine the internal consistency of responses.

In questions d) and e) students were asked to predict the effect of a change to barrier height or width on the probability of electron transmission and energy. This question was previously used by Morgan et al. (2004).

Question f) asked students to apply their knowledge of the tunneling process to explain the phenomenon of alpha decay. This topic was briefly discussed in lectures as an example of how tunneling applies to the real world.

6.2.2 Procedure

Sixty-four students taking the second year quantum mechanics course were surveyed informally during experimental laboratory at the conclusion of lectures and before the exam study period commenced. In first year, these students received twelve hours of quantum mechanics lectures. The second year lecture course comprised nineteen hours of lectures with an additional nine two-hour computational laboratories. Regular and Advanced streams attended common lectures for the majority of the semester, with approximately five periods in which the Advanced class received instruction in greater depth. It should be noted that at the time of surveying, a fraction of the Advanced class was working independently on laboratory projects and therefore was not surveyed. Students were asked to complete the questionnaire individually and were informed that it would not be used for assessment. It was indicated however that the questions might be similar to those on the final exam. Students took between ten and twenty minutes to complete the questionnaire, but no time restriction was enforced. This allowed students to demonstrate their knowledge as completely as they desired.

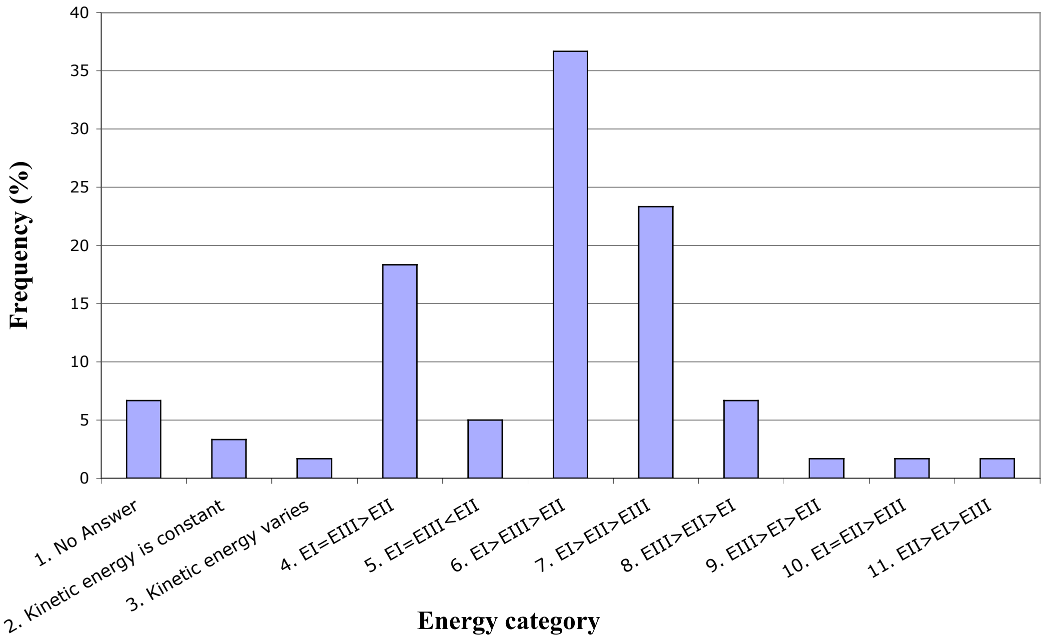


Figure 6.1: Frequency of energy ranking categories.

6.3 Results and analysis

6.3.1 Energy

Question a) The ranking of kinetic energies of a beam of tunneling particles lent itself particularly well to a categorical analysis as there were a finite set of possible answers.

As shown in Figure 6.1, student explanations of the variations in kinetic energy of the electron beam fell into eleven categories. Of these eleven, the five most common non-trivial explanations (categories 4–8) were interpreted as genuine and significant student conceptions. Based on the energy rankings in these categories they were labelled as shown in Table 6.1.

Category	Conception label
4. $E_I = E_{III} > E_{II}$	1. Barrier as 'quantum mechanical entity'
5. $E_I = E_{III} < E_{II}$	2. Barrier as 'well'
6. $E_I > E_{III} > E_{II}$	3. Barrier as 'classical obstruction plus potential'
7. $E_I > E_{II} < E_{III}$	4. Barrier as 'classical obstruction'
8. $E_{III} > E_{II} > E_I$	5. Barrier as 'energy screen'

Table 6.1: Student conceptions of kinetic energy during tunneling

4. Barrier as 'quantum mechanical entity' ($E_I = E_{III} > E_{II}$)

This category is regarded as the 'correct' answer as it is the conception that most physicists would possess. On either side of the barrier, the energy of the electron beam is given by E and therefore the kinetic energy of the beam in both regions must be equal. In Region II the energy of the electrons is less than that of the potential barrier and therefore, the electrons' kinetic energy must be lower than that of either of the neighbouring two regions.

5. Barrier as 'well' ($E_I = E_{III} < E_{II}$)

Students who indicated that the kinetic energy of electrons on either side of the barrier was equal but less than that of electrons in Region II likely misinterpreted the potential energy diagram. Students may have viewed the potential energy step as a barrier for positive particles and therefore a well for electrons, despite the wording of the question. This conception would be appealing for students due to their familiarity with well problems. Some of these students did, in fact, use the word 'well' to describe the barrier region in the short answer space in question b).

6. Barrier as 'classical obstruction plus potential' ($E_I > E_{III} > E_{II}$)

By far the largest fraction of the class expressed the belief that energy was lost during the tunneling process. Interestingly, students indicated that the kinetic energy was lowest in the barrier. This indicates that some correct reasoning from the potential energy diagram was then hybridised with more classical notions of energy dissipation through barrier penetration.

7. Barrier as 'classical obstruction' ($E_I > E_{II} < E_{III}$)

As the second most common response, the 'classical obstruction' view resembles a classical mechanics interpretation. The kinetic energy is highest in the incident beam, dissipated in the barrier, and therefore lowest in the transmitted beam. This conception arose saliently in an interview by Morgan et al. (2004). "One interview subject discussed her mental picture of snowballs flying through snow banks when she thought about tunneling" (p.4).

8. Barrier as 'energy screen' ($E_{III} > E_{II} > E_I$)

Perhaps the most surprising conception, barrier as an energy screen views the potential barrier as a selector for the highest energy electrons. Despite the fact that the incident beam was explicitly defined as "mono-energetic," students seem to have attributed a range of energies to the incident electrons. It has been noted by Ambrose (1999) that students don't identify mono-energetic beams with well-defined wavelengths. The fastest electrons are therefore the only ones capable of penetrating the barrier. This would be analogous to rolling balls up a hill; balls with high kinetic energy would travel furthest and only those with the highest energy could make it over the hill.

6.3.2 Wave functions

The general categories that emerged from the wave function data are summarized in Figure 6.2. Within categories, some variation exists; the example column depicts a typical response for each category.

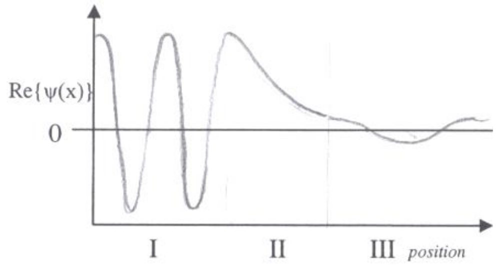
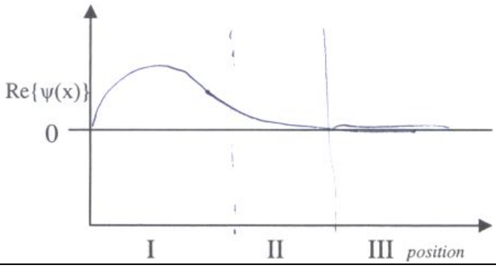
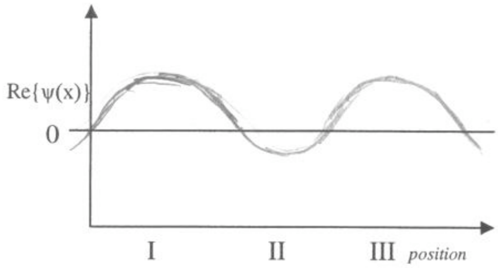
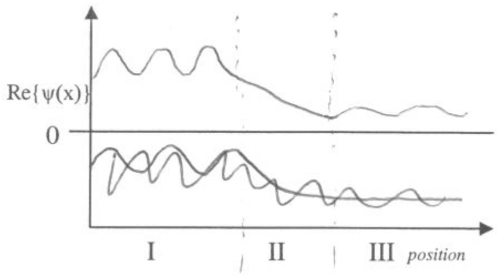
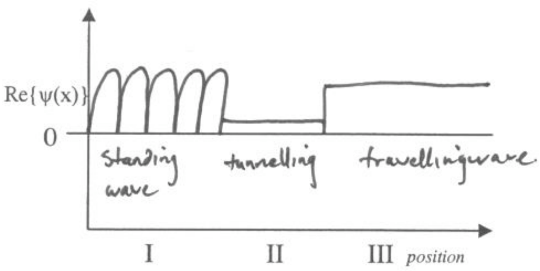
Category	Frequency (%)	Example
1. Standard (Most features consistent with a 'correct' wave function)	28	
2. Single maximum (One large maximum, centred over Region I or II)	9	
3. Large sinusoid (Typically 1.5 wavelengths, negative values in Region II)	13	
4. Conflated wave function and probability density (Some features of a 'correct' wave function plus features of probability density)	28	
5. Other	11	
6. Blank	11	

Figure 6.2: Summary of student wave function categories.

1. Standard

The largest fraction of students drew a wave function that closely resembles those found in textbooks and quantum lecture notes. Despite this encouraging statistic, few drawings were entirely correct. Most featured shorter wavelengths in Region III than in Region I. Students who did sketch equal wavelengths typically described their drawing as "I: amplitude high, II: amplitude falling, III: amplitude small; wavelengths equal in I and III."

2. Single maximum

With a maximum in Region I or II, these students' drawings look like wave functions of particles trapped in finite square wells. Some students seemed to view the y-axis as part of the barrier. To describe his drawing (shown as example 2 in Figure 6.2) one student wrote, "[half of a] wavelength is just larger than the length of I," indicating the significance of this region in determining the wavelength. This phenomenon has been reported elsewhere (Wittmann & Morgan 2004) with students making analogies to alpha decay. Another conception that was manifested in this category was that particles were slowed by the barrier and therefore were more likely to be found there. Some representative explanations are: "higher probability in high potential energy regions, travels slower, less kinetic energy," and "amplitude is highest in barrier, wavelength is largest in barrier since it slows down and $p = \frac{h}{\lambda}$."

3. Long wavelength sinusoid

This category is characterized by positive wave function values in classically allowed regions and negative values in Region II. This specific type of drawing has not been reported elsewhere in literature but certainly implies a correspondence between kinetic energy and the value of the wave function. For example, one student wrote, "amplitude should be smallest within the well, relatively the same outside." This student seemed to confuse the value of the wave function with its amplitude since she drew positive curves in Regions I and III and a negative curve in the barrier.

4. Conflated wave function and probability density

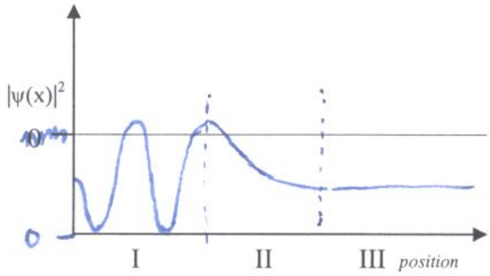
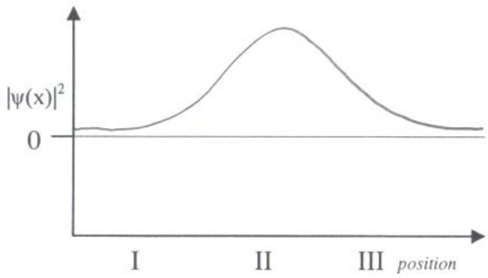
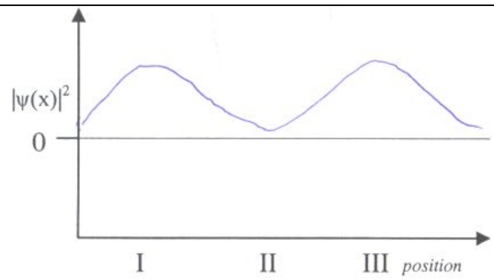
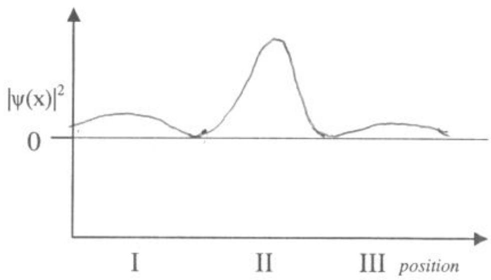
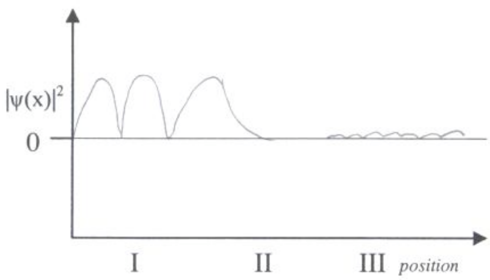
Drawings of this type were as common as 'standard' drawings and included the 'axis shift' response. One student who drew this type of picture seemed again to confuse the amplitude of the wave function with its value. She wrote, "Wave starts off with a pretty big amplitude (Region I), then looks like an exponential function (Region II) then ends up with a small amplitude." Another set of drawings that were included in this category depicted the wave function in Region III as a horizontal line characteristic of probability density. A student who gave this answer indicated that it was only interference that produced the sinusoid in Region I. "In region I there will be reflection, therefore get a wavelength from interference. In region III, no reflection therefore constant amplitude." This student clearly seems to be conflating the ideas of probability density with the wave function. Similarly, another student used probability density terminology to describe his wave function that was a horizontal line in Region III. "Region I: there is an incident and reflected wave giving a 'standing' travelling wave interfering, hence a varying amplitude. Region II: an evanescent 'wave.' Region III: a travelling wave with no interference."

6.3.3 Probability density

A similar approach was taken to analysing the probability density plots from question c). General categories were again arrived at by grouping the data according to overall similarities. These are described in Figure 6.3.

1. Standard

Less than ten percent of students successfully produced a qualitatively correct drawing of the probability density. Characteristics of this answer were a sinusoid in Region I, decaying exponential in Region II and horizontal line in Region III.

<table>
<tr><th>Category</th><th>Frequency (%)</th><th>Example</th></tr>
<tr><td>1. Standard (Most features consistent with a ‘correct’ probability density plot)</td><td>9</td><td>
</td></tr>
<tr><td>2. Single maximum (One large maximum, centred over Region I or II)</td><td>23</td><td>
</td></tr>
<tr><td>3. Double maxima (Two maxima, centred over Regions I and III)</td><td>19</td><td>
</td></tr>
<tr><td>4. Triple maxima (Three maxima centred over the three regions)</td><td>6</td><td>
</td></tr>
<tr><td>5. Rectified sinusoids in Regions I and III</td><td>13</td><td>
</td></tr>
</table>

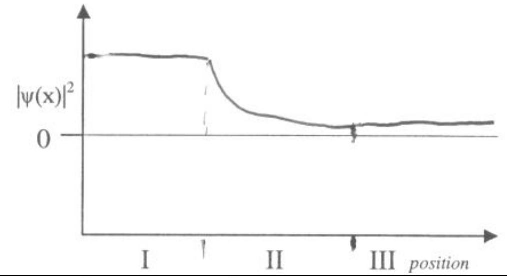
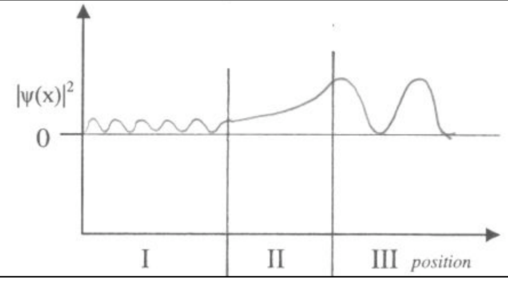
6. Standard without interference in Region I	9	
7. Other	14	
8. Blank	6	

Figure 6.3: Summary of student probability density categories.

2-4. Maxima

The single, double, and triple maxima responses are quite surprising since they have not previously appeared in published studies but account for almost half of responses in this survey. These drawings seem to imply ‘well’ misconceptions or the belief that the particles slow down (and cluster) in Region II as explained in the single maximum wave function case. Further investigation is required to fully understand student mental models that give rise to such answers.

5. Rectified sinusoids

This answer is also unique to this study. Typically, students in this category have taken the absolute value of their answer from b) as being the square modulus. One hypothesis for the cause of this was that students viewed the modulus squared operation as a simple square, giving rise to the rectified sinusoids. However, in discussions with the lecturer for this course, it emerged that the CUPS simulation program used in lectures and computational laboratory depicts the modulus of the wave function and not probability density. This gives rise to the discontinuities in the first derivative where a sinusoid would be expected. A question asked by one student during lectures expressing his confusion over when the slope of the probability curve is continuous. The lecturer interpreted this as a question about the modulus of the wave function, not the square modulus, and replied that the continuous case becomes discontinuous when there is very little transmission and hence a standing wave exists in Region I.

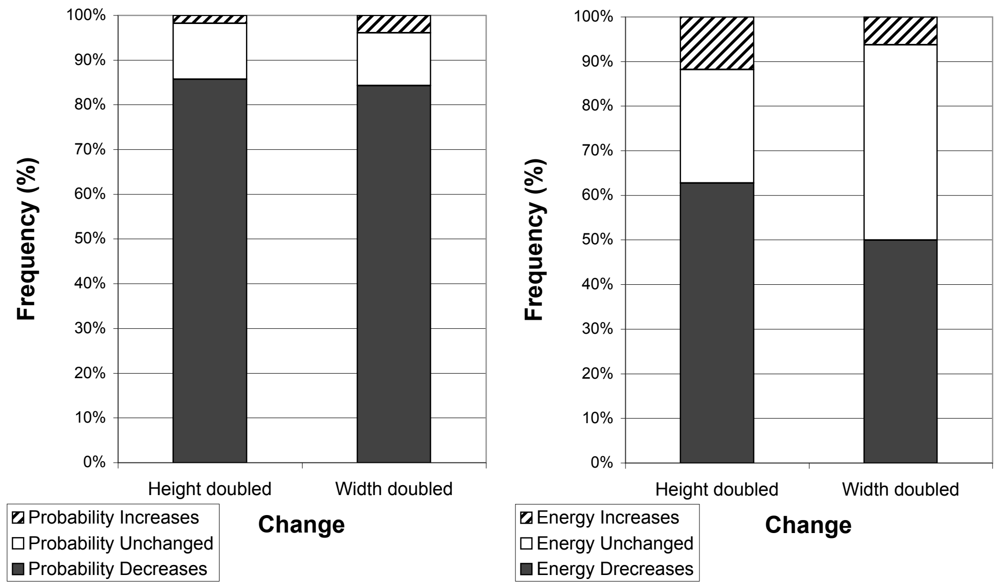
6. Standard without interference

Drawings in this category are similar to those classified as 'standard,' however they exhibit no interference between the incident and reflected beams in Region I. This could be an indication that students are failing to attribute wave properties, such as interference, to particles or that they have simply forgotten about the reflected beam.

6.3.4 Changing barrier parameters

Questions d) and e) dealt with the relative probabilities and energies of the transmitted beam through barriers of different dimensions. Students' answers were initially classified into a set of categories maintaining the association of the energy answer with its corresponding probability. Due to some vague answers and the already high number of possible solutions, a large range of categories was produced. To obtain more meaningful insight into the data, probability and energy answers were considered separately, yielding only three categories for each property.

From Figure 6.4, it is evident that students' conceptions of probability are more stable and 'correct' than their notions of energy. This is consistent with the finding of Morgan et al. (2004) that students are more likely to answer correctly about probability than energy when barrier parameters are altered. It is interesting to note that students were more likely to believe that transmitted electrons would have a different energy with a doubling of barrier height rather than a doubling of barrier width. This is opposite to the prediction made by the macroscopic analogy suggested in



(a) conceptions of probability (b) conceptions of energy

Figure 6.4: Student conceptions of tunneling when barrier parameters are changed.

Morgan et al. (2004), wherein a wider barrier induces a greater change in energy than a higher barrier.

6.3.5 Application to alpha decay

Question f) seems to have been the most difficult for students as only 38% made a serious attempt to answer. Students' fragmented knowledge is implicated by their lack of confidence in applying their understandings to real world contexts. Those who did make an attempt typically discussed a potential barrier through which the alpha particle must tunnel. They did not, however, explicitly mention the source of this potential. Students also discussed the varying widths and heights of potential wells though much of these explanations seemed to build on questions d) and e). A sample of the range of answers is given below:

"Dependent on the size of the radioactive element: the more tightly bound, the bigger the potential hill, therefore less probability of alpha tunneling, hence the decay constant would be larger."

"Binding forces between the alpha particle and nucleus act as a barrier. Different elements have different binding forces, thus probability of barrier penetration differs for each element for alpha particles of the same energy."

"Varying atomic potentials give different potential heights and widths which drastically affect the tunneling probability (and thus the half life)."

"Different particles have different potential wells which affect the decay constants."

6.4 Discussion

At the beginning of this study, two questions were posed to help guide this investigation and understand the learning that takes place in the School of Physics. It was hoped that through sampling a large number of students on a different continent to other related studies, common international misconceptions could be uncovered, providing insight into their origins. The survey predictably turned up many of the same conceptions that have been illuminated in previous interviews. Energy misconceptions, confusions of probability and energy, and difficulties with diagrammatic representations that have been reported elsewhere were observed in the local context. However, several prevalent misconceptions that are unique to this study were also identified. Thus, although much of the results can be understood in the context of international studies, the novel conceptions arising here must be further investigated.

This survey again confirms the most common misconception students possess: total energy is lost through quantum tunneling. Over sixty percent of students reported the 'incorrect' view that the energy of electrons in Region III was less than that in Region I. As the question was phrased in terms of kinetic energy, it was impossible to identify students who held the 'inconsistent' energy view (in which total energy is lost in the barrier but regained in Region III). Students who indicated that the kinetic energy was least in the barrier but regained in Region III were perceived as possessing the correct energy view. This brings to light an engaging possibility: as students are more comfortable with kinetic energy, isn't it likely that the 'inconsistent' view is really a manifestation of the 'correct' view when students confuse total and kinetic energy?

A possibility that is worthy of investigation regarding the 'incorrect' energy view is: are students conceiving of the total energy of the beam and not its constituent particles when thinking about the relative energies in the incident and transmitted regions? The beam itself is attenuated and therefore its total energy is reduced in the tunneling process.

The wave function is clearly a problematic representation for students because it "cannot in any sense be considered to 'look like' the object. It is merely a device for predicting the object's physical behaviour" (Johnston et al. 1998, p.431). From the wave function drawings for a beam of tunneling particles, students were divided into three roughly equal groups: the standard picture, the conflated picture, and something else entirely. Previously the conflated picture, which includes the 'axis shift' drawing, has been interpreted as a depiction of energy loss and only once considered a muddling of the wave function and probability density (Wittmann & Morgan 2004). Researchers have advanced the 'axis shift' picture as a consequence of the superimposed wave function and potential energy diagrams commonly used in instruction. In this study, however, it seems far more likely that this prevalent conception is due to the combination of probability density ideas with the wave function. This is evidenced by students' use of interference to account for the sinusoidal oscillations of the wave function and by the apparent confusion between amplitude and average height above the x-axis. Another observation that is important to note is that students are much more likely to use probability density features in their wave function drawings than vice versa (in fact no students made this error). The obvious explanation for this phenomenon is that the probability density looks more like what it is representing than the wave function. It is more comprehensible and tangible in terms of the probability of finding a particle. Hence, probability density is more appealing and concrete for students. This explains why students believe the amplitude of the wave function corresponds not to the amplitude of oscillations but the average offset of the wave function from the x-axis. This offset represents the particles in that region; the larger the offset, the more particles in the beam.

Student drawings of the wave function and probability density were scanned for the misconceptions summarized in the introduction. Only one student drew a picture of probability density that was non-zero in Regions I and III but non-existent in the barrier region. This phenomenon, which was found in (Ambrose 1999), indicates the reluctance of some students to believe that particles can exist in classically forbidden regions where conservation of energy is apparently violated. An unexpected and interesting feature of some students' drawings of probability density was a line with slightly negative slope in Region III, rather than a horizontal line. To explain this behaviour, students used phrases like "amplitude tends to decrease to right," or "dies in III." This again seems to affix classical particle behaviour to the tunneling electrons.

6.5 Conclusion

Quantum mechanics teaching has been maligned by some as unchanged since the time of its inception over seventy years ago (Fletcher 2004), however in the wake of new technologies it is certainly evolving and incorporating innovative methods (Muller 2005). Despite these changes to instruction and the undoubted variety of methods across languages and continents, similar misconceptions have arisen in all studies pertaining to quantum tunneling. Through a survey of a large group of intermediate physics students, this study confirmed several common misconceptions pertaining to energy, probability and graphical representations. Some of these misconceptions have been examined from a different perspective with the help of students' explanations. In addition, new misconceptions were uncovered particularly in regards to pictorial representations. These merit further qualitative study. Overall, the conceptions revealed in this study paint a picture of student understanding in the area of quantum tunneling that is fragmented, inconsistent, and unstable. It is no wonder that students are clinging to classical misconceptions and confusing probability and energy or conflating probability density with the wave function. These seemingly distinct entities are shrouded by quantum mechanics in novel mathematical and graphical formalisms. Students' conceptions are, as yet, very green shoots, grasping at footholds in a landscape rich with misconceptions and alternate paths.

Chapter 7

Learning theories

In Chapter 3, I outlined a few related models that describe how learning can occur with the presentation of words and pictures. Examining the current teaching and learning practices in physics, however, it is clear that learning doesn't proceed as smoothly as might be expected. Motivated by these results, in this chapter I take a second look at the learning process through the lens of constructivism, the dominant science education ideology of the past three decades.

The finding that students can complete a lecture course and even perform well on final examinations while remaining ignorant of some of the fundamental concepts of the course is unfortunately common in physics education. Students do not necessarily lack ideas about these central concepts, they just have alternative ideas to those presented in class. When teachers and educational researchers began realizing this in the 1970's, the epistemology of constructivism was ushered in. Constructivism is said to counter the notion that learning is a transmission process, explaining why students fail to 'get' the ideas presented to them. In this framework, students are the active architects of their own knowledge, so it should be unsurprising that, depending on their pre-existing knowledge structures, they come up with ideas that do not directly match those presented in class.

In this chapter, I provide an outline of constructivism and discuss some of the outcomes of related research programmes. I also highlight some of the weaknesses of constructivism and consider how various interpretations of the theory lead to conflicting implications for instructional design. Stemming from the observations of teaching and learning practices, I consider the implications of social constructivism and other social learning theories on the design of multimedia. These ideas then give rise to the experiments conducted in Chapters 8–10.

7.1 Constructivism

What is constructivism?

Constructivism is an epistemology that comes in a range of forms, but all assert that human knowledge and the methods by which it is created are constructions. This means that babies are not born with knowledge or epistemological criteria, nor can they experience the world in such a way that allows for a direct internalization of 'objective reality.' Instead, they must undergo a series of experiences that they interpret in their own unique way, developing knowledge structures with each new experience.

Constructivist learning is most often explicated in stark relief to what it is not: a direct transmission of ideas from the textbook or the mind of the teacher to the mind of the student. It is debatable whether anyone really takes this view of learning but this contrast provides a starting point for understanding how a constructivist views a learning scenario. The teacher creates a learning environment that she believes will give rise to meaningful experiences for the students, based on her prior knowledge. The students then engage with the environment, attempting to make meaning of their interactions with it in light of their prior experiences.

This view of learning brings into focus several aspects of the learning process. First, it is clear that the actions of the learner are more important for successful learning than those of the teacher. Second, each student's conceptions are unique since they are individually constructed and there is no mechanism for directly comparing one's concepts with another's. Third, the prior experiences of the learner are central to the effectiveness of the learning experience. Not only are prerequesites essential, conflicting ideas must be taken into account for they may inhibit learning. Finally, the idea that the learner is the main agent of learning implies that he or she must experience directly the phenomena of interest for education to have its greatest effect.

What have been the outcomes of constructivist research?
Many successful research programs evolved based on the constructivist framework. Broadly, three main objectives have occupied constructivist researchers: documenting alternative conceptions, theorizing about their nature and origins, and attempting to change them. These three objectives are discussed in the following sections.

7.1.1 Documenting alternative conceptions

Copious amounts of literature document the preconceptions or naïve frameworks of students as they enter physics classrooms (Duit 2007). If these ideas persist through instruction, or if new but unscientific ideas arise, then they are called misconceptions or alternative conceptions. Ardent constructivists argue that the latter is a better term because misconceptions may be entirely plausible and fruitful for the person who holds them. In this thesis, the terms are used interchangeably because I believe their meanings to be unambiguous, regardless of one's theoretical position.

Perhaps the topic on which students' conceptions have been studied most often, and in the greatest depth is kinematics and Newtonian mechanics. Since students are exposed daily to macroscopic objects and their motions, Newtonian mechanics is the area of physics students have the most experience with and therefore preconceptions of before entering the classroom. The following discussion of alternative conceptions pertaining to Newtonian mechanics will serve as useful background for Chapters 9 and 10.

A comprehensive study of students' concepts of velocity and acceleration was conducted by Trowbridge & McDermott (1980, 1981). The researchers interviewed over 300 introductory physics students individually as they worked through a series of demonstration activities. Many students had some notion of distance, position, speed, velocity, and acceleration, though they had significant difficulty applying these ideas consistently. For example, students often expressed the belief that when two rolling ball bearings were side by side, their velocities were equal. Furthermore, when one ball bearing was in front of the other, many students said it was going faster. Similar, but perhaps more pervasive difficulties were observed with acceleration. Students commonly failed to differentiate acceleration from velocity, despite substantial assistance from the interviewer. Many could state the textbook definition of acceleration but could not apply it correctly. When one ball bearing was catching up with another, students often said it had a greater acceleration; they did not consider that it might just have a greater velocity. Researchers termed these vague and seemingly inconsistent student notions 'nondifferentiated protoconcepts.'

One of the most common misconceptions studied, and perhaps most deeply held pertains to objects that, due to the force of gravity, slow down, reverse direction, and then accelerate. Examples include pendulums, ball bearings rolling up and down ramps, and projectile motion.

Many have studied novices' ideas about these phenomena (e.g. Trowbridge & McDermott 1981, diSessa 1982, Clement 1982, McClosky 1983, diSessa 1996), which are often some version of what is called 'impetus theory.' This theory dates back to Aristotle, but has been common in various guises among thinkers up until the time of Galileo and Newton. The 14th century philosopher Buridan succinctly articulated the central tenets of impetus theory:

> When a mover sets a body in motion, he implants into it a certain impetus, that is, a certain force enabling the body to move in the direction in which the mover starts it, be it upward, downward, sideward, or in a circle. It is because of this impetus that a stone moves on after the thrower has ceased moving it. (Buridan as cited in McClosky 1983, p.123)

McClosky found that many of his students both articulated common misconceptions about motion, and employed these ideas when attempting to perform tasks. For example, students were asked to walk across a room and drop a golf ball on a target marked on the floor as they walked. Almost half of the students released the ball directly above the target, expecting it to fall straight down. In another experiment, one quarter of students attempted to slide a puck in a curved path by moving it in an arc before releasing it. These experiments suggest that the problems encountered by students are not just semantic; they do believe what they say they believe and act accordingly. Furthermore, these studies highlight potential practical detrimental effects of retaining intuitive ideas about motion.

Once physics alternative conceptions were catalogued, researchers set about quantifying them. Studies sought to answer questions like: how common are misconceptions? Are they the same across cultures and languages? Do they arise with students of all ages and levels of education? Are they just extensions of general academic difficulty?

Halloun & Hestenes (1985) developed the first large-scale instrument, called the Mechanics Diagnostic test (MD), to assess understanding of Newtonian mechanics. The test was validated in an initial study with over 1500 students. Students scored around 30% on the pre-test in high school and 50% on entry to university, giving an idea of the extent of misconceptions at both levels before instruction. The test was shown to be a good predictor of course achievement, independent of other factors like mathematical competence.

After a series of studies including interviews, the MD developed into the Force Concept Inventory (FCI, Hestenes et al. 1992). This test underwent a further revision in 1995 (see Mazur 1997). Some educators and researchers have voiced concerns about the FCI, believing it does not measure a 'force concept' per se, but rather a loosely grouped set of ideas (Huffman & Heller 1995). Nonetheless, many teachers and researchers have found the test a useful tool in evaluating teaching and learning and many concerns have been alleviated by further research (Henderson 2002).

A separate test, the Force and Motion Conceptual Evaluation (FMCE, Thornton & Sokoloff 1998), was developed to perform a similar function to the FCI. In general, the FMCE uses simpler questions than the FCI and therefore has better face validity as a measure of Newtonian understanding. The two tests correlate to a very high degree, however.

7.1.2 Theorizing about alternative conceptions

Cataloguing students' misconceptions has been fairly straightforward, especially compared to the task of theorizing about their nature. On this issue, three main perspectives have emerged. One characterizes novices' alternative conceptions as coherent theoretical entities, while another claims they have very little, if any, systematicity. A third contends that it is not the coherent or fragmented nature of misconceptions that is important, but rather the way in which these ideas are categorized ontologically.

A striking feature of students' alternative conceptions is their resemblance to scientific theories from an earlier time. This 'recapitulation' of older ideas has led many to consider parallels between the history of science and the processes involved in conceptual change. Impetus theory is an obvious example but many other misconceptions resemble previous beliefs, like the idea that the sun goes around the earth. Carey (1986) believes it is remarkable not that students' ideas are difficult to change, but that they can be changed at all. "The reason that students' misconceptions are so resistant to tuition is that learning mechanics requires a theory change of the sort achieved by Galileo—indeed, even more than that achieved by Galileo, all the way from impetus theory to Newton" (p.1127).

The comparisons between student ideas and historical scientific theories have drawn criticisms because student conceptions are manifestly not as considered as scientific theories. However, these criticisms do not negate the possibility that student conceptions are internally consistent. Vosniadou (1994) studied young learners' conceptions of the earth, uncovering several well-defined, coherent models. These included square or circular flat earths, 'synthetic' models like earth as a hollow, spherical fishbowl, and the correct scientific model. Interview subjects each articulated only one of these views and were able to answer a range of questions consistent with this view.

These findings lead Vosniadou to propose that in early development children establish a 'framework theory,' which accounts for their ontology and epistemology. All new experiences are then interpreted in light of this framework. Misconceptions are learners' attempts to interpret new experiences within their framework in ways that do not align with scientific views. Through the learning process, learners' fundamental theories are modified and augmented. Changes to aspects of theories on the periphery of learners' cognitive structures are relatively easy to achieve. However, alterations to the central foundations of the framework theory are understandably very difficult to make.

In sharp contrast, diSessa (1982, 1996, 2006) argues that students' conceptions don't resemble scientific theories in their coherence or sophistication. Rather, he proposes, they can be modeled as little pieces of knowledge that require no further explanation, which can be applied across a range of contexts. He calls these smallest units 'phenomenological primitives' or 'p-prims' for short, because they are derived from simple sensory experience.

DiSessa notes that p-prims are difficult to express in language because they are much simpler than theories, beliefs, or even concepts, but he offers some examples called 'balance,' 'overcoming,' and 'dying away.' Each of these ideas can be applied to a range of different phenomena. Balance refers to a sense of equilibrium and stability, or perhaps equality of competing influences. Overcoming refers to any situation in which two influences interact, with the stronger one eventually 'getting its way' over the weaker one. Dying away is a common occurrence in nature; examples include echos dissipating and objects returning to a state of rest.

Using p-prims, diSessa explains impetus theory as follows. At the peak of the toss, there is some apparent stability with velocity being zero. The visually salient feature is that the ball is momentarily stationary. This, it is claimed, cues the balance p-prim. The change of direction at the top suggests a form of overcoming. Gravity provides the obvious downward force, but it must be overcoming some other influence that originally propelled the ball upwards. A force transferred from the hand to the ball at the beginning of the toss is the default candidate. The dying away p-prim can then be used to explain why the hand force decreases as the ball travels upwards, such that it balances gravity at the top, before gravity overcomes this force and the ball accelerates downwards.

Chi and colleagues (Chi 1992, 2005, Chi & Slotta 1993, Chi et al. 1994, Slotta, Chi & Joram 1995, Reiner, Slotta, Chi & Resnick 2000) take a somewhat different view of misconceptions, asserting that appropriate ontological categorization is the central obstacle to conceptual change. In this model, learners are said to possess at least three ontological trees: matter, processes, and mental states. Within each tree, categories are broken down into subcategories; for example, matter is broken down into natural and man-made objects. Further assumptions are that learners have a strong sense of ontology that governs their thinking about the world, and that some physical processes involve entities that are readily miscategorized.

Chi argues that robust misconceptions involve the assignment of entities to incorrect ontological categories. For instance in physics, processes such as force and current are frequently classified by learners as matter. This gives a simple explanation why misconceptions like 'current is stored in the battery' and 'force is used up,' are common. Conceptual change then involves a conscious shift of entities from one ontological category to another. The greater the magnitude of this shift—for example between rather than within ontological trees—the harder the conceptual change is to achieve. In addition, if the target category is not available or poorly formed, it will have to be better established before the conceptual change can take place.

The ontological view of misconceptions combines aspects of the theory view and 'knowledge in pieces.' In order for learners to have a well-developed ontology, they must have an established framework to delineate category boundaries. On the other hand, the ontological view acknowledges that learners' conceptions are not nomological in the same sense as scientific theories. That is, they do not represent a set of principles from which generalizations to new settings can readily be made.

7.1.3 Conceptual change

Applying an understanding of concepts and conceptual change to the design of effective learning environments is ultimately the most important challenge for educators. This challenge has been addressed both theoretically, based on the views outlined above, and practically using teaching interventions in authentic classroom settings.

Theoretical models

It is important to first discuss how constructivist theory describes learning in general because conceptual change should be understood in relation to more common forms of learning. Piaget (1970) developed one of the most popular accounts of learning involving two related processes which he called called *assimilation* and *accommodation*. Assimilation refers to the incorporation of new sensory experiences by existing cognitive structures without the alteration of those cognitive structures. Thus assimilation may involve the modification of stimuli to fit with existing constructions. By contrast, in accommodation, existing schemas are adjusted to better match stimuli. In practice, both assimilation and accommodation take place in parallel to adapt the learner to his or her environment.

This model is similar to others proposed by Ausubel (1968) and Norman & Rumelhart (1975, Norman 1976). Ausubel also used the term assimilation but in a way that encompasses both processes envisioned by Piaget. In Ausubel's assimilation process, a new idea is interpreted and assimilated by an established cognitive structure. Both the new idea and the cognitive structure are modified as a result, forming what is called an *ideational complex*. Over time, these structures dissociate and can be used independently of each other.

In Norman & Rumelhart's (1975) model, three processes are involved in learning: *accretion*, *tuning*, and *restructuring*. During accretion, new information is interpreted and stored by pre-existing schemata, similar to Piaget's assimilation. Tuning refers to the slow process of modification undergone by schemas through their continued use. Small changes are proposed to occur periodically to schemas so that they become better and better adapted to the environment. When a schema cannot be modified only slightly to fit with experience, a more dramatic form of schema evolution, restructuring, must occur. This may involve a significant alteration of existing schemata or in some cases the creation of an entirely new schema. It is this most dramatic shift that is usually defined as a conceptual change.

One of the most cited frameworks for understanding conceptual change is that of Posner, Strike, Hewson & Gertzog (1982). They propose that four conditions are necessary to achieve conceptual change: (1) learners must experience a dissatisfaction with their existing conceptions; (2) new, intelligible conceptions must be available; (3) these conceptions must also seem plausible; and (4) they must provide fruitful ways of both solving new problems and understanding old phenomena.

This model is consistent with a view of learner conceptions as theoretical. The above conditions assume, for example, that learners are aware of their existing conceptions and their consequences, and they are capable of judging them to be inadequate. A later revision of the theory gave more attention to motivational and affective aspects of conceptual change (Strike & Posner 1992).

Chi and colleagues similarly believe some nature of explicit dissatisfaction with existing categorization schemes precedes conceptual change. Learners must see the ways in which their current classifications are inadequate and consciously reorganize their ontological assignments.

Others working with the coherent theoretical view of student conceptions believe argumentation is vital to conceptual change. Students' ideas should be elicited and used to make predictions about an experiment, with comparisons made to the predictions of scientific theories. Resolution is then achieved by performing the experiment and discussing how different theories account for the observed results. This method leads to a 'gestalt' view of conceptual change—learners' ideas should transition instantly to the correct scientific view.

The theory view is complicated, however, by the idea of 'incommensurability' from the history and philosophy of science (Kuhn 1996). This idea holds that some competing scientific theories are not only incompatible, but also incommensurable in the sense that the concepts or assertions of one theory cannot even be described in the language of the other. In such cases the notion that two theories can be judged by comparing their predictions to the outcomes of experiment must be abandoned. The two theories will likely disagree on which phenomena are relevant and which methods are valid.

It could be argued that conceptual change is difficult for learners for the same reasons revolutionary science is so problematic for scientists. Students' concepts and theories may be entirely incommensurable with those of science, making arguing for scientific views futile.

Proponents of the ‘knowledge in pieces’ view dispute the recommendations of confrontation, argumentation, and replacement of misconceptions, products of what they call the ‘standard model’ of conceptual change (diSessa & Sherin 1998). They claim that theoretical and ontological assumptions about the nature of alternative conceptions may seem reasonable from an expert perspective, but they do not necessarily reflect student cognitive architecture. If, instead, misconceptions are the result of a momentary arrangement of p-prims, attempts at confrontation, argumentation, and conceptual replacement will prove fruitless. A slight change in context might alter the p-prims cued or the way in which they are related.

As a competing strategy for conceptual change, diSessa and colleagues propose that p-prims be used productively to construct correct scientific conceptions. In some areas, it is argued, p-prims provide a useful starting point for scientific conceptions. For example, in impetus theory the decreasing upward force can instead be interpreted as decreasing momentum. In cases where intuitive p-prims don’t align in any way with scientific theories, their realm of applicability must be clearly understood. According to Smith, diSessa & Roschelle (1993) this view of conceptual change better aligns with constructivist principles, in which new knowledge must be built on the basis of old knowledge.

Of course some pedagogic methods are not committed exclusively to one theoretical position or other. As diSessa (2006, p.14) points out,

> both adherents of knowledge in pieces and of theory theories advocate student discussion, whether to draw out and reweave elements of naïve knowledge, or to make students aware of their prior theories in preparation for judgement in comparison to instructed ideas.

Practical methods

Existing approaches to teaching and learning have clearly failed to achieve the desired learning outcomes of teachers and educational researchers. Although some alternative conceptions are significantly reduced by standard instruction, it is often found that more entrenched misconceptions are much less affected. For example McCloskey (1983) found that traditional instruction reduced the frequency of the impetus misconception from 93% to 80%. At the high school level, Halloun & Hestenes (1985) found that gains on the Mechanics Diagnostic test were rarely greater than 20%. Pre-test scores ranged from 30–44% with post-test results around 52%. In college, gains were no better, with averages between 50–53% before instruction and 64–65% after the lecture course. These results were quite independent of the lecturing style or student perceptions of instruction.

Compared to these dismal results, many innovative teaching strategies, or 'reform methods,' have demonstrated dramatic improvements. Unfortunately, research underpinning new teaching strategies has been theoretically simplistic and therefore has done little to advance theory (diSessa 2006).

Innovative teaching strategies have been developed by many different physics education research groups, though the methods often share a degree of similarity. At the University of Washington, Tutorials in Introductory Physics were developed to involve students in small group discussions usually while conducting hands-on activities (McDermott & Shaffer 2001). A similar approach, called Workshop Tutorials, was designed and implemented at the University of Sydney (Sharma et al. 1999, Sharma et al. 2005). A research collaboration led to Interactive Lecture Demonstrations (ILDs), a specific sequence of demonstrations embedded in a predict-observe-explain cycle (Thornton & Sokoloff 1998). Time is also allocated for discussion of predictions among students and between students and lecturers before the demonstration is carried out. At Harvard University, a strategy called Peer Instruction was developed to encourage discussion among students and lecturers (Mazur 1997). Highlights of this approach are conceptual questions asked during class, which students must answer using personal keypads after discussing their answer with neighbors.

Many other interventions have been tried, usually achieving a measure of success. These methods are sometimes collectively called 'Interactive Engagement' (IE) techniques. "Interactive Engagement methods are those designed at least in part to promote conceptual understanding through interactive engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate feedback through discussion with peers and/or instructors" (Hake 1998, p.65). In a study involving over 6,000 students, Hake (1998) found IE methods yielded superior gains to traditional courses by almost two standard deviations. Despite this and other successes, IE has been slow to catch on in physics departments around the globe.

7.2 The trouble with constructivism

Many researchers, even those critical of constructivism, acknowledge the positive contributions the theory has made to the field of science education. Constructivism has focused attention on learners, their conceptions, and the specific processes involved in learning. It has forced many educators to critically evaluate their teaching, often coming to the realization that learning is not occurring as presumed. Furthermore, teachers and researchers now take more seriously the preconceptions or naïve views of students instead of dismissing them as simple mistakes. Moreover, a set of innovative methods and resources have been developed that focus on conceptual learning.

On the negative side, many have claimed that constructivism's flaws have been ignored by the research community (Millar 1989, Solomon 1994, Phillips 1995, Osborne 1996), limiting the possibilities for research and the extent of critical thought about learning. It is also debatable whether some of the pedagogy promoted as constructivist truly follows from the fundamental principles of constructivism. In the next two sections, I consider first how constructivism has maintained its position as the prevailing education ideology and second, how its limitations have negatively affected both research and practice.

7.2.1 Constructivism as a meme

Constructivism can be characterized as being a particularly persistent 'meme.' A meme is a unit of cultural evolution, just as a gene is a unit of biological evolution.

The ‘memetic’ theory holds that the cultural practices we observe today exist because they appeal to us in ways that get them repeated (Dawkins 2006). Memes that have intrinsic appeal because of the way they interact with our psyches and with other memes become more prevalent over time, whereas memes with less appeal decrease in frequency and eventually go extinct.

Constructivism can be seen as a meme because of the way it has arisen in different fields and sustained itself in a position of hegemony. Constructivism traces its roots back to philosophies in Ancient Greece, but is more commonly attributed to Kant and Vico. In 1708, Vico declared: “the norm of truth is to have made it.” Since this time, constructivism has been a recurrent theme in philosophy. More recently, the ideology has made inroads in education and even international relations (e.g. Wendt 1992). Despite many criticisms of constructivism in education and declarations that the theory has run its course (e.g. Millar 1989, Solomon 1994, Phillips 1995, Osborne 1996, Swartz 1999), it remains dominant.

It is worthwhile to consider why constructivism is such a persistent meme. Below, I suggest three possible explanations: constructivism is simple and intuitively appealing; it can be interpreted (or misinterpreted) to support a range of pedagogy; and, some of this pedagogy has resonated with the prevailing idea that traditional modes of teaching must be overthrown. Likely all of these reasons act in concert to give the theory its staying power.

First, constructivism is relatively easy to understand and its assertions are not especially controversial. “The notion that knowledge is personally constructed is at one level a trivial truism” (Osborne 1996, p.66). Constructivism “embodied a truism that the teacher cannot ‘learn’ the student” (Swartz 1999, p.330). The central premise of constructivism, that learning is an active, conscious activity of a learner is not generally open to debate. The suggestion that objective reality cannot be known separately from our sensory experience of it is again something that most will accept. These ideas are not particularly complicated nor confronting if given a moment’s thought (Millar 1989).

A second reason constructivism is so persistent is that these seemingly innocuous principles can be used (or misused) to support a range of suggestions about teaching and learning. One of the most common suggestions is that students must be 'doing' in order to learn.

At first glance, this does not seem an unreasonable conclusion. Since students are the actors in the learning scenario, they must be the ones reasoning, synthesizing, and investigating. For this to occur, constructivists argue, they must be challenged to work it out for themselves rather than be told the answers. However, a more detailed consideration of the principles of constructivism reveals that these recommendations overstep the limited guidance the theory can provide for pedagogy (Millar 1989, Solomon 1994).

The active construction of knowledge is something that occurs in the learner's head. The learner's behavioral activity is not necessarily relevant to his or her cognition. Saying that a learner learns by constructing does nothing to explain how effective construction can be facilitated. It is possible that learners can only construct effectively when their hands are on objects of interest like carts and pulleys. Or perhaps constructing most often proceeds fluidly when learners are engaged in discourse. A further possibility is that learners can actively and effectively construct while experiencing a presentation of words and pictures. The principles of constructivism cannot adjudicate among these methods. Experiments would need to be carried out to establish which method (if any) were more effective at encouraging effective constructions. As Millar (1989) cogently summarized,

> the process of eliciting, clarification, and construction of new ideas takes place internally, within the learner's own head. This occurs whenever any successful learning takes place and is independent of the form of instruction. (p.589)

The idea that constructivism prescribes discovery-style teaching has been called the *constructivist teaching fallacy* (Mayer 2004*a*), and is perhaps the most damaging influence of constructivism on education. "It could be argued here that a weak or at least controversial epistemology has become the basis for a strong pedagogic policy" (Phillips 1995, p.11).

The depth to which this idea is ingrained in constructivists is evident in their writings. For example Dean & Kuhn (2007) believe it is self-evident that direct instruction cannot be effective in teaching experimental skills. "The counterintuitive nature of the proposition, however—in suggesting the superiority of a method other than involving students in activities that demand inquiry as a means of fostering inquiry skills—means that claims for such superiority should be especially well-documented" (pp.385–386). This view illustrates the deep and unremarked paradox in constructivist reasoning: "it is like arguing that the best way to learn poetry is by writing poetry" (Ogborn 1994).

The promotion of activity-based, inquiry-style methods that run counter to many traditional practices, is a third reason why constructivism may be so tenacious. This is not to say that these methods are necessarily wrong-headed or ineffective, but that they are different to the practices that have come before. This appeals to educational researchers who have a vested interest in discrediting previous learning theories and methods of teaching. After all, if previous methods were sufficient, then educational research would serve little purpose. Test results from conceptual inventories seem to validate this claim, but some argue that the desirable outcomes of a physics course are not necessarily the same as what is measured by the tests.

The preceding discussion helps understand why many educational researchers are eager to accept (a) the conclusion that previous teaching practices were ineffective and (b) the notion that constructivism supports inquiry-based learning. "However we account for its popularity and growth, there can be little doubt that the alternative conceptions movement has resonated exceptionally well with the mood of the times," (Millar 1989, p.588).

7.2.2 Marginalizing other perspectives

Unfortunately for science education research, constructivism's position of dominance has limited the scope of research and the number of different perspectives in the field. Any research that explored students' responses to lectures or other non-interactive multimedia has been open to the criticism that it is 'transmissionist' or out-dated. Furthermore, the idea that learning occurs through new constructions being built on old ones has led few to consider exactly how the construction process takes place. Research on memory and psychology, for example, has rarely been cited in science education studies.

> The problem is that other paths and viewpoints are not just ignored, they become disused and impassible. If constructivism obscures other perspectives, either by its popularity or its blandness, that could be damaging. (Solomon 1994, p.17)

Something that could be particularly damaging for students learning under the constructivist paradigm is pedagogy that disregards the cognitive limitations of working memory (Section 3.2). "Constructivist pedagogy often makes a fallacious connection between the manner in which new scientific knowledge is created and the manner in which existing scientific knowledge is learned" (Osborne 1996, p.54). The idea that learners should learn science by doing science is appealing to many—indeed if a learner is to become a scientist, at some point he or she must practice the methods of the discipline. However, due to cognitive restrictions on working memory, this transition should come later on in a student's development, once he or she has already mastered most of the fundamental science. Sweller (2004) notes that pedagogic practices with less guidance may produce robust learning but they may also be very slow and time-consuming. New knowledge must be built this way but that does not make it the best method of learning old knowledge (Kirschner, Sweller & Clark 2006). One could argue that such an approach might lead to germane cognitive load (mental effort required for meaningful learning), but would also likely involve excessive extraneous cognitive load (mental effort that does not result in learning).

Students have often reacted negatively to constructivist physics education initiatives for reasons that stem from a lack of guidance and cognitive support for learning. A notable example comes from the Technology Enhanced Active Learning (TEAL) project implemented at the Massachusetts Institute of Technology (Belcher 2003). In the TEAL program, students worked in groups of up to nine for five hours per week with very little lecture instruction. They participated in activities involving simulations, problem solving, and hands-on experimentation. Reading and reading assignments were required before every session and marks were awarded for attending classes.

Despite improved gains on standard conceptual inventories, students were immediately critical of TEAL. One hundred and fifty students signed a petition requesting lectures be offered as an alternative. Some watched recorded lectures from the previous year's course on the Internet during the time allocated for group work. Vocal student opposition to the new teaching identified a lack of guided instruction as the reason the program was so unpopular.

> Basically, the idea behind TEAL is that students 'learn on their own.'... The problem, of course, is without a lecturer, students who didn't understand the reading cannot gain anything from the experiments or the workshops. Thus all this time is wasted... The workshops, for example, usually simply get done by one person, because the others are often clueless. (Agarwal 2003)

A renowned professor at the university, Walter Lewin, was aware of the difficulties students were having with the course. "Most complain that TEAL is not helping them to learn, so they are on their own. Without recitation, the students are missing the ins and outs of problem solving" (Lewin as quoted in LeBon 2003).

This is a typical example of reactions to constructivist teaching that fails to provide enough scaffolding for learners. In their eagerness to embrace the notion of students as the active constructors of their knowledge, educational researchers have much less considered the impact of learners' well-documented cognitive limitations.

> The advocates of constructivist methods of teaching have failed to recognize that there is a role for telling, showing, and demonstrating. Teachers, we are told, should 'negotiate,' 'facilitate,' 'coconstruct,' 'mediate,' 'socialize,' 'provide experiences,' 'introduce,' and 'make the cultural tools of science available,' but never ever will they *tell*. (Osborne 1996, p.67)

7.3 Social learning theories

Not all constructivist research promotes such minimally guided practices of teaching and learning. Social constructivism, for example, emphasizes the social interactions involved in learning, many of which involve extensive guidance. While scientific knowledge is still seen as constructed, it is viewed as the cumulative product of many scientists' efforts, working together with a shared set of understandings and practices. Learning science is then a process of enculturation by which students come to know the symbol systems, methods, and language of the scientific community. Listening and observing are valuable to this pursuit because, in this view, science is not only about understanding natural phenomena, but also about the socially agreed upon conventions used to express that understanding.

Social learning theories are especially relevant for this study because social processes are so limited in the observed current teaching practices yet they figure centrally in reform methods and conceptual change theories. In virtually all of the methods developed by physics education research groups, discussion among students and between students and lecturers is seen as essential. Disparate theories of conceptual change also support these practices because they can be used to elicit student conceptions either to confront them or to build upon them. In the quantum mechanics lectures discussed in Chapter 5, social interaction among students or between students and teachers was minimal. Lecturers rarely forced students to make predictions about demonstrations performed in class and when they asked questions, students rarely answered. Of the few questions asked by students, only a small fraction were seen as productive. These observations make it essential to understand the functions of social processes in learning and how they can inform the development of multimedia.

In the learning process, the scientific community plays the central role of maintaining: (1) a body of accepted knowledge, (2) the social tools for recording that knowledge, and (3) the practices for uncovering new knowledge. In order to perform these functions, language, symbols, and methods must be continually refined by scientists. Klein (2006) claims that one direct outcome of the social negotiation of knowledge is the commonly observed lexically dense nature of scientific text. Scientists from different specializations and cultures must be able to communicate effectively with each other. Carefully defined words are essential to this purpose. Textbooks and other recorded summaries of scientific findings therefore take on a particular form: words are given clear, specific meanings and then used in lexically dense passages both for brevity and clarity.

When a learner attempts to join the scientific community, literacy initially poses a great challenge. Being new to the culture, he or she does not share the definitions nor the implicit or explicit assumptions of the discipline. Learning involves a 'cognitive apprenticeship,' with significant guidance from experienced members of the group (Collins, Brown & Newman 1987).

Social learning theories recognize the importance of observation to this apprenticeship, especially in the early stages of learning. Although the practices of the community may involve trial-and-error experimentation, the learning of those practices needs not take place in the same way. Natural social processes, which involve listening, observing, and modeling, offer a more efficient means of enculturation. These activities form a central core of Bandura's (1986) social cognitive theory.

> If knowledge could be acquired only through the effects of one's own actions, the process of cognitive and social development would be greatly retarded, not to mention exceedingly tedious. The constraints of time, resources, and mobility impose severe limits on the situations and activities that can be directly explored. Without informative guidance, much of one's efforts would be expended on costly errors and needless toil. Fortunately, most human behavior is learned by observation through modeling. By observing others, one forms rules of behavior, and on future occasions this coded information serves as a guide for action. Because people can learn approximately what to do through modeling before they perform any behavior, they are spared the costs and pain of faulty effort. The capacity to learn by observation enables people to expand their knowledge and skills on the basis of information exhibited and authored by others. Much social learning is fostered by

> observing the actual performances of others and the consequences for them. (p.47)

The social constructivism of Vygotsky (1978) has also been extremely influencial in recent decades. Like Bandura, Vygotsky believed that teaching must run ahead of development. Through interactions with more experienced peers or teachers, he argued, learners could achieve cognitive feats beyond what they could accomplish unaided. Vygotsky called the area between what is cognitively possible with and without guidance the *zone of proximal development* (ZPD). It is within the ZPD that a learner can make progress towards learning goals.

The upper limit of the ZPD can be seen as a direct consequence of cognitive load theory. Because of their limited experience in the domain, novices have few schemas and even fewer that are automated. Therefore, their abilities are understandably much more restricted than those of experts. These restrictions can be reduced when novices work with more experienced tutors. In effect, tutors serve to expand the working memory capacities of novices. Of course tutors also perform other important functions, offering additional concepts, perspectives, and interpretations when appropriate, but their ability to direct the learning activity is key. Tutors can maintain an overview of the problem and remind learners of important information in a timely fashion. This guidance can only be successful up to a point, however. When the learning task greatly exceeds a novice's working memory span, a tutor cannot provide enough assistance without first spending time developing requisite schemas. This gives an upper bound to the ZPD.

Vygotsky (1978) believed social processes serve even more important functions in personal learning and development than the scaffolding role discussed above. He claimed that all higher cognitive functions originate in observed social interactions. Vygotsky proposed his theory in terms of two planes: the 'social plane,' which refers to discussions between people; and the 'intrapersonal plane,' or inner monologue. According to Vygotsky, new ways of thinking can be developed by observing social interactions. In this manner,

> an interpersonal process is transformed into an intrapersonal one. Every

> function in the child's cultural development appears twice: first, on the social level, and later on the individual level; first between people (interpsychological) and then inside the child (intrapsychological). All the higher functions originate as actual relationships between human individuals. (Vygotsky, 1978, p.57)

This oft-cited quote outlines how the schemas of others, verbalized in social interactions can act as templates for novices. Vital to learning is not only the information contained in the utterances but the ways in which dialogue participants respond to each other. Embedded in these interactions are ways of thinking and reasoning. It is important to note that Vygotsky did not see this process constituting a transfer of ideas from the social plane to a pre-existing internal plane. One of his contemporaries, Leontiev, made it clear that "the process of internalization is not the transferral of an external activity to a pre-existing 'internal plane of consciousness.' It is the process in which this plane is formed" (Leontiev 1981 as cited in Scott, Asoko & Leach 2007, p.40).

The idea of internalization bears striking similarity to the 'borrowing principle' outlined by Sweller (2004). In short, this idea is that learners unfamiliar to a subject area must draw on the schemas of those more experienced, or else expend incredible amounts of energy testing random combinations of information for appropriateness.

The idea that learners depend on the schemas of others during the enculturation process leads social constructivism to another point of departure from individual constructivist accounts. According to mainstream constructivism, learners should be exposed to various phenomena and then asked to make meaning of their experiences. During the interpretation stage the teacher may help guide students towards the accepted scientific view. Social constructivism, by contrast, recognizes that students' existing conceptions heavily influence their perceptions. Therefore they should be acquainted with the schemas of experts *before* they directly experience any phenomena.

> Perceptions are guided by preconceptions. Observers' cognitive competencies and perceptual sets dispose them to look for some things but

> not others. Their expectations not only channel what they look for but partly affect what features they extract from observations and how they interpret what they see and hear. By giving coherence and meaning to available information, cognition is very much involved in perception. (Bandura 1986, p.53)

It is perhaps for this reason that Carlsen (2007, p.59) concludes "some scientific concepts may never arise from hands-on experience, no matter how creative or time consuming that experience may be."

These ideas are not unique to social views of learning, however. It is well-acknowledged in the cognitive psychology literature that long-term memory structures play a pre-eminent role in working memory.

> The information processed in the short-term store has already made contact with information stored in long-term memory. For example, our ability to engage in verbal rehearsal of visually presented words depends on prior contact with stored information concerning pronunciation. Thus, access to long-term memory occurs *before* information is processed in short-term memory. (Eysenck & Keane 2005, p.193)

Although cognitive science models offer potential insight into the phenomenon of misconceptions, they have rarely been considered by science education researchers. In Section 3.2 the most common mode of forgetting from long-term memory, interference, was mentioned. Retroactive interference is said to occur when newly learned ideas result in the forgetting of older memories. Learning one's new phone number or address, for example, may inhibit recall of the old number or address. Proactive interference refers to the opposite process by which previously learned schemas inhibit the learning of new knowledge or skills. Anyone who has travelled between countries that drive on opposite sides of the road knows the difficulty of learning to look in the correct direction at a crosswalk. It takes active, conscious thought (and often many repetitions) to reject one's almost reflexive response and learn to look in the right direction. Proactive interference appears to account well for many observations of misconceptions and conceptual change.

Kane & Engle (2000) performed a study on proactive interference in which they recruited participants with high and low memory spans. They found that those with high memory spans were much better able to overcome interference than those with low memory spans. However, when the participants were asked to perform a finger-tapping task concurrently with the memory task, high-span learners performed no better than their low-span counterparts. This implies that in the first part of the experiment, high-span participants were using some of their working memory to actively block interference. The implication of this experiment for achieving conceptual change is that learners must focus almost exclusively on the learning task and extraneous cognitive load must be minimized. This is particularly important in physics where most concepts involve a large intrinsic cognitive load.

From this perspective, imagine the processes that must occur when a learner tries to unlearn a misconception. First, when unfamiliar ideas are presented, the learner will have difficulty interpreting the ideas in light of his misconception schemas. This would arguably involve more intrinsic cognitive load than if the learner had no prior knowledge of the subject at all. If he is capable of accommodating a new idea, he will then be faced with additional ideas that build upon this one. However, when he goes back to retrieve this idea from long-term memory, it may suffer interference due to the prior conception. Older ideas have greater robustness than newer ideas of similar rehearsal. The process of learning a new conception is therefore like trying to understand little pieces of incomprehensible new knowledge with the preconceptions acting like a schema quicksand, swallowing up newly learned ideas as the learner turns his attention to the next construct. This interpretation supports Halloun & Hestenes's (1985) original conclusions about their data from the Mechanics Diagnostic. "This means that throughout the course the students are operating with a seriously defective conceptual vocabulary, which implies that they continually misunderstand the material presented" (p.1048).

Of the recommendations for teaching made by Vosniadou, Chi, diSessa, and colleagues, implications of cognitive science most closely parallel those of the 'knowledge in pieces' approach. Like diSessa & Sherin (1998), cognitive science advocates the activation and use of existing schemas in the learning of new knowledge.

This is not because the practice closely adheres to the principles of constructivism but because the idea of schemas and their impact on learning are well documented. Smith et al. (1993) criticized the idea of replacement because it does not fit with constructivist principles. Studies of proactive interference suggest that a direct 'replacement' of existing schemas is probably impossible since the old schemas are robust and will likely be activated in preference to newer ideas. It would seem that instead new ideas should be tethered to old ones so that they will be activated in turn. DiSessa and colleagues don't believe that confrontation is a necessary or worthwhile process in conceptual change. On this point the cognitive science model better agrees with Chi and Vosniadou. Confrontation is not necessary as a precursor to replacement, rather it is essential for raising the mental effort students invest in instruction. They must be made aware that their pre-existing ideas are insufficient and require significant thought.

7.4 Summary

This chapter was motivated by the results of Chapters 5 and 6, which looked at the current teaching practices and resulting learning of intermediate quantum mechanics in the School of Physics. The observed teaching practices were in many cases innovative, and in most, exemplary forms of multimedia. The learning that occurred as a result of these practices fell well below expectations, however, replicating the findings of the past three decades of physics education research.

The ideas collected under the umbrella of constructivism have been found useful by many researchers and educators when interpreting the challenges of physics teaching and learning. Constructivism has spawned many productive lines of inquiry and thought. It has focused attention on the activities of the learner. It has recognized the importance of prior or alternative conceptions, thereby developing a large catalogue of alternative conceptions. Conceptual inventories were designed and refined to accurately assess incoming knowledge states and determine the effectiveness of instruction. This has also led to theories about the nature of alternative conceptions, and how to change them.

These developments raised the awareness of teaching in physics institutions, and spawned the physics education research movement. Many new teaching strategies were developed and implemented, often producing much greater learning gains than previously observed.

Constructivism as it is commonly understood and practiced, however, has limitations that have gone undetected, allowing it to become the dominant theory in education. The main limitation identified in this chapter is constructivist teaching methods fail to account for cognitive restrictions in learning. Other limitations have been addressed by critics of the ideology, but this limitation is the most important for this thesis.

Social constructivism, unlike many other forms of constructivism, emphasizes the importance of observation in learning. It suggests that higher cognitive functions may in fact result from the observation of social interactions. Furthermore, it asserts that these experiences involve a lower cognitive load, and conserve learner effort.

Chapter 8

Quantum mechanics multimedia

8.1 Introduction

The next stage in the design experiment involved using the preliminary research and literature reviews to design and test a multimedia intervention.

The multimedia equivalence principle discussed in Chapter 1 suggests that different forms of multimedia can inspire the same thought processes necessary for learning if appropriate methods are employed. Social interactions among students and between students and lecturers are valued by students, educators, and theorists from different research paradigms. Social constructivists argue that discussions are important to the process of enculturation. Cognitive load theorists contend that observing discussions involves less extraneous cognitive load than discovery-style methods. Lecturers and tutors in reform programs use discussion as a tool for confronting prior conceptions and developing new ones. However, under the current system of physics teaching very few dialogues take place within the zone of proximal development.

The goal of this study was to incorporate dialogue into a multimedia resource and assess its effectiveness for changing student conceptions. The hypothesis was that this might prove an effective means for confronting alternative conceptions. Incorporating dialogue into a learning resource is a logistical challenge. Some researchers have attempted to empower computers to act as participants in learning conversations, but by most accounts the computer is still developing as a conversational partner (e.g. Andriessen, Baker & Suthers 2003). The obvious alternative is to have students watch modeled student-tutor discussions, which could act as triggers for questions or reflection. It is an open question whether this 'vicarious learning' method could encourage 'heads-on' learning.

The foreseeable application of vicarious learning in physics education is then twofold: first, as a lecture aid to activate prior knowledge, validate student concerns, concretize abstract phenomena and hopefully encourage participation; and second, as an online reference tool to encourage review, reflection and metacognition.

Three specific areas of research informed the development of the dialogue video and the methods employed in this study: vicarious learning, conceptual change refutation texts, and coping models. Research and implications for this investigation are summarized below.

Vicarious learning

Although dialogue is well accepted as an important tool for learning, the educational usefulness of watching or overhearing dialogue is a matter for debate. Some research indicates that vicarious learning can be as effective as didactic instruction (Lee, Dineen, McKendree & Mayes 1999). In the domain of sentence parsing, Cox, McKendree, Tobin, Lee & Mayes (1999) found that students who observed a 're-used' student-tutor dialogue with accompanying animation performed as well as those who listened to a direct tutor explanation. Furthermore, Cox et al. argue that vicarious learning may have benefits not measured by performance tests; vicarious learning is more student-centered and it provides a model for the kinds of questions and reasoning that are appropriate in a domain (see also McKendree, Lee, Dineen & Mayes 1998). Others contend that dialogue is useful only for direct participants in a conversation since listeners have different backgrounds and do not receive tailored feedback to meet their particular needs (Schober & Clark 1989). Further studies have investigated the educational value of learners watching human-computer tutor interactions. Craig, Driscoll & Gholson (2004) found observers yoked to recorded student-computer tutor interactions did not perform as well as the students who were directly involved. This could be due to a loss of fidelity or the lack of tailored feedback available to observers as suggested above.

The results of these studies illustrate how authentic dialogues may dilute the potential of vicarious learning. Authentic dialogues are complicated by the model student's concerns and mannerisms and do not reflect the range of alternative conceptions present in the intended audience. For these reasons, a manufactured dialogue was used in this experiment, ensuring the explicit inclusion of alternative conceptions identified by research.

Conceptual change refutation texts

In non-interactive media, studies have been carried out on so-called refutation texts, in which misconceptions are discussed and rejected (for a review, see Guzzetti et al. 1993). Although some conflicting findings were reported, in general text that attempted to create cognitive conflict resulted in greater learning gains than non-refutational text. A study of elementary science students found that a refutational text passage on energy that addressed two prominent preconceptions was much more effective as an adjunct to standard class teaching than an expository text (Diakidoy, Kendeou & Ioannides 2003). One extensive qualitative study (Guzzetti, Williams, Skeels & Wu 1997) found that refutation texts do induce cognitive conflict and over a period of months can help students develop correct scientific understandings. In some cases, however, students found support for their alternative conceptions in refutation texts when the refutation was not direct enough or the students lacked necessary reading strategies.

Vicarious learning presents a similar opportunity for students to engage their existing conceptions and to observe the necessary reasoning pathways that lead to scientifically correct conceptions. In conjunction with the results of vicarious learning studies, it is logical to expect that students should learn as much, if not more, by watching dialogue than by receiving traditional instruction.

Coping models

Researchers have investigated the effects on self-efficacy and performance of watching peer models complete tasks, either with ease (mastery condition) or gradually (coping condition). Schunk & Hanson (1985) found that young math students displayed higher self-efficacy and achievement after watching a peer model solve mathematical problems than after watching an adult model. Although no significant difference was observed between the mastery and coping condition, students in a later study (Schunk, Hanson & Cox 1987) who watched the coping condition judged themselves to be more similar to the model than students in the mastery condition.

This suggests that vicarious learning could improve students' self-efficacy and attitudes towards education, especially in difficult subject areas. In this study a coping model student was depicted in the dialogue to encourage students to relate to the model.

Local Context

This study was inspired by investigations into learning in the domain of quantum mechanics, a notoriously difficult subject for undergraduate physics students due to its abstract, counter-intuitive, and highly mathematical nature (Fletcher 2004). A survey of quantum mechanics lectures revealed that very little discussion occurs during class time despite significant use of simulations, visual materials, and real world examples (Muller 2005). Students were sometimes given substantial portions of a lecture to raise concerns, but their questions were rarely appropriate for the learning outcomes of the course.

The nineteen lecture second year quantum mechanics course follows on from a first year series that briefly introduces quantum tunneling. Quantum tunneling was addressed in greater detail in the second year course two weeks prior to the experiment. Two-hour weekly computational laboratories give students the opportunity to work on quantum mechanics problems with the aid of powerful mathematics software. The major benefit of these sessions, though, comes from the interactions students have with each other, lecturers and tutors in small groups as they work through the exercises. Even with opportunities for interactions in lecture and computational laboratory, students emerge from the second year quantum mechanics course largely confused about fundamental conceptions of the subject (Muller & Sharma 2005). This is the area in which the potential of vicarious learning was investigated.

Four questions guided the inquiry:

1. Can vicarious learning be as effective as didactic modes of instruction?
2. Can the alternative conceptions, dialogue, and representation of a student on-screen encourage learners to consider their prior knowledge and reflect upon their learning?
3. Can vicarious learning provide affective benefits, improving self-efficacy or validating students' concerns?
4. Do students perceive this strategy as potentially helpful for promoting question-asking in lecture?

To answer the first question, students were tested before and after watching either a Dialogue or Exposition video. To answer the latter three questions, interviews were carried out with students from each of the treatments.

8.2 Method

A mixed methods approach was adopted to determine learning gains quantitatively while exploring learner's perceptions of the two educational resources. The seven steps for developing resources, testing materials, and running interviews are outlined below.

1. Identifying alternative conceptions

A research-based questionnaire was administered to second year physics students in 2004 following regular lecture and computational laboratory instruction (Chapter 6). The test was designed to assess conceptual understanding of quantum mechanical tunneling, a problem that has received increased attention because of the way it draws together key ideas of quantum mechanics (Redish et al. 2000, Morgan et al. 2004, Domert et al. 2005). Results of this survey showed students hold a limited range of alternative conceptions, consistent with the findings of physics education research (Fraser & Tobin 1998).

2. Designing the video treatments

In as many ways as possible, all video material for this study was created in line with principles of multimedia learning (Mayer 2005). Verbal material was spoken rather than presented on screen, diagrams were kept simple, and information was presented contiguously both in time and space. However, the Dialogue video involved words and diagrams not directly related to the learning outcomes, in contradiction of the coherence principle (p.49). On the issue of extra material, Mayer writes (private communication):

> My interpretation is that the added material can have both a negative effect (by increasing extraneous processing) and a positive effect (by increasing the motivation to engage in germane processing). Thus, the effect on learning outcomes depends on whether the negative effects are greater or less than the positive effects.

In relation to multimedia learning literature, this study should help determine whether the benefits of the vicarious learning method outweigh the deficit of increased cognitive load (Sweller 1994) with intermediate quantum mechanics students.

The main alternative conceptions were then used to create a video Dialogue, simulating the discussion that might take place between an inquisitive student and a tutor on quantum tunneling. The Dialogue was specifically scripted so that the tutor did not provide direct answers, but rather questioned the student on certain parts of her reasoning so that she identified and resolved inconsistencies in her own logic. This Socratic method (Hake 1992, Edelson 1996) has been advocated in physics education reform as an important tool for promoting reflection and activating prior knowledge in students. The student in the Dialogue represented a coping model as she verbalized confusion and frustration when she encountered new information that contradicted her existing conceptions. The semi-authentic dialogue was filmed and supplemented with simple drawings and animations to illustrate the ideas raised in the discussion. The video was scripted rather than created from recorded student-tutor discussions to ensure the widest range of alternative conceptions were addressed in a clear manner.

For comparison, a video Exposition was created, which summarized the same correct physics information, simple drawings and animations as in the Dialogue but without alternative conceptions. This was similar to the presentation a well-prepared lecturer might make on the topic. To ensure both videos contained the same correct physics material, the course lecturer, an experienced physics education researcher, examined both treatments. An additional physics academic reviewed the scripts of both videos for clarity and consistency. The scripts for both videos can be found in their entirety in Appendix C.1 and the videos are on the DVD (Appendix D) on the back cover.

Since the Dialogue treatment included alternative conceptions and used a conversational format, it had a longer running time (13 min.) than the exposition treatment (7 min.) and contained twice as many words. Implications of this difference are considered in Section 8.4.

3. Video production

Both videos were produced using a consumer video camera (Sony HDR-FX1) and readily available animation (Flash MX) and editing (Final Cut Pro) software.

4. Design of pre/post-test

To evaluate students' conceptual understanding of quantum tunneling before and after watching the video, the same research-based questionnaire from 2004 was used (Appendix B.3). Additional questions were included from the Quantum Mechanics Conceptual Survey (McKagan & Wieman 2005). This test is being designed to perform a similar function as the FCI in the quantum mechanical domain. In line with the definitions used by Mayer (2001) the pre- and post-test included both questions of retention and transfer. Students were asked to draw and describe verbally characteristic graphical representations of quantum tunneling to assess consistency of conceptions. Along with the pre-test, students completed a personal information sheet on which they indicated their interest in physics and preferences for various modes of learning (Appendix A.3). With the post-test, students filled in an opinion form about the video they watched (Appendix B.4 with results in B.5).

5. Running the experiment

After lecture and laboratory instruction on quantum mechanics in 2005, the second year physics class was randomly divided into two groups during a standard lecture ($n(D) = 40, n(E) = 39$). The class consisted of Advanced and Regular students so each of these groups was split independently to obtain an equal ratio of Advanced and Regular students in each group. Advanced students generally have more background in physics and receive more in depth and mathematical instruction than regular students. In two lecture halls, students were read the same script outlining the procedure of the experiment. After filling out ethics consent forms, students were given fifteen minutes to complete the pre-test and personal information sheet. They were not informed that there would be a post-test at this time. This was done so that students watched the video without the objective of memorizing answers for the post-test. After the video was shown, students were allowed fifteen minutes to complete the post-test and opinion response form.

Two students in the Dialogue treatment and four students in the Exposition treatment were excluded from the analysis because their pre-test scores were three standard deviations greater than the means of the remaining samples. There for the final sample sizes were $n = 35$ for the Exposition and $n = 38$ for the Dialogue.

6. Student perceptions

Four interviews were held with between one and three volunteer students representing each of four populations (Advanced students from the Dialogue treatment

(AD), Regular students from the Dialogue treatment (RD), Advanced students from the Exposition treatment (AE), and Regular students from the Exposition treatment (RE)). These interviews, which ranged from thirty minutes to an hour in length, were videotaped and transcribed to gain insight into student perceptions of the videos and possible reasons for differences in achievement on the post-tests. Students were asked to fill in a worksheet about their perceptions of the multimedia (Appendix B.6) and to discuss their responses.

7. Long-term retention test (retest)

Two months following the original experiment, the post-test was again given to students at the end of a lecture to assess long-term retention. The sample size for the retest was smaller than the original experiment because some students do not attend all lectures. Furthermore, students who had participated in the interviews were removed from the sample. The final sample size for the retest analysis was 25 for each of the Exposition and Dialogue multimedia.

8.3 Results and analysis

1. Marking

The questionnaire consisted of three open-ended questions, two drawings, and four multiple-choice questions. The open-ended and multiple-choice questions were each worth one point and the drawings were each worth four points, reflecting their relative level of complexity. One of the multiple-choice questions (2a) was excluded from the analysis due to a high non-response rate ($> 30\%$) in both treatment samples. This question was located at the top of the back page and was likely overlooked by many students. The maximum possible score on the questionnaire was fourteen points.

2. Pre-test

Both Exposition and Dialogue pre-test scores were normally distributed (K–S $p > .05$) and were not significantly different from one another ($t(71) = -.21, p > .05$), with means of 2.6 and 2.7 respectively as shown in Figure 8.1.

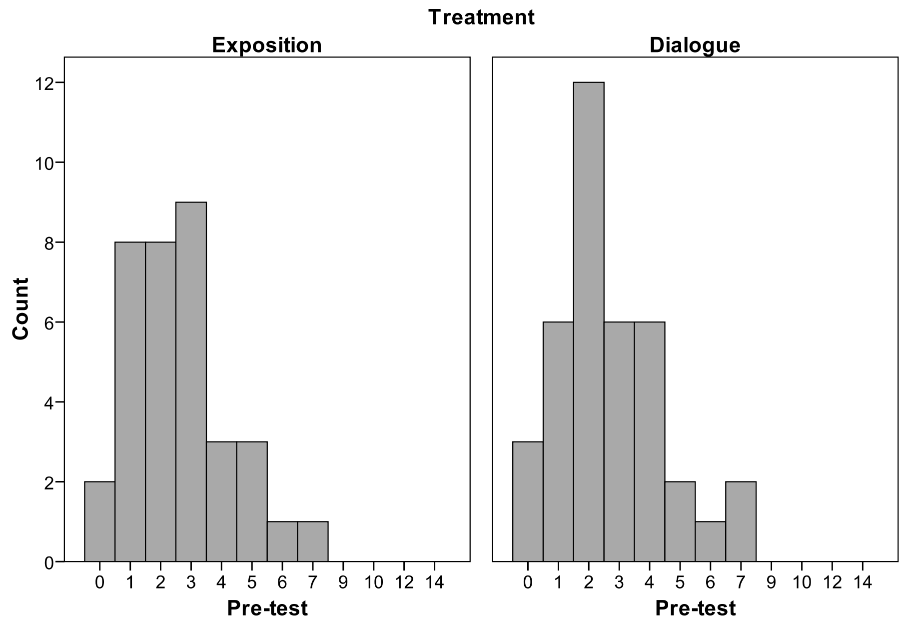


Figure 8.1: Pre-test distributions from the Exposition and Dialogue treatment groups were not significantly different.

3. Post-test

Both Exposition and Dialogue treatments produced substantially higher post-test mean scores of 6.8 ($SD = 3.3$) and 9.1 ($SD = 3.3$), respectively. The post-test scores were each normally distributed (K–S $p > .05$) and the Dialogue distribution was significantly higher than that of the Exposition ($t(71) = 3.01, p < .01$) as shown in Figure 8.2, with an effect size of $d = .71$.

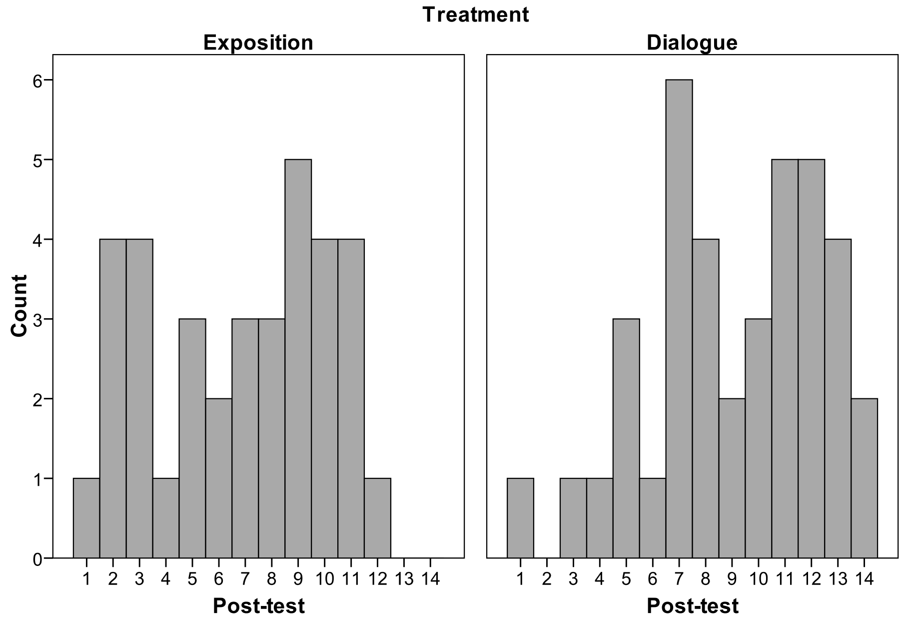


Figure 8.2: Post-test distributions with means of 6.8 for the Exposition treatment and 9.1 for the Dialogue were significantly different.

4. Correlations with pertinent variables

The samples were not significantly different with regards to gender, Advanced/Regular student ratio, or self-reported interest, understanding, revision or plans to continue in physics as shown in Table 8.1. Thus the random splitting of the class was effective. Regardless, these factors were not found to correlate with pre-test, post-test, or gain scores in either sample.

Treatment	Male	Female	Adv	Reg	Interest		Understand		Revision	
					M	SD	M	SD	M	SD
Exposition	23	12	15	20	5.5	1.1	3.9	1.1	3.3	1.5
Dialogue	31	7	21	17	5.9	1.0	4.3	1.1	3.7	1.4

Table 8.1: Descriptive statistics for the two treatment groups.

Pre-test scores correlated with post-test scores for both video treatments as shown in Table 8.2, but only correlated with normalized gain scores for the Dialogue group. This suggests that prior knowledge may be an important factor in vicarious learning. Students who had less prior knowledge were less likely to learn by watching the Dialogue than the Exposition. This result could be interpreted as evidence that the Dialogue treatment encourages reflection and the building of concepts in light of prior knowledge. The more prior knowledge a student possesses, the more able he or she is to activate and expand on existing conceptions. This is an intriguing result that calls for further investigation.

Treatment	Pre test correlation with post-test	Pre-test correlation with normalized gain[a]
Exposition	$r = .34$ $p = .045$	$r = .14$ $p = .408$
Dialogue	$r = .48$ $p = .002$	$r = .34$ $p = .037$

[a]Normalized gain is calculated by dividing the gain by the total possible gain $(post - pre)/(100 - pre)$

Table 8.2: Correlations between pre- and post-test scores and normalized gain.

8.3.1 Interviews

Interviews were used as a secondary source of data to help understand the quantitative findings. Below, interview participants are referred to by stream indicated by two letters (A or R for Advanced or Regular and E or D for the Exposition or Dialogue case) and a number to differentiate students in the same group.

As in the *Falling Cats* study, students felt the two modalities of information helped them learn.

AD1 Yeah I just liked that when you explained something, the little visual that comes up because then you're getting these two really really maximized kind of ways of understanding something at the same time, so yeah I thought that was the best part.

AD3 Yeah, I have to concur—like a description of... you've got a pictorial representation that's animated and you've got somebody describing it as it's going along so you're getting a description of what you're seeing which is always a helpful way 'cause you can put it together in your head.

These comments are concordant with dual coding theory and the *multimedia principle*, which states that instruction is more effective when presentations are made in two modalities rather than one.

Although it is impossible to conclude from this study whether showing either of the videos in lecture would increase the frequency of discussions, participants attitudes on the subject were solicited. Some students felt that the lecture environment would still inhibit discussion while others believed either video might crystallize some ideas for students prompting them to speak up.

AE2 Usually when the lecturers say 'any questions?' they're usually asking it at the wrong time. But after this video would be a right time to ask 'are there any questions?'

RD1 I think that will, for some personality types, help them actually overcome their sheepishness and actually ask questions when they are confused.

Students' acknowledged three main benefits of the discussion and alternative conceptions present in the vicarious learning Dialogue treatment. First, students found the difficulties modeled in the dialogue validated their challenges with the subject.

RD1 A student can very easily just sit there in a corner and have no idea and not realize that everyone else has no idea as well. Probably admitting that people do find this very confusing is a good thing.

RD3 It was like 'ah, I was thinking that too.' I should ask questions, and I didn't feel so stupid because I would never ask questions in a lecture.

Second, also evident in the second quote above, the dialogue format was found to be a more immediate trigger for activating prior physics knowledge.

RD3 I liked that it was asking me questions, so I had to actually think about it rather than just telling me stuff, because then it had more relevance to me.

AD2 We were watching a lot of dialogue. That's what I found most helpful about the video was the dialogue between the two characters, and the visual feedback was helpful, but by talking about the issue and outlining various places where you could get caught up or trapped or tripped up when thinking about it—that's what helped me answer the questions afterwards.

AD3 Whenever you encountered a problem you'd remember what they were talking about. You were talking around the various mistakes a person could make and issues you could encounter. I found that pretty helpful.

Students contrasted this format to what they see in lectures. One student felt the communication was much better in the video.

AD2 [It was] similar to a lecture but there's no real feedback. In a lecture it's really a one-way thing. You're sitting back watching a lecturer as they work through the example and occasionally they'll ask questions and...occasionally the lecturer will ask the audience to ask questions and, you know, some good ones will get brought up. But I find that in the video the communication between the person who is explaining and the person who's being explained to was a fair bit better.

And third, the research-based script seemed to resonate with students who felt the dialogue was authentic. This also suggests that students actively compared their prior knowledge and questions to those of the model student.

AD2 The student...would often ask questions which were quite relevant, I found they were relevant at least, when I initially attempted the questions which hadn't actually been properly approached in the lectures or hadn't really been explained in much detail when the lectures were done.

RD2 The questions the student asked were really the sort of questions I asked and what other people asked at the beginning of the semester. What [the student] was asking I thought were really genuine questions.

6. Retest

The results from the retest administered at the end of a lecture two months following the original experiment are shown in Figure 8.3. The sample size was reduced because not all students attend all lectures and those who participated in interviews were removed because they spent substantial additional time discussing the quantum tunneling problem. Statistics for the two groups are summarized in Table 8.3. The distributions for the two groups were not normal. A Mann-Whitney test revealed that retest scores were significantly higher for the Dialogue group ($Z = 2.04, p = .04$).

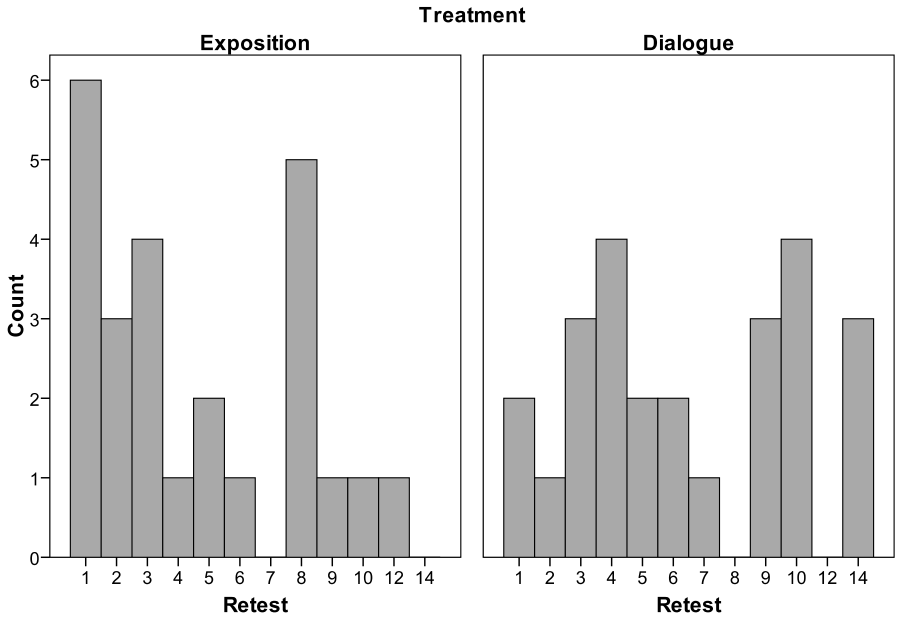


Figure 8.3: Retest distributions with means of 4.6 for the Exposition treatment and 6.7 for the Dialogue were significantly different

Treatment	*n*	M	SD	Median
Exposition	25	4.6	3.4	3
Dialogue	25	6.7	4.0	6

Table 8.3: Descriptive statistics from the quantum mechanics retest, two months following the original experiment. The maximum possible raw score was fourteen.

8.4 Discussion

Despite the additional information in the Dialogue video, students who received this intervention performed better on the post-test than those who received a more traditional summary. Since extra material in short instructional segments typically has a negative effect on learning performance (Mayer 2001), the Dialogue and alternative conceptions must serve a constructive purpose. Interviews with students who watched the Dialogue treatment revealed that they liked the conversational nature of the video. They felt the questions asked by the model student were relevant and similar to the questions they had. This made them feel more comfortable about the concerns they were having since the video addressed ideas they felt might be unique to them. Students also noted that the questions caused them to consider the material with a greater immediacy than if they had been presented with more expository material. Responses were mixed as to whether the video would encourage more participation in lectures.

Over the course of the study, theoretical and practical implications of vicarious learning were clarified. In this section, the validity of the study is discussed, along with the inherent advantages and disadvantages of vicarious learning. In light of these considerations, implementation issues and avenues for further study are explored.

Study validity

Some might argue that since the Dialogue video was longer than the Exposition, the greater gains were due to the extra time students spent thinking about the physics rather than the method. This is an appealing argument since it is obvious that more can be learned in a longer period of time. However, if the same instructional message were presented either briefly or spread out over time, interspersed with material that does not directly contribute to the learning outcomes, one would expect the concise version to be superior because its audience would be more attentive to less instruction. This has been confirmed through empirical studies, sometimes where extra material is added to increase interest (Mayer 2001). The spreading out of material in this study occurred to accommodate alternative conceptions in a dialogue format. Separately, this extraneous material would not have contributed to the learning outcomes, but together with the correct conceptions, it led to higher gains than the concise summary alone.

Advantages of vicarious learning

Vicarious learning provides a unique opportunity for students to reflect on the process of learning. Coupled with assessment activities, online video dialogues could lead learners to consider their own preconceptions and the supporting structures for their current conceptions.

This reflection could also lead students to new understandings of the scientific process. Science, as it is presented in textbooks, appears to be a collection of facts, accumulated gradually over time by hardworking scientists (Kuhn 1996). This gives students the artificial impression that science is a one-way progression rather than the complicated dialogue it is acknowledged to be. Students may be more open to science if they appreciate how competing hypotheses can co-exist and what standards of evidence and reasoning are required to favor one over another.

Disadvantages of vicarious learning

As with all methods that involve information besides the correct curriculum, vicarious learning runs the risk of confusing students or reinforcing misconceptions. This problem has been reported elsewhere in relation to refutation texts (Diakidoy et al. 2003). However in this study, despite the presence of many alternative conceptions, the Dialogue treatment students outperformed the straight-forward exposition group. One student who was interviewed admitted to being confused after watching the video, but overall only a couple of participants in each sample experienced negative gains. This suggests that learners do possess extensive coping mechanisms for dealing with the various points of view and ambiguities present in social discourse (Lee et al. 1999).

Practical disadvantages of implementation must also be considered. With lecturers concerned about having enough time to cover all of the important topics, vicarious learning is a burden because of the extra time it requires. Similarly, despite the demonstrable benefits of refutation texts, written instructional material remains largely expository. The practical disadvantages of refutation texts are obvious: they require more research and time to write, and result in heavier books and more reading for students. However, written dialogue that includes alternative conceptions has already made inroads in some physics education tutorials (McDermott & Shaffer 2001). Students are asked to consider a debate between two fictitious students and decide which, if either, has the right idea and why. Perhaps vicarious learning in lectures could serve a similar function, highlighting key issues and encouraging reflection on alternative ideas.

Future investigations

A more concrete domain may better serve future investigations of the efficacy of vicarious learning. Although students do commonly bring classical mechanical notions to the study of quantum mechanics, the field is so novel, mathematical, and abstract that their conceptions are invariably fragmented and unstable (Johnston et al. 1998). Therefore a truer test of vicarious learning in physics could be achieved in an area like Newtonian mechanics where misconceptions are well documented and have been repeatedly found to be systematic and robust.

8.5 Conclusion

In a lecture setting, students who watched a student-tutor discussion performed better on a test of conceptual retention and transfer than students who watched a direct expository summary. This result suggests that vicarious learning can be an effective method for confronting students' alternative conceptions and raising their levels of confidence. Further investigation is required to quantitatively measure changes in self-efficacy. This may best be facilitated in the domain of classical mechanics where alternative conceptions are common, well-researched, and robust. Further research will also be required to determine whether vicarious learning can stimulate discussions in lectures and how it can be practically utilized in classrooms and online.

Chapter 9

Newtonian mechanics multimedia

9.1 Introduction

The study detailed in the previous chapter provided a proof of principle. Students learned better with multimedia when common alternative conceptions were presented in a dialogue format than when only correct information was presented in a lecture style. However, the study raised new questions for investigation. Below, I discuss limitations of the quantum tunneling multimedia study and how they were addressed in the next design experiment iteration.

One criticism of the quantum multimedia study is that students' incorrect conceptions in the area are not necessarily like the misconceptions documented in previous physics education research. Quantum mechanics is an abstract domain with which students have very little experience. It could be argued that alternative conceptions are less ingrained in student thinking because they have not been reinforced through repeated experience. Student ideas are likely best represented by diSessa's (2006) p-prims, a set of disconnected thoughts cobbled together to form answers when necessary. This means that learning about quantum tunneling would not involve conceptual change of the sort envisioned by Carey (1986) and others, making it easier to achieve.

A better test of vicarious learning for promoting conceptual change through multimedia would require a learning domain full of well-established misconceptions. Therefore in this chapter the subject matter addressed in multimedia interventions was classical Newtonian mechanics, a topic fundamental to virtually all introductory physics courses, but notoriously difficult to learn (McClosky 1983). Since the subject matter is so common and widely misunderstood, extensive research has been carried out to document common student misconceptions (McDermott 1991, Mayer 2004*b*, see Section 7.1.1). Newton's first and second laws were selected as the focus of the multimedia treatments, with design informed by misconception research (Trowbridge & McDermott 1980, Trowbridge & McDermott 1981, Clement 1982, McClosky 1983, Halloun & Hestenes 1985, diSessa 1996).

Studying learning in the domain of Newtonian mechanics had the added benefit of increasing the sample size. All approximately 800 first year physics students receive lecture instruction on Newton's laws. In addition, these students are grouped into three streams based on their levels of prior knowledge. In the quantum tunneling multimedia study, most students had the same low level of prior knowledge. Despite the Regular and Advanced designations, students scored similarly and generally very poorly on the pre-test. It was important to investigate whether the vicarious learning method is as effective for students who have substantial prior knowledge.

Attempting to use all first year students as participants meant multimedia treatments could not be tested in a lecture environment. Instead, students watched multimedia streamed through the Internet on computers. Consequently, many variables were uncontrolled. Participants accessed the experiment wherever and whenever they liked. They may have used resources or consulted with peers when answering the pre- or post-test questions. Although the time between the start and end of the multimedia was calculated, an appropriate length of time was no guarantee that a student actually watched the treatment. These were features of the methodology. The ability of students to participate as they saw fit ensured the results could be generalized to authentic learning environments, a central requirement of the design experiment methodology.

Another question raised by the quantum multimedia experiment concerns the way in which alternative conceptions are presented in multimedia. Is dialogue an essential feature or could misconceptions be stated and refuted by a single speaker with equal effectiveness? This question was addressed by creating an additional treatment called the Refutation, in which alternative conceptions were raised in a lecture style.

The Dialogue was longer than the Exposition, leaving open the possibility that the extra time students spent thinking about the physics, rather than some benefit of the method, led to greater learning gains. This concern was addressed by creating another multimedia treatment called the Extended Exposition. To make this multimedia, additional interesting and related material was added to the Exposition. This extra material was not directly relevant, however, to the learning outcomes as measured by the post-test.

Finally, measurements of learning in the quantum study involved only comparisons of right answers between treatment groups. No information was obtained about the extent to which students were committed to their answers. Thus there was no way of telling whether difficulties reflected genuine, deeply held beliefs, or spontaneous guesses. On the Newtonian mechanics pre- and post-tests, students were required to report their confidence in their answers. This allowed for students' gains in confidence to be measured and compared among the treatments.

9.2 Managing cognitive load

One of the ways in which different instructional treatments can be categorized is by the way they attempt to allocate cognitive load. In this section, I consider multimedia design suggestions from different bodies of research that have different implications for cognitive load. I also indicate where the treatments used in this study fit into the broader framework.

Reduce cognitive load

On one end of the spectrum is the large body of multimedia research that suggests cognitive load should be minimized by instructional designers to free up cognitive resources for germane processing. The finding that reducing the amount of presented information leads to more learning has been confirmed in a variety of empirical contexts (Chandler & Sweller 1991, Mayer 2003). Instructional messages that contain redundant information sources inhibit learning in what is called the redundancy effect (Sweller et al. 1998). A related recommendation, that all non-essential information be removed from instructional messages is called the coherence principle:

> Adding extraneous words or pictures to a multimedia message can interfere with cognitive processes by encouraging learners to pay attention to words or images that are not relevant, by disrupting how learners organize words or pictures into a causal chain, and by priming inappropriate schemas to be used to assimilate the incoming words and pictures. (Mayer 2003, p.133)

These effects have been observed in student learning of scientific topics like blood flow in the heart (Chandler & Sweller 1991), the formation of lightning storms (Mayer, Bove, Bryman, Mars & Tapangco 1996), and deep-sea waves (Mayer & Jackson 2005). The Exposition embodies the recommendations of multimedia research and cognitive load theory. It includes only accurate scientific information presented clearly with simple diagrams, graphs, and animations. Verbal information is presented as narration rather than on-screen text and it is synchronized with corresponding images. Extraneous sounds and images are avoided as much as possible. This treatment resembles a well-prepared lecture or textbook presentation.

A limitation of much multimedia research is that it was conducted with psychology students in learning laboratories rather than with students enrolled in the learning domain in a more naturalistic setting. In addition, multimedia researchers have considered the effects of relevant and correct prior knowledge on learning with multimedia (e.g. Kalyuga, Chandler & Sweller 1998, Kalyuga, Ayres, Chandler & Sweller 2003), but have neglected the effects of learners' beliefs that are at odds with scientifically accepted views. Arguably the central obstacle to science education, alternative conceptions can have an inhibitory effect on understanding scientific principles.

Raise cognitive load

Science education researchers rarely build upon multimedia and cognitive load research as they generally regard presentations as ineffective for promoting learning (Solomon 1994, Osborne 1996). The constructivist focus on active learning has led many to equate listening or observing with being passive. “Learning is an effortful and mindful process and students should be encouraged to construct their own knowledge and skills through active processing, rather than being passive listeners” (Vosniadou, Ioannides, Dimitrakopoulou & Papademetriou 2001, p.382). A possible reinterpretation of this concern is that for various reasons in some presentation settings learners do not invest sufficient cognitive effort to engage with the material. The constructivist solution is to withhold information or guidance to force students into investing significant mental effort before they can continue. When used appropriately, this technique can encourage germane cognitive load leading to robust learning. An example is the prediction component of Interactive Lecture Demonstrations (Thornton & Sokoloff 1998), without which demonstrations have been shown to be ineffective (Crouch et al. 2004). However when misused this method may have both cognitive and affective drawbacks, producing extraneous cognitive load and reducing student motivation. The TEAL project (see p.131) is one such example.

Constructivists working on the problem of multimedia design have focused on simulations since they require learners to be physically active in the learning process. Simulations allow the learner to explore and construct his or her own ideas. They can slow down processes and allow learners to directly see the consequences of their adjustments of different variables. Furthermore, simulations can represent visually and dynamically important concepts that would otherwise be invisible.

Simulations with intricate controls and minimal guidance represent the other end of the cognitive load spectrum. On the topic of electric circuits, Ronen & Eliahu (2000) found that simulation feedback helped students recognize and change their misconceptions. A further study confirmed that over the course of a semester, students who used both real and virtual experiments developed a stronger conceptual understanding than those who used real experiments alone (Zacharia 2007).

However, an important finding in interactive settings is that learners often require more scaffolding to focus on conceptual issues (Lowe 2004, van Joolingen, de Jong & Dimitrakopoulou 2007). Ronen & Eliahu (2000) found that some of their students did not have the requisite prior knowledge to learn with the electric circuits simulation. With a projectile motion simulation, Yeo, Loss, Zadnik, Harrison & Treagust (2004) found that students interacted superficially and retained their intuitive conceptions. Only after researcher intervention did they focus on the salient conceptual issues in the program.

A case study of two students working collaboratively with a Newtonian mechanics simulation revealed that it took more than an hour to develop a scientifically accurate definition of acceleration, during which time many ultimately unsuccessful lines of reasoning were explored (Roschelle, 1992). Cognitive load researchers would argue that much of the interaction with the simulation was resulting in extraneous rather than germane processing, and that scaffolding could be used to redress this imbalance.

Raise germane cognitive load

More recently, multimedia researchers have considered ways of increasing germane cognitive load during instruction (see the special issue of *Learning and Instruction*, *16*(2)). Misconceptions added to multimedia treatments have the potential to increase germane cognitive load by highlighting possible differences between scientific theories and a learner's prior knowledge. On the other hand, they may impose an extraneous cognitive load on the learner, inhibiting him or her from building a correct, coherent mental model.

Posner et al. (1982) proposed that dissatisfaction with existing mental models is the first step towards conceptual change. This has been supported by a number of experiments in which different methods have been used to create cognitive conflict (Guzzetti et al. 1993, Limon 2001). Although there is debate over which methods are most effective, many cognitive conflict tactics employed in classrooms have demonstrated improved performance compared to traditional instructional approaches (Duit & Treagust 2003, Vosniadou & Verschaffel 2004). Therefore, misconception treatments should activate students' prior knowledge and help them recognize any disparity between their ideas and correct scientific theories.

The Dialogue and Refutation were hypothesized to engender more cognitive load than the Exposition but not as much as a simulation. Since linear multimedia does not require learners to make selections or enter input, it should not be as demanding as a simulation. By including ideas that conflict with accepted scientific knowledge, however, these multimedia require learners to pay attention to more information and discriminate alternative from correct ideas. In addition, since the alternative ideas in the multimedia are commonly held by participants, they should be better able to understand the presentation and therefore pay more attention to the ensuing discussion.

There is a risk, however, that adding misconceptions to a concise scientific presentation may interfere with learning. When viewing a misconception treatment, learners must select from a greater number of words and pictures to form coherent mental models. They must also pay attention for a longer duration to see the same amount of correct scientific information. Furthermore, in both of the misconception-based treatments, force diagrams and animations were shown to illustrate multiple misconceptions. In the Dialogue, misconceptions were presented as the genuine beliefs of one of the dialogue participants without cautioning students that not all of the information in the treatment was correct. Resolutions were reached later, through discussions between the student and tutor.

In addition, not all students have the same misconceptions, so a discussion that might be useful for some students would likely be irrelevant for others. Multimedia researchers have found that some instructional guidance that benefits novices can hinder learners experienced in the area, in what is called the *expertise reversal effect* (Kalyuga et al. 1998, Kalyuga et al. 2003). Therefore, one might expect a discussion of misconceptions to be beneficial for novice learners, like Fundamentals, but detrimental for those more experienced, like Advanced students.

Multimedia containing misconceptions is significantly different from refutation text because multimedia is transient in time. While reading refutation texts, learners can easily refer back and forth between misconceptions and correct scientific ideas. In contrast, viewers of misconception-based multimedia must develop their understandings as the multimedia progresses. This increases the likelihood that the added material may misdirect or overload learners.

The Extended Exposition aimed to increase cognitive load in a different manner to the Dialogue and Refutation, by increasing interest. Schank (1979) argued that interest plays a key role in allocating limited cognitive resources, making it essential for learning. Mitchell (1993) extended this line of thought, proposing that a learner's interest can be caught and held during instruction to improve retention. This purpose could be well served by including highly interesting but unimportant information, often called 'seductive details,' in instruction (Schraw & Lehman 2001). The benefits of seductive details have not been empirically verified however. Relevant studies of seductive details in multimedia have shown the extra material to have detrimental effects (Mayer, Heiser & Lonn 2001).

9.3 Method

9.3.1 Participants and design

The participants were 678 first year students at the University of Sydney from three physics streams: Fundamentals, for students with little prior formal instruction in physics; Regular, for students with senior high school physics backgrounds; and Advanced, for students who excelled in senior high school subjects and physics in particular. Students in the three streams are from a wide range of degree programs including Engineering, Medical Science, and Arts, and almost all completed high school locally. The three streams are comparable in most respects except gender ratios. Approximately 60 percent of the Fundamentals stream is female compared to 30 percent in the other streams.

Since the experiment was conducted in an authentic setting and participants were allowed to withdraw at any time, the data required filtering prior to analysis. Participants were removed from the data set for failing to complete the post-test (116), watching more than one multimedia treatment (75), not watching the multimedia in its entirety (30), spending less than four minutes completing the pre- or post-test (57), failing to answer all questions (6), or scoring higher than 95% on the pre-test (30). Students were able to watch more than one multimedia treatment by using the back button in their browser or by manually changing the website url.

Using a website to administer the materials for this study had several advantages. It allowed for large numbers of students from authentic lecture courses in physics to be surveyed. The times of submission for each question were easily recorded and participants were able to complete the study in their own time at their own pace. Assignments to video treatments were completely randomized using a Hypertext Preprocessor (PHP) script. Some drawbacks of the website setup included the high bandwidth required to view the multimedia treatments over the Internet. Participants either required broadband at home or had to complete the study from an on-campus access lab.

9.3.2 Pre/post test

The same 26-question multiple-choice test was used as a pre- and post-test. Twenty-two questions were drawn from the Force and Motion Conceptual Evaluation (FMCE, Thornton & Sokoloff, 1998; questions used were: 1–4, 7–10, 14, 18–21, 23–25, 27–29, 41–43); three questions were from the Force Concept Inventory (FCI, Hestenes, Wells, & Swackhamer, 1992; Halloun & Hestenes, 1995; questions 13, 17, 25); and one question was written by the researchers. All assessed different aspects of Newtons first and second laws and were therefore considered as a coherent assessment tool. The FMCE and FCI have been used with thousands of students in physics education research (e.g. Hake, 1998), and their validity is well established (Henderson, 2002). The answers to 12 questions were explicitly stated in each multimedia, while the other 14 required the application of physics principles to new situations. Four of the questions were presented on separate web pages, while the other 22 were grouped onto six pages because they shared a common stem (e.g. three questions pertaining to a coin toss were on the same page). Each question was worth one mark. See Appendix B.7 for the pre- and post-test as it appeared online.

9.3.3 Multimedia treatments

To explicate the physics concepts, all of the multimedia treatments examined three examples: a book pushed across a table at constant speed, a juggling ball thrown upwards and caught, and a toy car rolling up and down a ramp. The treatments ranged in length from seven to eleven and a half minutes. The scripts were written with reference to several textbooks (Hewitt 1997, Young & Freedman 2000, Halliday, Resnick & Walker 2003) and were critiqued by a panel of three physics educators, each with over 30 years of experience. The scripts were iteratively compared and contrasted throughout the writing process to ensure all treatments contained exactly the same accurate physics information. After the treatments were completed, they were again critiqued by physics educators to ensure there existed no inconsistencies in physics content.

The Exposition was designed to be very similar to a concise presentation a well-prepared lecturer might make on the topic of Newton's Laws. It included graphs, force diagrams, animations, and live action demonstrations, along with 'talking head' narration. The Extended Exposition and the Refutation consisted of the Exposition plus additional material. Interesting information beyond the assessed learning outcomes was added in the Extended Exposition to make it equal in length with the Dialogue. Thus it served as a control for the time students spent thinking about the physics and, in conjunction with the Exposition, it allowed for an investigation of the applicability of the coherence principle (Mayer 2001, Mayer 2003) in an authentic physics setting. Common misconceptions were explicitly raised and refuted in the Refutation to investigate this method of recognizing anomaly between prior knowledge and scientific theory. The most common alternative conceptions were selected from the literature (Trowbridge & McDermott 1980, Trowbridge &

McDermott 1981, McClosky 1983, Halloun & Hestenes 1985, diSessa 1996), including:

- motion requires a force
- confusing[1] velocity with acceleration
- confusing[1] position with velocity
- acceleration is zero if velocity is zero
- increasing force is required to achieve constant acceleration.

Previous studies on refutation texts were used to inform the writing of this script (Diakidoy et al. 2003, Guzzetti et al. 1997).

The Dialogue was entirely different in structure to the other three treatments, utilizing a simulated discussion between an inquisitive student and a tutor. Over the course of the discussion, the student's misconceptions, the same as those in the Refutation, were revealed and corrected. These ideas were not only stated by the student in the Dialogue with no warning that they were alternative ideas, they were also animated and shown on graphs and diagrams. Only through a Socratic dialogue with the tutor did the correct scientific ideas emerge (Hake 1992). The added material in the Dialogue ran counter to the recommendations of established multimedia design principles because it risked confusing students with alternative ideas they may not have been entertaining. Despite the potential of these ideas to induce an extraneous cognitive load, I suspected based on the results from the previous chapter and theoretical underpinnings that for most students the increased load would be germane. With the inclusion of alternative conceptions, the total length of the Dialogue was eleven minutes and 22 seconds. Parts of the dialogue script were inspired by transcripts of a student's interviews on Newton's Laws (diSessa 1996). Where possible the same phrases as in the Expositions were used in the Dialogue.

A summary of the similarities and differences among the four treatments is shown in Table 9.1. The scripts are included in Appendix C.2 and all treatments are on the DVD on the back cover.

[1] It should be noted that the word 'confusing' is used in a limited sense to denote specific, well-documented difficulties in differentiating one idea from another.

Treatment	Exposition	Extended Exposition	Refutation	Dialogue
Number of speakers	1	1	1	2
Length	7:02	11:22	9:33	11:22
Addresses misconceptions	No	No	Yes	Yes

Table 9.1: Summary of multimedia treatment characteristics.

9.3.4 Procedure

All students taking first year physics at the University of Sydney were asked to access a website for one mark towards their first assignment. This assignment was due one week following the experiment announcement. A consent form on the opening page informed students that

- the study would take between 30–45 minutes to complete,
- the study should be completed individually without referring to textbooks or online resources,
- performance on the pre- and post-tests would be kept confidential,
- participation in the study was voluntary and that they could withdraw at any time with no penalty.

Between pre- and post-tests, participants were randomly assigned to one of the four multimedia treatments. They could watch their prescribed treatment using Windows Media Player or QuickTime. Students had the ability while watching the multimedia to pause, rewind, and replay the presentation as desired but the total amount of time spent was logged to a database. Participants' answers and the times at which they were submitted were written to a MySQL database. This allowed for determination of the time spent on the pre- and post-tests and the time spent watching the multimedia treatment. After completing the post-test, each student received their mark on that test along with helpful suggestions about resources they could use to improve their understanding. Students who scored below 40% were not told their exact mark, and additional aids were recommended. A record of all students who accessed the website, regardless of whether they completed the study, was sent to the course coordinator who allocated participation marks.

9.4 Results and analysis

The scores on the pre- and post-tests were not normally distributed. This was due to the large number of students with widely varied abilities. Non-parametric tests revealed no significant differences among the pre-test results across the four treatment groups, however post-test results were significantly different (Kruskal-Wallis $\chi^2 = 8.625, p = .035$, Median Test $\chi^2 = 9.565, p = .023$). Gender composition was not significantly different across the four treatment groups, nor was the time spent completing the pre- or post-tests.

9.4.1 Differences between treatments

To determine the relative effectiveness of the multimedia treatments, a gain score for each student was computed by subtracting their pre-test mark from their post-test mark (each of which had a maximum of 26 marks). Gain was normally distributed for each treatment group. The sample size, mean gain, standard deviation, and Kolmogorov–Smirnov Z and p-value (to test for normality) are shown in Table 9.2.

Treatment	Sample size (n)	Median pre-test	Median post-test	Gain		K–S	
				M	SD	Z	p
Dialogue	92	8.5	16	4.77	4.59	1.057	.214
Refutation	86	7.5	14	4.41	4.01	0.914	.373
Extended Exposition	95	8.0	12	2.41	3.72	1.300	.068
Exposition	91	8.0	9.0	1.77	2.65	1.075	.198

Table 9.2: Summary of dependent variables for Newtonian mechanics multimedia treatments.

Using a one-way ANOVA the gains of the treatments were compared yielding a significant difference between treatments ($F(3, 461) = 13.625, p < .001$). The Games-Howell post-hoc procedure, which doesn't assume equal variance, showed the gains for students who watched the Dialogue or the Refutation were significantly greater than those who received the Exposition ($p < .001$) or Extended Exposition treatments ($p = .001$ for the Dialogue, $p = .004$ for the Refutation). The effect size for these differences in comparison to the Exposition was $d = .83$ for the Dialogue and $d = .79$ for the Refutation.

9.4.2 Gain dependence on prior knowledge

Students from the three physics streams had different levels of prior physics instruction, allowing for an investigation of the dependence of gain on prior knowledge. It was expected that the Fundamentals students, with the least prior physics instruction, would hold the most misconceptions and therefore benefit most from misconception-based instruction. Regular students, having completed Newtonian mechanics in high school, represented a mixture of prior knowledge. It was therefore unclear which treatment would be most advantageous for them. Advanced students, with significant accurate prior knowledge, were expected to achieve greater learning gains with the concise treatment. The mean gains for each treatment, separated by physics stream, are shown in Figure 9.1.

Fundamentals students who watched the Dialogue or Refutation had significantly greater gains than those who watched the Exposition ($F(3, 109) = 6.609, p < .001$). Similarly, in the Regular stream the Dialogue and Refutation students achieved significantly greater gains than Exposition students ($F(3, 163) = 7.262, p < .001$). Students from the Advanced stream did not show significantly different gains between treatments, though the trends in means observed are similar to those above ($F(3, 83) = 2.069, p = .111$). The lack of significant difference might be due to the small sample size in this stream and a possible ceiling effect on the post-test. The median score for the Advanced stream on the post-test was 85%, compared to 23% and 54% for the Fundamentals and Regular streams, respectively.

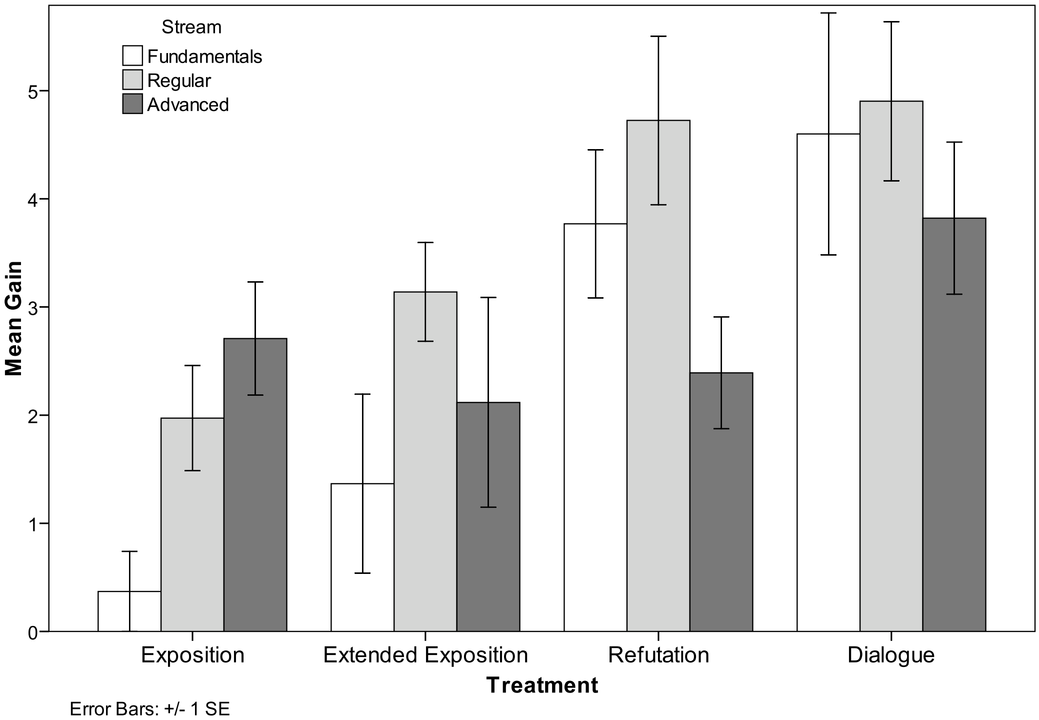


Figure 9.1: Gains for each treatment by physics stream.

9.4.3 Gain in confidence

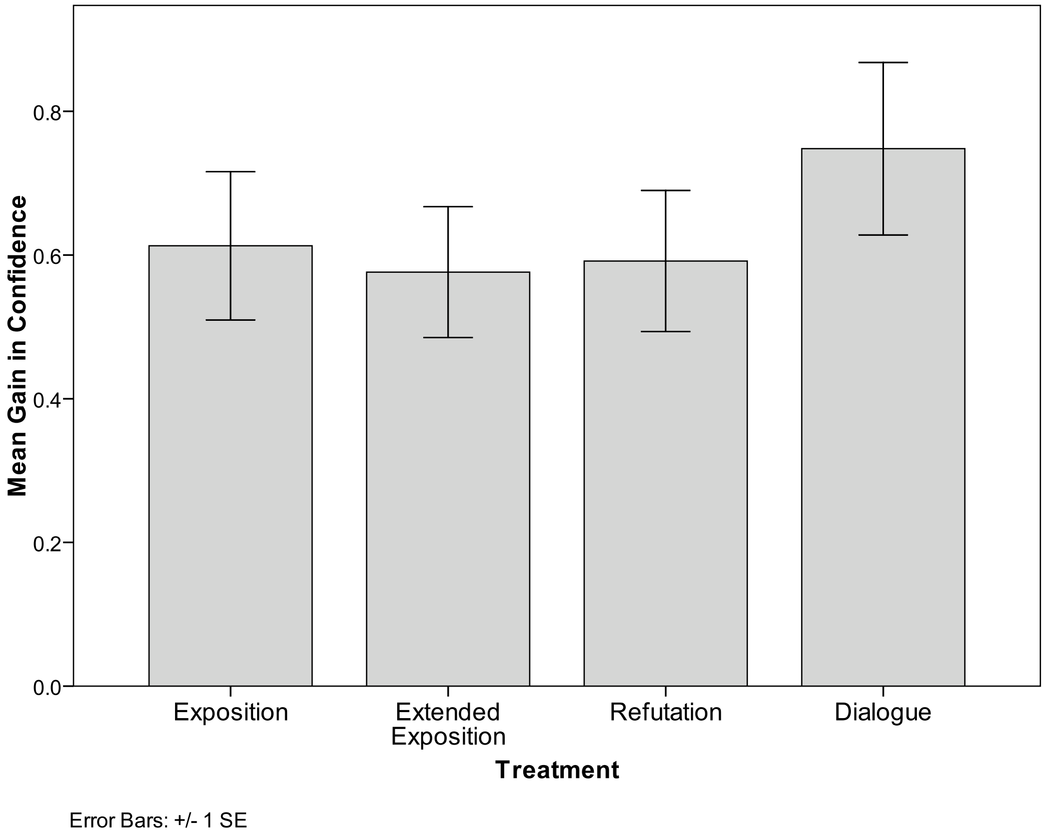


Figure 9.2: Gain in confidence from pre- to post-test for each treatment

Changes in confidence were measured by subtracting a student's average pre-test confidence rating from his or her average of post-test confidence rating. Figure 9.2 shows that gains did not differ significantly among the different treatment groups.

9.5 Discussion

9.5.1 Theoretical Implications

Using questions from standard mechanics conceptual inventories, students from three lecture courses in first year physics were tested before and after watching a short multimedia treatment about Newton's first and second laws. Results show that overall students achieved greater gains by watching a treatment that addressed misconceptions than one which presented only correct scientific information. This suggests that the increased cognitive load incurred with misconception-based treatments was germane rather than extraneous on the average for students with all levels of prior knowledge. The results of this study are consistent with the findings of conceptual change research that suggest cognitive conflict is essential to conceptual change (Guzzetti et al. 1993). As both Refutation and Dialogue treatments produced similar effect sizes, it seems that both methods of recognizing the anomaly between prior knowledge and scientific theory are equally effective in non-interactive multimedia.

The findings highlight the need to consider what constitutes extraneous information in the context of cognitive load theory. In the standard practice of teaching physics, misconceptions are not considered essential teaching material. They are addressed when the need arises, in response to student questions or answers on assessment. Even then, feedback may only address the specific problem without clearly explaining a misconception in its entirety. Almost all textbooks, including those used by the students in this study, do not include discussions of misconceptions. However, the addition of incorrect information to form the Dialogue and Refutation treatments was essential for students to engage in germane processing; it did not impose an onerous extraneous load on students.

One might expect the discussion of misconceptions to be particularly irrelevant for the Advanced students, given their experience with Newtonian mechanics and excellent performance on high school assessment tasks. However no analogue of the expertise reversal effect was found. Advanced students in the misconception treatments achieved non-significantly greater gains than their peers in the concise treatment. Although studies have shown misconceptions to be quite persistent, it is unlikely that Advanced students held misconceptions to anywhere near the same degree as Fundamentals students. From these results, the explicit discussion of misconceptions seems to be an effective instructional strategy whether students actually hold the misconceptions or not.

With respect to the Exposition and the Extended Exposition, the effects of the coherence principle were not observable using the multiple choice pre/post-tests in this setting. This result suggests that interesting but irrelevant information might encourage students to pay attention to online multimedia when they are watching it in their own time. Alternatively, the multiple-choice tests may not have been sensitive to the differences in learning between the two treatments. Replications of laboratory studies that investigated the impact of additional interesting information conducted in authentic learning environments could shed light on the issue. This highlights another possible area in which seemingly extraneous information in a laboratory setting might yield a germane cognitive load in authentic learning contexts. In a laboratory, a learner's attention is focused and therefore his or her motivation to engage with instructional material is less important than in an unstructured environment. A previous study has found that some well-established multimedia principles fail to generalize easily to authentic settings (Tabbers et al. 2004). Further research is required to determine whether the coherence principle holds in authentic settings (Muller, Lee & Sharma 2007).

In future studies, an attempt to measure the cognitive load of students may help to understand and interpret results. In a setting like that of the present experiment, this would most likely be achieved through self-reported rating scales, however other techniques could be used in a laboratory setting (Paas, Tuovinen, Tabbers & Van Gerven 2003).

This study helps to understand an 'active ingredient' in the reform methods developed to achieve conceptual change through lecture instruction (Hake 1998). Physics education research has been criticized for comparing instructional strategies where several variables have been altered simultaneously (Guzzetti et al. 1993). Reform teaching methods include various combinations of hands-on activities, discussions with peers, increased instructor feedback, demonstrations involving learning cycles, written worksheets, and classroom communication systems, leaving in doubt the essential factors that enhance learning. The results of this study suggest that part of the benefit of interactive lecture classes and tutorials is likely derived from students observing discussions between other students and tutors in which misconceptions are addressed. Discussions of this sort are quite rare in traditional lecture classrooms (Graesser & Person 1994, Muller 2005).

9.5.2 Practical Implications

The results of this study suggest that, unlike other non-essential information, discussing misconceptions does not interfere with learning when added to multimedia. When designing multimedia for science education areas, developers should therefore address common misconceptions explicitly in their explanations of appropriate topics. Although their inclusion in multimedia results in longer interventions with more words and diagrams, they serve a useful pedagogic purpose, aiding learners to consider scientific conceptions in light of their prior knowledge.

In addition, although interactive methods in lectures have demonstrated substantial gains in conceptual understanding over traditional methods, this study suggests that raising misconceptions in traditional-style lectures should increase student conceptual understanding. This is an important result for teachers who find it difficult to implement interactive methods due to restrictions on time, money, and technology, often coupled with large class sizes. Multimedia interventions that address misconceptions, like those investigated in this study, could be used in lectures to highlight key conceptual difficulties. Alternatively, they could be used to provide conceptual scaffolding for interactive multimedia.

Despite the verified advantages of reform teaching methods and refutation texts, uptake of these strategies has been quite limited. The practical drawbacks of refutation texts are clear: they require more research and writing to produce and they result in heavier, bulkier books. Challenges of implementation for reform teaching methods are similar. They require substantial investments of time and money, specialized training, and often result in a decrease in the number of learning outcomes that can be achieved in the same number of contact hours. The Internet offers a new means of circumventing some of these difficulties. Multimedia is almost as readily available as text and is not cumbersome to carry like a textbook. Adding misconceptions only increases the duration of instruction, which as demonstrated in this study can dramatically increase the learning gains in an authentic setting.

The success of conceptual change interventions is often heavily dependent on the expertise of the teacher (Limon 2001). Addressing misconceptions through multimedia rather than teacher-led discussion reduces the burden on teachers and increases the likelihood of success. Teachers are often hesitant about conducting conceptually challenging discussions due to concerns about time constraints or their own mastery of the subject (Weaver 1998).

It is important to note that although misconception-based multimedia on average resulted in greater learning gains, it is not a stand-alone solution to conceptual difficulties. The process of moving from alternative ideas to a coherent scientific view is complex and it remains only partially understood. Undoubtedly, discussions among students and between students and teachers are important for developing accurate conceptual understandings. Multimedia that addresses misconceptions is simply one resource that may help students along the path to scientific reasoning. Furthermore, the misconception-based techniques presented in this study may be useful adjuncts to simulations or online discussions, to help focus learner attention on salient conceptual issues.

Chapter 10

Newtonian mechanics multimedia: Second iteration

10.1 Introduction

The study reported in the previous chapter established that addressing misconceptions in multimedia can promote conceptual change in an area where misconceptions are common and robust. Effect sizes were large especially considering the pre- and post-test measures were comprised of multiple-choice questions from validated conceptual inventories, the multimedia treatments were short, and students took part in uncontrolled learning environments. Interestingly, gains in confidence did not differ among the four treatments despite the large differences in learning.

A significant concern with the experiments conducted in Chapters 8 and 9 was that the addition of alternative conceptions was hypothesized to raise cognitive load but cognitive load was not directly measured. In the third iteration of the design experiment, described in this chapter, the cognitive load induced with each treatment was measured directly.

The previous Newtonian mechanics study confirmed that the extra material in the Dialogue and Refutation is beneficial for learning. Therefore a new matter for investigation was whether the inclusion of alternative conceptions has any special status compared to other forms of relevant additional information. In this study, the

Extended Exposition was replaced with a Worked Examples treatment in which the additional information was directly relevant to the post-test questions.

All multimedia treatments were filmed again with different actors and modified diagrams. Gender roles were reversed from the previous study, with a female playing the tutor in all treatments, and a male acting as the student in the Dialogue. This allowed for a confirmation that the method, rather than some uncontrolled variable accounted for the different gains observed in the first Newtonian mechanics study.

Interviews were conducted with small groups of students to investigate perceptions of the multimedia and explore the types of conceptual change that are possible through simply viewing multimedia. The language students use to describe the conceptual change process is a rich source of data that can be used to interpret the large-scale quantitative data (Guzzetti & Hynd 1998).

In this experiment each of the three streams was used to test a different hypothesis. In what amounts to a replication of the previous Newtonian mechanics study, Fundamentals students were randomly assigned to one of the four new multimedia treatments. Students from the Regular class were assigned to either the Dialogue or Exposition and only half were randomly assigned to take the pre-test. The pre-test is an aspect of these studies that might not be used in authentic learning environments. Therefore, the goal of the Regular student study was to decouple the effect of the pre-test from the treatment effects. Finally, in the Advanced class, only the Dialogue and Refutation were compared. For lower prior knowledge groups, I hypothesized that the inclusion of alternative conceptions would lead to improved gains regardless of the way in which they were presented. However, for Advanced students the presentation format may be more important since learning opportunities are more limited. It was hypothesized that the Dialogue would create increased germane load because the situation better resembles a social learning environment.

Thus the third and final design experiment iteration involved four topics of research:

1. Do students invest more mental effort when watching the Dialogue than the Exposition? How do post-test scores compare for students in the two groups?

2. Does misconception-based multimedia result in higher mental effort and post-test scores than non-misconception-based multimedia?
3. How does the pre-test affect the mental effort invested by students and their post-test scores?
4. Is there a difference in effectiveness between the Dialogue and the Refutation for Advanced students?

To answer each of these questions, different student samples were used. The general method was the same for all experiments and is summarized in the next section. The samples of students employed and the results observed are grouped into the four numbered research question sections that follow.

10.2 Method

Pre/post-tests

The same 26-question multiple-choice test described on page 169 was used as a pre- and post-test (see Appendix B.7). Four of the questions were presented on separate web pages, while the other 22 were grouped onto six pages because they shared a common stem (e.g. three questions pertaining to a coin toss were on the same page).

Procedure

In 2007, as part of an assignment, students were asked to access a website and participate in the study. Students received credit for visiting the website regardless of whether or not they participated. They were informed that they could withdraw from the study at any time with no penalty.

After logging in, students completed the multiple-choice pre-test. Regular and Advanced students answered all 26 pre-test questions. Fundamentals students, however, answered only ten questions because they had seen the other sixteen three weeks previous on an unrelated diagnostic test. This difference was not felt to impact on the study because a) simply taking mechanics conceptual tests does not improve results on later repetitions of the same test (Henderson 2002), and b) pre-test results were only used to establish the homogeneity of samples before the multimedia interventions.

After instruction, students were asked to rate the mental effort they invested in the multimedia on a nine-point semantic scale ranging from ‘extremely low mental effort’ to ‘extremely high mental effort.’ Similar rating scales have been used to measure the cognitive load of instruction in numerous previous studies (Paas et al. 2003).

All students then completed the standard 26-question post-test. On each of the ten pages of questions, students were asked to rate the mental effort they invested in answering the question(s) using the same nine-point scale as above. At the end of the post-test, students were informed of their score on the test and helpful resources were recommended.

Multimedia treatments

All multimedia except the Worked Examples treatment was very similar to that described in the previous study. The same basic scripts were used though some phrases were changed for clarity or to shorten the longer treatments. The multimedia were encoded as QuickTime and Windows Media Player files with dimensions of 640 by 360 pixels, to be streamed through any standard web browser. A summary of the similarities and differences among the four treatments is shown in Table 10.1. All treatments are on the DVD on the back cover.

Treatment	Exposition	Worked Examples	Refutation	Dialogue
Number of speakers	1	1	1	2
Length	7:30	10:00	10:00	10:22
Addresses misconceptions	No	No	Yes	Yes

Table 10.1: Summary of new multimedia treatment characteristics.

The additional Worked Examples treatment was created by adding material to the Exposition. In the Worked Examples treatment, two numerical problems were solved. In these examples, the time of flight of the juggling ball and force required to push the textbook across the table at constant speed were determined. These worked solutions involved the manipulation of formulae, and the repetition of diagrams and concepts directly evaluated on the post-test. For example, to solve the textbook problem, the balance of forces was shown diagrammatically and with formulae. Numbers were substituted into the equations and the final solutions for the forces were shown to be equal and opposite in the vertical and horizontal directions. To calculate the time of flight of the ball, the forces on the ball were repeated and used to find its acceleration of $-9.8m/s^2$ while in the air. The extra instruction increased the duration of the Worked Examples multimedia to ten minutes.

10.3 Research question 1

The primary goal of this study was to compare the effects of the Dialogue and Exposition treatments. Both Fundamentals and Regular students viewed these multimedia so the two groups were considered together in the analysis. Although students in these two streams have different levels of prior knowledge, distributions on diagnostic tests indicate that their understanding of Newtonian mechanics is similar. Regular students who did not receive a pre-test were excluded.

Participants

One hundred and eighty-five first year physics students formed the sample for this analysis, with 108 from the Regular stream and 77 Fundamentals. Since participation in the study was voluntary and students could discontinue at any time, the data were filtered prior to analysis. Students were removed from the sample for failing to watch the multimedia treatment in its entirety (6), taking less than two minutes to complete the pre-test (2), leaving three or more questions blank on the post-test (10), or achieving a gain more than three standard deviations greater than the mean of the remaining sample (30 – this outlying group was clearly identifiable in histograms of gain and post-test score). Of the remaining 137, 72 students were randomly assigned to the Dialogue treatment while 65 received the Exposition.

10.3.1 Results and analysis

The lengths of time spent on pre- and post-tests were not significantly different for students in the two treatment groups. A Mann–Whitney test revealed a significant difference in the time spent watching the multimedia ($U(135) = 1558, p < .01$). Median viewing times for the Exposition and Dialogue were 10:12 and 11:48 respectively. A *t*-test revealed pre-test scores were not significantly different for the two treatment groups.

Did students differ in mental effort or post-test score?

A *t*-test revealed that Dialogue students reported investing significantly more mental effort than Exposition students ($t(135) = 2.50, p = .014, d = .43$) while watching the multimedia. Post-test scores were also significantly different, favouring the Dialogue treatment ($t(135) = 2.60, p = .010$). The effect size for this difference was $d = .45$. No difference between the groups was observed in the mental effort invested on the post-test. The means and standard errors for invested mental effort during instruction and the post-test scores are shown in Figure 10.1.

10.4 Research question 2

The second objective of this experiment was to determine whether treatments that involve misconceptions (the Dialogue and Refutation) result in greater learning and mental effort than those that do not (the Exposition, and Worked Examples).

Participants

All 164 participants in this experiment were from the Fundamentals stream. Again, due to the voluntary nature of the study, students were removed from the sample prior to analysis for: failing to watch the multimedia in its entirety (4), completing

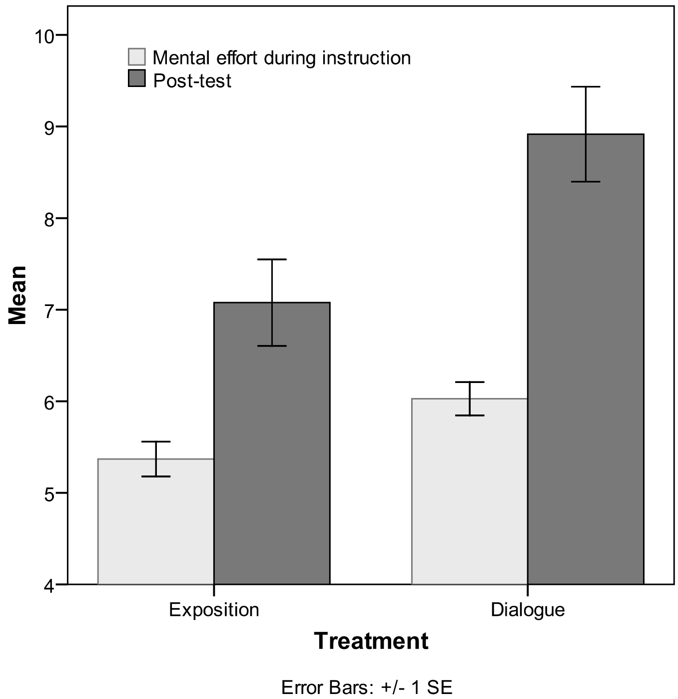


Figure 10.1: Post-test scores and mental effort invested in instruction for students who received the Exposition and Dialogue multimedia. The maximum possible post-test score was 26, while mental effort ranged from 1 (extremely low mental effort) to 9 (extremely high mental effort).

the pre-test in under two minutes (2), leaving three or more post-test questions blank (11), and achieving a gain more than three standard deviations greater than the mean of the remaining sample (2). The remaining 145 students were randomly assigned to view the Exposition ($n = 33$), Worked Examples ($n = 34$), Refutation ($n = 42$), and Dialogue ($n = 36$) treatments. Since only Fundamentals students participated, all completed the ten-question pre-test.

10.4.1 Results and analysis

The lengths of time spent on pre- and post-tests were not significantly different for students in the four treatment groups. A Kruskal-Wallis (non-parametric) test revealed that the time spent watching the multimedia was not significantly different across the groups; median viewing times were 11:06, 12:30, 12:12, and 11:36 for the Exposition, Worked Examples, Refutation, and Dialogue treatments respectively.

Results from this study were analysed in two ways. First, the treatments that involved alternative conceptions (the Refutation and Dialogue) were compared to the treatments that did not (the Exposition and Worked Examples). Second, all treatments were compared with each other using a one-way ANOVA.

Did the inclusion of alternative conceptions affect mental effort or post-test scores?

T-tests revealed that students who watched a multimedia treatment involving alternative conceptions invested significantly higher mental effort ($t(143) = 2.62, p = .010, d = .44$) and achieved significantly higher post-test scores ($t(139.7) = 2.17, p = .032$) than those who watched a multimedia treatment without alternative conceptions. The effect size for the difference in post-test scores was $d = .36$. This finding is shown in Figure 10.2. Again no difference was found in the mental effort invested by students in the two groups on the post-test.

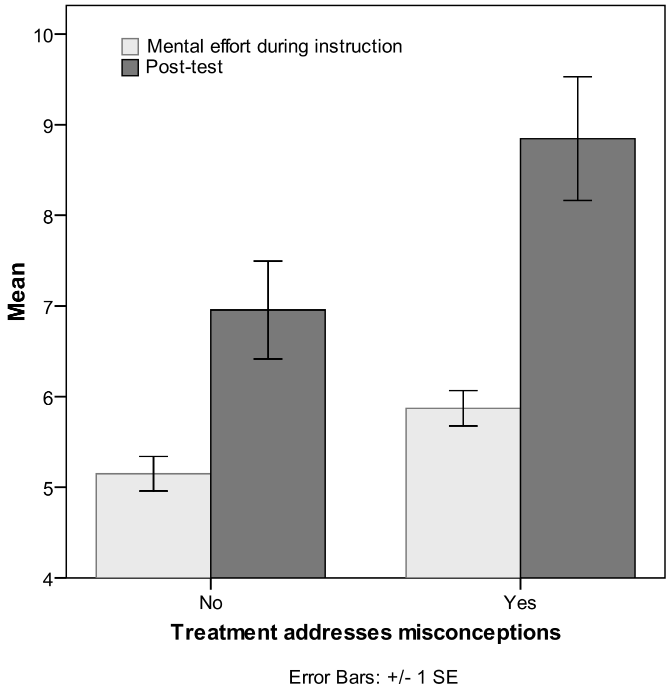


Figure 10.2: Post-test scores and mental effort invested in instruction compared between treatments that addressed alternative conceptions and those that did not.

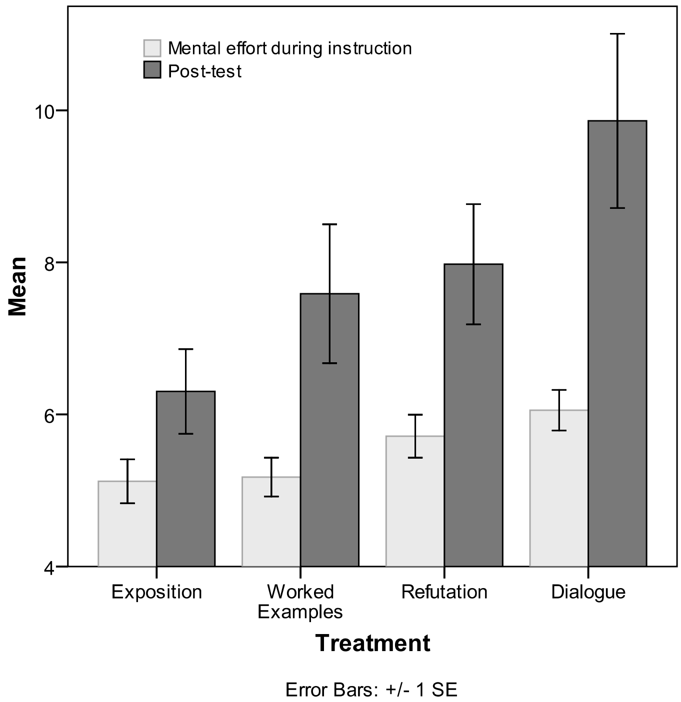


Figure 10.3: Post-test scores and mental effort invested in instruction compared across all four treatments with Fundamentals students.

Did mental effort or post-test scores differ across all four treatments?

A one-way ANOVA yielded no significant difference in invested mental effort during instruction, post-test scores, or invested mental effort in the post-test across the four treatment groups. Results approached significance, however, for mental effort invested during instruction ($F(3, 141) = 2.55, p = .056$), and post-test score ($F(3, 141) = 2.65, p = .051$). Means and standard errors for these measures are shown in Figure 10.3.

10.5 Research question 3

The Regular students participated in a study to determine the effect of the pre-test on post-test scores and the investment of mental effort. Half were randomly assigned to complete the test before the multimedia, and half were directly assigned to an instructional treatment. Only the Dialogue and Exposition were used to ensure a reasonable sample size.

Participants

Two hundred and thirty-one Regular students participated in this study including the 108 students used to answer research question 1. Students were removed from the sample prior to analysis for: failing to watch the multimedia in its entirety (11) or leaving three or more post-test questions blank (7). Participants could not be screened based on their pre-test times or gain scores because half of the sample did not complete the pre-test. The remaining 213 students were randomly assigned to view either the Exposition or the Dialogue with half of each group completing the pre-test. The number of students in each condition and the times spent on the multimedia and post-test are summarized in Table 10.2.

Pre-test condition	Treatment	Sample size (n)	Time spent (min) on multimedia		on post-test	
			M	SD	M	SD
Pre-test	Exposition	42	14.1	11.3	10.2	6.7
	Dialogue	57	15.2	7.5	11.9	7.4
No pre-test	Exposition	57	10.5	7.5	19.6	7.2
	Dialogue	57	12.9	2.8	19.6	9.1

Table 10.2: Regular students randomly assigned to the pre-test or no pre-test condition and either the Dialogue or the Exposition.

10.5.1 Results and analysis

A two-way ANOVA revealed that students who completed the pre-test took less time to complete the post-test ($F(1,209) = 63.93, p < .001, \eta_p^2 = .236$), but more time to watch their multimedia treatment ($F(1,209) = 7.858, p = .006, \eta_p^2 = .037$).

The treatment they watched had no effect on the time spent watching the multimedia or completing the post-test.

There was a small but significant interaction effect of the treatment and pre-test condition on the post-test score ($F(1, 209) = 5.265, p = .023, \eta_p^2 = .025$). No main effects were observed for either treatment or pre-test. Mental effort scores both during and following instruction were not significantly different. A main effect for the pre-test on post-test mental effort approached significance ($F(1, 209) = 3.080, p = .081, \eta_p^2 = .015$).

Pre-test condition	Treatment	Post-test		Mental effort invested during:			
				instruction		the post-test	
		M	SD	M	SD	M	SD
Pre-test	Exposition	11.4	8.0	5.50	1.47	4.73	1.26
	Dialogue	13.8	7.4	5.54	1.81	4.62	1.33
No pre-test	Exposition	13.5	8.4	4.93	1.77	4.79	1.15
	Dialogue	11.3	6.6	5.67	1.87	5.20	1.17

Table 10.3: Summary of results for the Regular class watching the Exposition or Dialogue treatments under pre-test or no pre-test conditions.

10.6 Research question 4

The last objective of this experiment was to determine whether the format in which misconceptions are presented affects mental effort or learning for high prior knowledge students.

Participants

All 76 participants in this experiment were from the Advanced stream. Students were removed from the sample prior to analysis for: failing to watch the multimedia in its entirety (5) or leaving three or more post-test questions blank (3). The remaining 68 students were randomly assigned to view the Dialogue ($n = 33$) or the Refutation ($n = 35$).

10.6.1 Results and analysis

Students in the Dialogue and Refutation conditions spent similar times on the pre- and post-tests and on the multimedia. As shown in Table 10.4, mental effort invested during and after instruction, and gain scores were nearly identical for the two groups.

Treatment	Pre-test		Post-test		Gain		Mental effort invested during: instruction		Mental effort invested during: the post-test	
	M	SD	M	SD	M	SD	M	SD	M	SD
Dialogue	17.5	7.1	20.3	6.5	2.84	4.22	4.59	1.74	4.05	1.41
Refutation	18.9	6.3	21.8	4.7	2.86	3.55	4.66	1.85	4.02	1.77

Table 10.4: Summary of variables for the Advanced class watching the Refutation or Dialogue treatments.

10.7 Interviews

In addition to the quantitative data collected in this study, four interviews with small groups of Fundamentals students (who were not involved in the previous experiments) were conducted, each focusing on a different multimedia treatment. Interviews ran for approximately one hour and were divided into four stages. First, students discussed their ideas about force and motion as they pertained to two examples: a pen thrown in the air and caught, and a book pushed across a table at constant speed. The goal of this stage was to introduce the interview topic and assess students' prior knowledge. Second, one of the multimedia treatments was shown. Third, students completed a questionnaire including three physics questions selected from the post-test and a few questions about the multimedia itself. Lastly, students discussed their answers with each other.

The aim of the interviews was to help understand the quantitative findings by answering several questions:

- Could students recall accurately the ideas presented in the multimedia?
- Did the inclusion of alternative conceptions cause students to engage differently with the instruction?
- What were students' perceptions of the alternative conceptions presented in the Refutation and Dialogue?

These questions are addressed separately in the sections below. Individual students are identified by the multimedia treatment they watched with the letters E (Exposition), W (Worked Examples), R (Refutation), and D (Dialogue), and a number to differentiate students from the same interview. The letter 'I' is used for the interviewer.

In the first stage of the interviews, no students were able to answer questions about Newtonian mechanics correctly. All either expressed common alternative conceptions or a general confusion about key physics terms.

10.7.1 Accuracy of recall

Students in all focus groups had difficulty accurately recalling information directly presented in the multimedia. In fact, in all of the interviews combined, only a few questions were answered correctly following the multimedia on the questionnaire. Often, students used technical terms loosely when describing what they watched, suggesting that a misunderstanding of terms may have contributed to their inability to recall what was said. Not surprisingly, many students in the no-misconceptions treatments believed that the multimedia supported the views they held at the start of the interview.

W2 In the video it said as the ball leaves your hand there is a decreasing amount of energy acting upon the ball, therefore it stops at the origin and then it comes down and it comes down because of the force of the earth. So I know that B and C are really similar but I chose B because it says 'a steadily decreasing upward force.' Because the ball is slowly decreasing in force so therefore it stops at one point and comes down.

Although energy was not discussed at all in the multimedia, this student had a strong inclination that something was decreasing as the ball travelled upwards. The way the word energy is used, as in "energy acting upon the ball," suggests that the student is thinking more in terms of force than energy, and in fact she equates these two when selecting her answer. This is despite the fact that this exact scenario was addressed in the multimedia. Along with an animation showing one arrow on a juggling ball, the narration explained: "only one force acts on the ball throughout its flight. This is the force of gravity which is constant and downward." Furthermore, this reasoning was repeated when it was used to calculate the acceleration of the ball at all times while in the air.

In the Exposition interview, three students agreed that the force by an elevator cable must be greater than the force of gravity to keep the elevator ascending at constant velocity.

E1 Because it's moving up.

E3 It wouldn't move if they were both equal would it?

E2 There's also the book example like she said the book would not move because it's got the force pushing back this way [friction]. The force of you pushing has to be greater than the force pushing it back in order for it to move – so if you made that from horizontal to vertical it would be an elevator.

The last student (E2) perceptively related what he saw in the multimedia to the elevator question in a good example of transfer. Unfortunately, he did not recall accurately the idea that the forces on the book must be balanced, and instead retained his intuitive notion that motion requires an unbalanced force. Describing the same book scenario, a student in the refutation group accurately recalled both the misconception and the correct scientific explanation presented in the multimedia.

R1 Well the normal force and gravity were equal and also the friction and the force from her hand were equal. And also, she was saying about the misconception – that was a good point – you would think that you would have to be pushing harder than the friction whereas it turns out that you don't have to.

I Does that make sense?

R1 Sort of. I guess if you're going at the same... If the force is the same magnitude... I don't know.

Even though this student was uncertain about the new conception, he could clearly recall the idea presented in the video. He also recognized that the new idea was different to his previous conception and used the misconception as an important counter example to his intuition. Given the persistence of alternative conceptions and the short duration of the multimedia intervention, this is an impressive result that may form the first step towards conceptual change for this student.

A student in the Dialogue interview also correctly recalled the book example and related it to the elevator question.

I Can you take us through your thought processes for why you picked B [force from the cable = force of gravity]?

D2 So I had A and I thought it was right and then I went down [the list] and the rest were wrong except B, which I wasn't sure about because I thought, hang on, didn't they say it was equal on the book even though it's moving – cause the arrows were the same and it was still moving. And I was thinking 'that doesn't make sense,' and then I had a look at it and I thought of F equals MA and I thought it's not accelerating because it's at a constant speed [points to question] and so if force equals mass times acceleration and acceleration equals zero then force equals zero – so they [the forces] must be equal because they're opposite vectors – cancel each other out. That's what I thought.

From this quote, the student recalled correctly that the forces on the book were equal, but this conflicted with her previous beliefs, resulting in confusion. She resolved this confusion by reasoning effectively with Newton's second law. This exemplifies how conceptual change may be facilitated by alternative conceptions in multimedia. Confusion created by alternative ideas may cause students to consider different lines of reasoning, making possible a cascade of thought processes that precipitates the conceptual change. It should be noted that the diagram of the book with four forces acting on it was present for the same amount of time in each of the multimedia, as was the formula $\vec{F} = m\vec{a}$. However the initial confusion that led to this student challenging her beliefs only occurred because she accurately recalled the scientific conception presented in the multimedia. Students who watched treatments without alternative conceptions did not seem to be similarly confused because they did not often correctly recall the material presented in the multimedia.

It would be incorrect, however, to characterize recall from the multimedia as entirely accurate for the misconception groups and inaccurate for the others. There were examples of inaccurate recall in the misconception groups and at least one conceptual changing recall in a no-misconception group. In the Dialogue interview, one student recalled the intuitive impetus theory for the juggling ball thrown in the air:

D2 Because gravity's always going to be exactly the same, but the steadily decreasing, like the force up wears out and so it's decreasing and then when it becomes equal it stops and then when it's less, it's still decreasing but gravity's more so it pulls down.

It is unclear whether this student is remembering the misconception discussed by the model student in the Dialogue, or her own preconception. Regardless, she did not recognize the disparity between her ideas and the scientifically accurate conceptions presented in the multimedia.

In contrast, a student who watched the Worked Examples treatment correctly recalled the forces on the book.

W1 Well at the start, doesn't the force of the hand have to be greater than the frictional force to get it to initially start to move? To overcome the frictional force at the beginning to initially start it moving. And then, I don't know, are they equal?

I What did she say in the video? Do you remember?

W1 I thought that's what she said in the video – that they have to be equal. If the force of her hand were greater than the force of friction, it would speed up.

The examples above suggest that alternative conceptions in multimedia are not always necessary to promote conceptual change nor are they always effective. However the general trends in the interviews, and the quantitative data from hundreds of students suggest that multimedia treatments that include alternative conceptions are more often accurately recalled than treatments without alternative conceptions.

10.7.2 Engagement with the multimedia

Students in all focus groups reported enjoying the multimedia they watched and learning from it. The graphs, diagrams, and animations were selected as key features that students felt helped them learn.

In the no-misconception multimedia interviews, students generally expressed the view that the multimedia was clear and simple. Some said that they already knew the information presented in the multimedia and therefore didn't have to focus on it as intently.

W2 Yeah, definitely, very simple explanations. So it makes physics look really simple. And it was also very clear and concise – to the point, didn't go around in circles.

W1 It wasn't that hard to pay attention to, I think – because I knew already what she was talking about. So I was listening, but I wasn't really paying utmost attention.

W2 Newton's first law I knew already. I guess it was revision from two years ago.

These views help understand why the self-reported mental effort scores were lower for the no-misconception groups. Students felt they had seen the material before. The subject matter appeared simple and straightforward since the speaker explained it methodically. Therefore, students invested less effort in understanding deeply what was presented and checking this against their existing ideas.

Students in the misconception treatment interviews were much less likely to view the multimedia as simple. Across all of the interviews, students used the words 'simple' and 'clear' seven times each and 'concise' four times; all instances occurred in the no-misconception interviews. The phrase 'easy to understand' was used twice, once in the Exposition interview, and once in the Worked Examples interview. In contrast, the word 'confused' was used five times, all by misconception viewers. The phrase 'hard to understand' was used once in the Refutation interview. Since students who watched the misconception treatments realized that the presented ideas differed from their own, they were more likely to invest mental effort during the instruction.

10.7.3 Perceptions of alternative conceptions

Even though alternative conceptions may have confused them, students who watched the Dialogue or Refutation believed they benefited from this aspect of the multimedia.

R1	Saying all the common misconceptions – that was really helpful. So you know what it is but you also know what it's not. So you can know that if you end up with that, you're like 'no, can't be that.' So that helped a lot.
I	Did you have anything like that – where you thought, it's like this but she said it wasn't?
R1	Yeah, the juggling ball one – the misconception is there's a force and that it's slowly decreasing until it reaches the top and then it disappears or whatever. That was the misconception I think. It was just good, and with the car one . . . as it went up the hill, it was actually the velocity of the car that pushed it up the hill but it was always gravity acting upon it and the force from the hill.

An additional benefit of the Dialogue was that it involved a character to whom students could relate.

D2	I liked that the guy was just as confused as I was [laughs]– to begin with. The fact that he was confused kind of helped the whole explanation process, in me [points to head] to understand.

This quote highlights two benefits of the dialogue approach. First, this student relates to the emotions of the student in the Dialogue, something observed in the interviews following the quantum tunneling multimedia. Schunk & Hanson (1985) found that students who watched a peer demonstrate a particular mathematics skill displayed higher self-efficacy and performance than those who did not. A later study showed that students rated themselves more similar to a peer who had difficulty demonstrating the skill than one who performed the skill with ease (Schunk et al. 1987). The second benefit this student eludes to is that the confusion helped provide context for the explanations. This is in line with Sweller's (2004) borrowing principle and Vygotsky's (1978) ideas about vicarious learning. The schemas verbalized by the two dialogue participants served as guides for the organization of novel information in the learner. The schema of the student in the Dialogue was particularly helpful because it was similar to the learner's schema, meaning instruction was taking place within the zone of proximal development (p.135). Under the right circumstances, the reasoning in the Dialogue observed by the learner could be internalised and applied in future situations.

10.8 Discussion

Students from the Fundamentals and Regular classes who watched multimedia that included alternative conceptions achieved higher post-test scores than those who viewed multimedia that strictly followed established design guidelines. These results help extend the findings of refutation text studies (Guzzetti et al. 1993, 1997, Diakidoy et al. 2003) to the area of linear multimedia. Linear multimedia is different to refutation text because it is transient in time. Learners cannot easily refer back and forth between scientific and alternative conceptions as they can with refutation text. Instead, they must construct their understanding as the multimedia progresses. In spite of this, students on average were able to achieve higher post-test scores with misconception-based multimedia than with concise expository treatments. This study is in good agreement results in previous chapters, confirming that a measure of conceptual change can occur with learning resources that some would consider 'passive.'

The novel finding reported here is that the inclusion of alternative conceptions measurably increased the cognitive load on learners during instruction. Furthermore, it appears that this increase in cognitive load was germane since students who watched misconception treatments achieved greater post-test scores than those who saw expository treatments. Follow-up interviews support this conclusion. Students reported that more traditional instructional approaches demanded less attention because they appeared simple and clear. Consequently, students did not necessarily consider how the presented information fitted with their prior knowledge. When asked about the contents of the multimedia they watched, they were more likely to say that their preconceptions were presented than those who watched a misconception treatment. It is remarkable that small changes in the content and presentation of a short multimedia treatment can have a pronounced effect on the way in which the presentation is viewed, and the learning that results from it.

Some might question whether the methodology allowed some students to consult with other resources or peers before or during the post-test, or to not pay attention to the multimedia, thereby undermining the results. These possibilities cannot be ruled out since students were allowed to participate in their own time as they saw fit. However, it is incredibly unlikely that Refutation and Dialogue students, all randomly assigned, would behave in such a way as to artificially inflate their scores. More likely is the possibility that unanticipated student behaviours resulted in an underestimation of the benefits of the Dialogue and Refutation. Regardless, the results of this study should be applicable to real learning settings with learners enrolled in the subject matter, a central objectives of this work.

The effect sizes observed in this experiment ($.36 - .45$) were smaller than those in the previous study ($.79 - .83$). This may be due to a technical difficulty that arose during this study. Some students emailed complaints that they were not able to complete the pre- or post-test, receiving an error message when they tried to advance to the next screen. The likely reason was their connection to the server was timing out, though the source of this difficulty was not identified. The PHP code was not changed from the 2006 study, however upgrades to server software, firewalls, and new versions of web browsers may have been to blame. Once the problem was known, all error messages were changed to inform students about how they could continue with the study. This required returning to the login screen, logging in and then manually changing the url to the web page where the error first occurred. Because of this difficulty, some students did not persist. Others may have been distracted by the extra instructions.

Qualifying the inclusion of discussions of alternative conceptions as germane cognitive load still begs the question of how learning from only correct expositions is different from learning with the inclusion of alternative conceptions. Although the data do not bear directly on the different learning processes, schema theory appears to offer a satisfying explanatory framework for the observed effect. The kind of learning taking place differs depending on if and what kind of prior knowledge is activated. For the case of no relevant prior schema being activated, offering learners a discussion of alternative conceptions allows for schema induction (analogous to well understood concept learning mechanisms), with counter-examples being an important resource for induction (Langley & Simon 1995). For the case of a schema being activated, but one that is not entirely in accordance with the correct explanation offered in the presentation, the schema representing prior knowledge can be structurally modified and/or the belief strength can be reduced (Holland, Holyoak, Nisbett & Thagard 1986). The Refutation and Dialogue provide important information for which part of an existing schema to modify (or to replace it completely). Finally, for the case where an existing schema is activated and it is identical in content to the correct explanation being offered in the presentation, that schema can be seen as being 'strengthened' and the probability that it will be used again by the learner is hence increased.

In comparing the two misconception-based multimedia treatments, the Refutation and Dialogue seem to be equally effective at promoting conceptual change. Even with high prior knowledge learners, the key variable influencing learning appears to be the inclusion of alternative conceptions. However, interviews suggest that in addition to remedying alternative conceptions, the Dialogue treatment may provide affective benefits. Students related to the model student in the Dialogue. They appreciated seeing their ideas presented in the multimedia and felt less alone in their confusion as a result.

The most surprising result in this study was that the pre-test appeared to have an effect on post-test scores, and a different effect on students who watched the Dialogue and the Exposition. It was expected that the pre-test would have little if any effect. Taken at face value, the results suggest that the pre-test helped Dialogue students to pay attention to the important parts of the multimedia while distracting Exposition students from key information. It is unclear how the pre-test might have had these dramatically different effects on two similar populations of students.

Students may have played a role in creating this effect by logging in repeatedly and, by doing so, choosing their experimental condition. Although they were asked to participate in the experiment individually, it was clear that some students discussed it with their friends. Some students found out that the pre-test was optional. They emailed to complain that after viewing the multimedia they were asked to do the test again. Since their friends only completed the test once, they didn't understand why they should do it a second time. Some participants were observed to login repeatedly especially if they received the pre-test the first time. Once they were assigned directly to a multimedia treatment, they appeared to complete the experiment in earnest. The effect of the pre-test is important and should be investigated in further studies.

Science education researchers might doubt the extent to which conceptual change can be facilitated by multimedia. When considering this concern, it is important to keep in mind the range of experiences students have, and the different amounts of mental effort they invested (Cook 2006). For some students, misconception multimedia certainly failed to make any impact on their conceptions. However for others, the evidence suggests that some degree of conceptual change was achieved. After all, conceptual change is not regarded as an all-or-nothing process but rather as a gradual shift over time (diSessa 2006). Discussions in interviews indicate that even when a student has taken a step towards conceptual change, it may not be apparent in his or her answers to multiple-choice questions. This is a second way in which the results obtained in this study may be a conservative estimate of the learning that occurred.

It is important to note that the claim is not that online multimedia is the best method for changing students' conceptions, nor should it be a stand-alone solution to conceptual difficulties. I do suggest that a) linear resources can be more effective than interactive simulations in cases where students have little accurate prior knowledge, b) linear resources can be improved by including alternative conceptions, because c) this results in students investing more mental effort that d) means they are more likely to recognize discrepancies between their extant knowledge and correct scientific conceptions. I therefore recommend that misconception-based multimedia be used as a resource in conjunction with interactive teaching methods. Furthermore, it can be worked into simulations to provide scaffolding and help focus on conceptual issues (Rieber et al. 2004).

Chapter 11

Discussion

The series of experiments described in this thesis have identified obstacles to the conceptual learning of physics and potential methods of overcoming them with multimedia. Below I consider the theoretical and practical implications of this work, as well as limitations and potential future research.

11.1 Theoretical implications

It is worthwhile to consider how the theory discussed in this thesis accounts for the learning experiences that took place with the misconception-based and non-misconception-based multimedia.

Learning with expositions

First, consider a Fundamentals student with little formal prior physics knowledge, watching the Exposition treatment on Newtonian mechanics. The technical words used in the multimedia presentation, like force, acceleration, velocity, and mass would activate schemas in the student for these concepts. However, these schemas would be quite unlike a scientist's schemas for the same ideas. The student's schema for velocity, for example, would be a 'nondifferentiated protoconcept' (Trowbridge & McDermott 1980). Althought the student might be able to appropriately define velocity, his or her concept would likely change depending on the context. Observing two objects moving beside each other, position, a more visually salient feature, might be 'confused' with velocity. This would make a coherent interpretation of the presentation very difficult.

Even with appropriate, fairly well-formed schemas, the Exposition presentation would require a high intrinsic cognitive load. Ideas like $\vec{F} = m\vec{a}$ not only require an appreciation of the elements in the formula, but also of the relationships among the elements. In the description of a particular event, for example the flight of a juggling ball through the air, a student's pre-existing understanding of the situation would guide his or her perception of the salient features of the explanation (Osborne & Wittrock 1983). Thus some statements that did not fit with pre-existing conceptions would be processed at a shallow level. An example would be "while the ball is in the air, only one force acts on the ball. This is the force of gravity, which is constant and downwards." Meanwhile, statements in partial or complete accord with the schemas in long-term memory would be much better attended to and possibly extended beyond their realm of validity. For example, "the upward force of the hand is greater than the downward force of gravity, accelerating the ball in the upwards direction," and "as the ball goes up, it travels slower and slower upwards, and its velocity decreases."

If part of the presentation were perceived correctly, even though it conflicted with a student's prior knowledge, it would probably not make a lasting change to the student's long-term memory. Once attention were directed to a new concept, the recently perceived idea would be susceptible to proactive interference. So, if the student tried to access this conception again, he would much more likely activate the older, more robust, alternative conception than the newly perceived scientific idea.

This explains why, following the multimedia presentation, students reported that their preconceptions were presented in the multimedia. Those aspects of the presentation that agreed with their prior knowledge, for example the idea that the force from the hand is greater than the force of gravity, were correctly recalled and even extended beyond the context in which they are valid (in this case, to include the time after the ball has left the juggler's hand to the peak of its flight). With velocity as an nondifferentiated protoconcept, the section describing the decreasing upward velocity could be recalled afterwords as decreasing upward force.

A key feature of this description of learning is that at no time is the learner aware that what he or she is perceiving is at odds with what is being presented. The words are familiar as are the situations discussed. The sentences used to describe the phenomena are not long or convoluted, and the presenter does not struggle to make the concepts plain. And, importantly, the learner believes he already has the general idea. Therefore meticulous attention is not paid to every detail of the multimedia presentation and deep processing is not generally encouraged.

This effect can be viewed as a type of metacognitive impairment. A colleague working with the confidence data and the psychological constructs of over- and under-confidence (see Kleitman & Stankov 2001) described the students in the 2006 Newtonian mechanics study as supremely overconfident. Furthermore, the constructs showed that Advanced students displayed much less over-confidence than the poorer performing Fundamentals students. The general implication for such a finding would be that the Fundamentals require more training in metacognitive strategies and evaluating their knowledge structures.

I believe this view flips the problem on its head, but it is instructive when envisioning the effects of alternative conceptions. The problem is flipped on its head because I do not believe Fundamentals students are inherently less effective at executing metacognitive strategies than Advanced students. Both groups of students entered university with similar high school marks and admissions rankings. If, for example, their ability to judge the extent or correctness of their understandings in a domain area apart from physics were evaluated, I think Fundamentals and Advanced students would perform equally well. Their deficiency lies not in the knowledge evaluation process but in the knowledge with which they are evaluating. It is useful, though, to consider this phenomenon as though it *were* a metacognitive impairment. Students with alternative conceptions view the multimedia in the same way as they would if they had no way of evaluating whether the presented information matched or differed from their prior knowledge.

This explains why students in the Fundamentals and Regular streams reported high levels of confidence on the pre-test, even when answering less than 50% of questions correctly. It also accounts for the similar gains in confidence that were observed across all treatments, regardless of whether much learning occured or not. Because students could not accurately evaluate whether their post-test answers matched with the information presented in the video, their gains in confidence resulted simply from the experience of seeing some instruction. Furthermore, this view helps understand why students who viewed a non-misconception-based multimedia treatment invested less mental effort than their peers who viewed the Dialogue or Refutation. They were not aware that there was any need to invest more effort.

Learning with the Dialogue or Refutation

In contrast, consider how learning occurs when a student watches a multimedia treatment involving common misconceptions. Intrinsic cognitive load is again a concern, however comprehension may be easier to achieve because the presented conceptions may better match a student's conceptions. As the student in the Dialogue describes the two forces on the ball as it travels through the air, the learner feels she understands the explanation. It confirms what she previously believed and therefore is likely to be remembered accurately. Proactive interference should not hinder remembering for most novice students because the prior conception aligns with the presented idea. However, when the tutor in the Dialogue points out inconsistencies in the student's reasoning, the learner is forced to reconsider the explanation she just saw. Both the confusion expressed by the model student and the clearly remembered explanation may cause the learner to pay more attention to the discussion.

The discussion serves three very important functions. First, it alerts students to the fact that their previous ideas may need reconsidering. As observed in interviews, students felt they understood Newton's laws if they could produce textbook definitions. Posner et al. (1982) proposed that the first requirement for conceptual change is a dissatisfaction with existing mental models. However it seems that a prerequisite is an explicit awareness that current conceptions require more thought. It is this zeroeth order requirement that the discussion of misconceptions must achieve. How else will a student ever become dissatisfied with his or her existing models if these models are used to selectively perceive incoming information, and evaluate whether any discrepancies exist?

Second, as proposed by Vygotsky (1978), the interpersonal process of consensus building reveals the ways of thinking appropriate for resolving discrepancies between two different views in the domain. This process is rarely observed by students because most teaching, which takes place in the form of a monologue, does not seriously consider alternative views. In addition, when other views are taken into account, they are not usually presented with supporting arguments. Observing this reasoning process may later give rise to an inner dialogue within the student that functions as an analytical reasoning tool.

The third significant function of the discussion is that it tethers the new conception to the older robust conception. As described by Ausubel (1968), the assimilation process creates an ideational complex consisting of a modified old idea and the new conception. When a student is asked a question about the material, the old misconception schema will be cued, but it will also likely activate the scientifically correct conception. As a student who watched the Dialogue explained in an interview (p.196), she first thought of her preconception, but almost immediately remembered that the multimedia showed something different. Over time, through repeated reinforcement of the correct conception, the misconception can become weakened and dissociated from the scientific idea (Holland et al. 1986).

New theoretical contributions

Although alternative conceptions have been of central importance to science education researchers over the past three decades, the older idea of proactive interference from studies of memory has received little mention in science education studies. A Google Scholar search for “proactive interference” and “science education” yields only thirteen hits[1]. Science educators and learning theorists have developed their own ways of explaining alternative conceptions, with framework theories, p-prims, and ontological categories, but proactive interference seems to make these models redundant. Numerous well-designed studies have documented the effects of proactive interference and how it can be overcome. In a study with particular relevance to this thesis, Kane & Engle (2000) demonstrated that learners with larger memory spans were better able to overcome proactive interference. If, in the studies described in Chapters 8–10, the mental effort invested by students is regarded as a measure of the effective memory span allocated to the learning tasks, then the results can be understood as a practical, natural demonstration of the findings of Kane & Engle. This thesis represents only a preliminary attempt to investigate the applicability of proactive interference research for understanding alternative conceptions in science education. Future studies should draw inspiration from this body of literature and investigate possible parallels with misconception research.

The studies described in this thesis have advanced cognitive load theory in naturalistic settings. Existing methods of measuring cognitive load have been applied online with results consistent with expectations. One question that has been expanded upon is: what constitutes an extraneous cognitive load? Some extraneous information is easy to differentiate from material that is important and relevant to the learning outcomes. For example, Newton's life story, though potentially interesting, has no bearing on one's understanding of his laws of motion. Misconceptions, though certainly not required to understand the correct physics, appear to help in the learning process. Because they are non-essential they are often left out of instruction. Few textbooks include them and few lecturers discuss them in class. With students of all ranges of ability and different prior conceptions, a discussion of misconceptions would seem to involve an extraneous load for most. However the studies reported here have shown that on average the load induced is germane.

This finding also bears on the claim of constructivists that learning must be 'active' rather than 'passive.' When pressed to unpack these loaded terms most concede that it is cognitive rather than physical activity that is required. However in

[1] as of September 7, 2007

most constructivist writings there appears a tacit assumption of what generates useful cognitive activity. Answering questions, hands-on experiments, or discussing with peers, not listening, observing, or watching is thought to promote the cognitive processes required for learning. The latter activities are thought too similar to traditional teaching methods to be effective. This is what gives rise to the constructivist teaching fallacy (Mayer 2004*a*). The results in this thesis clearly demonstrate how it is, in fact, a fallacy. Depending on the methods employed in multimedia, instruction can be viewed in different ways by students, encouraging different levels of learning. Much is made of the difficulty in changing student conceptions, but at least some measure of conceptual change was achieved by these short multimedia messages, as evidenced by interviews and extensive pre- and post-test data.

Cognitive load, as utilized in this study, may be a useful construct when considering methods of science teaching in a range of settings, not just with multimedia. Often the idea of 'heads-on' learning is used to describe the desired result of innovative pedagogy (e.g. Hake 1998), but it is not clear exactly what this type of learning is or how it can be measured (without using academic tests). Germane cognitive load resembles active learning in that it refers to conscious, effortful activity on the part of the learner that results in measurable long-term learning. Cognitive load researchers have shown that this load can be measured in a variety of ways including physiological measures and self-reported rating scales (Paas et al. 2003). Perhaps the established methods in cognitive load research can be used to evaluate teaching practices and learning in other areas of physics. Science educators should consider whether cognitive load offers a useful framework for understanding active learning.

The results of the three design experiment iterations are consistent with the equivalence principle, that media are not inherently beneficial for learning, but they can be made equally effective by the appropriate choice of methods. Unlike numerous other studies in which the 'no significant difference' phenomenon has been regarded as an unfortunate consequence of deficient methods or materials, the equality of all media in this case shows its potential for multimedia learning. For students who are not in Peer Instruction classrooms, or who do not have peers with whom to discuss their ideas, vicarious learning with multimedia provides a substitute for the natural social interactions demonstrated to improve learning. Discussing alternative conceptions in multimedia brings instruction into the zone of proximal development. In effect, it makes established reform methods available to more students and it allows for the investigation of aspects of instruction on learning.

11.2 Practical implications

The most significant implication of this research for practice is that all linear multimedia created in subject areas where misconceptions are common should include discussions of misconceptions. In the domains of Newtonian and quantum mechanics, with nearly one thousand students, participants who watched a misconception-based multimedia treatment performed better on tests of retention and transfer than those who viewed an expository treatment.

Traditional lecture-style presentations, at least in the areas physics investigated in this thesis, seem not only ineffective, but detrimental for student learning. In general, students learned very little from clear, concise, multimedia, especially low-knowledge learners. However, students became more confident in their alternative conceptions as a result of viewing the multimedia. They believed they learned the same amount as students with double their learning gains. Thus the expositions actually strengthened misconceptions.

These findings are significant and widely-applicable given the multimedia explosion now being witnessed on the Internet. The Internet is routinely hailed as a tool that will democratize education, bringing knowledge to those who previously would not have had access. However the format of instruction seems at least as important as the availability of content. Clear, concise, presentations are not necessarily the best to learn from, though this is often what is created when instructional designers are left to use their intuition.

The research findings also call into question pre-readings assigned as part of reform teaching methods like Peer Instruction. Students may gain little by reading traditional expository passages before coming to class, if refutation text results and those of the previous three chapters are any indication. Students likely comprehend little during the reading and may in fact feel that their preconceptions are confirmed. Currently, very few refutation texts are available or used in physics education. These would be essential for pre-reading to be an effective adjunct to in-class discussions.

11.2.1 Understanding reform methods

The view of learning presented in Section 11.1 helps understand existing physics education reform methods in a novel theoretical framework. Below I discuss one reform method and how its advantages can be explained by this framework.

At the University of Washington, *Tutorials in Introductory Physics* were designed to help novice students develop scientifically accurate conceptions of introductory physics topics. Key features of these tutorials are that they a) require students to fill out worksheets, b) break complicated physics ideas down into a sequence of small steps, c) were written in light of voluminous research on students' alternative conceptions, d) involve tutors signing off on the progress of groups at crucial stages, e) get students to work in small groups, and f) involve students in hands-on activties. Why each of these features is effective can be explained by the theory outlined above.

By requiring students to fill out worksheets, the load on working memory is reduced. When asked to draw conclusions at the end of each section, students can reflect on the words they've written rather than their memories. This also reduces the problems of poorly formed schemas and proactive interferece. Because students are asked to write down an operational definition for each term, like velocity, they can (and are encouraged to) refer back to it when answering questions involving the term. Thus students who don't have a readily accessible schema for each term, can rely on their worksheet as a durable external memory store. Proactive interference is less of a difficulty because each answer stored on the page is not susceptible to interference from prior knowledge. Therefore, although students may have alternative conceptions, their written answers can provide scaffolding while they construct new, scientifically accurate knowledge.

By breaking down the physics into a sequence of small steps, the tutorials limit the intrinsic cognitive load of the task. Again this is done by offloading memory functions to the worksheets. By writing and sketching, learner's may also be forced to consider their conceptions explicitly and undertake thought processes germane to the learning task. The creation of external representations is widely acknowledged to aid knowledge creation and reflection (Rivard 1994).

Since the tutorials were written after in depth research into students' alternative conceptions, they focus on important and commonly misunderstood topics. They ask questions that might seem irrelevant to a physics lecturer, but which, from a student's perspective, are of central imporance. In other words, the tutorials are written with the common schemas of students rather than lecturers in mind.

The tutors serve an important role by signing off on students' progress at the completion of each section. They ensure that students' worksheets contain clear operational definitions and are not affected by alternative conceptions.

Group work has many accepted benefits. Due to the social interaction, students may be more engaged in the task than they would be alone. This could be seen as encouraging germane cognitive load. Group members can also serve as scaffolds. All students in a group likely have different prior schemas; negotiating meaning consolidates the understandings for all participants.

Finally, involving students in hands-on activities may encourage germane cognitive load the same way group work does.

11.3 Applications of misconception-based multimedia

Misconception-based multimedia has the potential to be incorporated into a range of learning environments to enhance learning. As shown in this study, it may be sufficient in itself to promote conceptual change. It may be important to use prompting questions before the multimedia and follow-up questions afterwards, a role served in these studies by the pre- and post-tests.

This type of multimedia could also be used in classroom settings either as an introduction, review, or part of a concept testing question. It has the potential to promote question-asking in lectures though this may be hindered by student aversion to speaking up in large classes.

Since simulations are often found to overload learners, linear multimedia could be used as scaffolding in interactive simulations. Rieber et al. (2004) studied how short expository multimedia segments could improve learning with a Newtonian mechanics simulation. Misconception-based multimedia could likely provide even better scaffolding.

Dialogues involving alternative conceptions are not new in instruction, though this is because of their intuitive appeal rather than scientifically demonstrated merit. Already, vicarious learning inspired tools are being used in physics reform methods. The *Tutorials In Introductory Physics* worksheets include scripts of common dialogues involving alternative conceptions between hypothetical characters (McDermott & Shaffer 2001). Students are asked which character they agree with, if either, and why. This leads to reflection on important concepts, which appears to be beneficial for both novice and more experienced learners. Peer Instruction is another method that has improved student performance by involving them in dialogues about conceptual questions (Mazur 1997). The results of this study suggest that it is not just discussing but observing discussions that leads to these impressive conceptual gains. It could be argued that observing should precede engaging in dialogue to set the groundwork for ideas to come and limit faulty effort (Vygotsky 1978, Bandura 1986). A film program has even addressed the issue of modern physics in a dialogue format between a student and a physics lecturer located inside a television screen (BBC 1970). Of course one of the most famous dialogues is that written by Galileo (1954) on the topic of mechanics.

11.4 Limitations and future research

Overall, the design experiment methodology provided a worthwhile framework in which to conduct this research. The cycles of design, evaluation, and re-design allowed for the development of the multimedia interventions and of theory. An unusual aspect of these studies was that theory development was the primary focus rather than the creation of effective classroom interventions.

One limitation of the research conducted in this thesis is that almost all results were collected in naturalistic settings. Although this is important for the design experiment methodology, a balance should be established with more controlled laboratory studies. To better establish the theoretical claims about cognitive load and vicarious learning, more laboratory studies should be undertaken. This would verify that the observed learning benefits weren't a result of different behaviors of students working from home, or self-selection into different treatment conditions.

Studies were conducted with a modest number of participants. To confirm the findings of this thesis, more experiments should be undertaken with a larger sample of students from different backgrounds.

Besides the interview data, the studies reported in this thesis relied heavily on multiple-choice evaluations. This was necessary to ensure that students answered most questions and that differences in responses more likely reflected understanding than a willingness to answer short-answer questions. As I found in a related study of the coherence principle in astronomy (Muller et al. 2007), in authentic online learning environments, learners are unlikely to invest effort in answering short-answer response questions, especially when participation in the study is voluntary.

Fortunately, the multiple-choice questions used were well-researched and validated (Hestenes et al. 1992, Thornton & Sokoloff 1998). The list of distractors for each question was developed from interview data and thus the majority of alternative student conceptions were present. Further research including interviews with students completing these tests have shown that the answers selected on mechanics conceptual inventories do, in a vast majority of cases, accurately reflect students' conceptions (Henderson 2002).

In future studies it would be beneficial to solicit more responses to open-answered questions to better measure learning. These types of questions have been frequently used in other multimedia learning experiments (e.g. Mayer 2001). Furthermore, different types of questions, like two-tiered multiple-choice items, would allow for deeper probing of student conceptions without requiring short-answer responses (Treagust 1987). Two-tiered questions involve a traditional multiple-choice question followed by a multiple-choice set of supporting explanations. Considering the two responses together helps identify random guessing and inconsistencies in reasoning. This technique has been used successfully in a number of other studies (e.g. Kearney & Treagust 2001, Sharma et al. 2005).

It is likely that because the multimedia experiments were performed in real learning settings, they yielded an underestimate of the actual benefit of misconception-based multimedia. Since students were allowed to partake in the experiments at their convenience from home, some probably completed the pre- and post-tests without watching the multimedia. Although the link to continue to the post-test was hidden for the duration of the treatment, some participants may done other tasks while the multimedia played. They would then likely enter the same responses on the post-test as they did on the pre-test. If these participants were part of the Exposition sample, their results would not greatly affect the average gain of the group since most students achieved a gain near zero anyway. However these students in the Dialogue or Refutation samples would significantly reduce the overall average, with students gaining an average of four to five marks on the post-test. Since these non-participants were impossible to differentiate from the 'earnest' sample, they likely narrowed the performance gap between students who watched misconception-based and traditional multimedia.

In future studies it would be important to further explore the role of the pre-test. One of the most unexpected findings was that completing the pre-test appeared to have a positive effect on students who viewed the Dialogue but a negative effect on those who watched the Exposition. Part of the problem may have been an effective self-selection of students into the different treatment conditions when they realized that some involved less work. It would be interesting to replicate the study in a controlled learning laboratory where participants' random assignments to different treatment conditions could not be affected by their unanticipated behaviors.

I believe that it is important to study short linear multimedia interventions as they provide a confined arena in which to investigate significant questions for teaching and learning. However, in order to truly understand meaningful learning, multimedia experiments like those reported here could be expanded along a range of dimensions. Interventions could run for longer durations or involve multiple short sessions. Interactive video can be annotated, which may help students with memory load considerations and proactive interference. More longitudinal data should be collected with students using multimedia throughout a course to establish its long-term implications for learning. Finally, the instruction used in this study consisted solely of linear re-playable presentations. Many more diverse types of multimedia exist with different opportunities for interaction and these should be investigated as research on simpler forms accumulates and theory develops.

Appendix A

Ethics and participant information forms

A.1 Quantum tunneling information sheet, 2005

The University of Sydney

PARTICIPANT INFORMATION SHEET
Research Project

TITLE: The use of video for increasing student understanding and appreciation of quantum tunneling and related phenomena

(1) What is the study about?

This study is about student conceptions of quantum mechanics, specifically quantum tunnelling. Goals of the study are to identify areas of student difficulty with the subject matter and determine how to confront alternative conceptions.

(2) Who is carrying out the study?

The study is being conducted by Derek Muller and it will form the basis for the degree of PhD at The University of Sydney under the supervision of Dr. Manjula Sharma, senior lecturer of physics and head of the Sydney University Physics Education Research group.

(3) What does the study involve?

This part of the study involves watching an instructional video treatment and filling out a short survey on quantum tunneling by intermediate physics students.

(4) How much time will the study take?

The study will take approximately 40 minutes.

(5) Can I withdraw from the study?

Being in this study is completely voluntary - you are not under any obligation to consent.

(6) Will anyone else know the results?

All aspects of the study, including results, will be strictly confidential and only the researchers will have access to information on participants. A report of the study may be submitted for publication, but individual participants will not be identifiable in such a report.

(7) Will the study benefit me?

Completing this survey will give you some insight into what types of questions you may see on the final exam and therefore help you revise. The video contains information about tunneling with a particular instructional strategy. Also, if you are interested in discussing the results of the study, you are invited to do so. The contact information of the researcher is given below.

(8) Can I tell other people about the study?

Yes, by all means. There is no reason to keep this study a secret.

(9) What if I require further information?

When you have read this information, Derek Muller will discuss it with you further and answer any questions you may have. If you would like to know more at any stage, please feel free to contact Derek Muller by email at muller@physics.usyd.edu.au or by phone on (02) 9351 2553. He is also available in his office, room 233A in the Physics Building, A28.

(10) What if I have a complaint or concerns?

Any person with concerns or complaints about the conduct of a research study can contact the Manager, Ethics Administration, University of Sydney on (02) 9351 4811.

This information sheet is for you to keep

A.2 Quantum tunneling consent form, 2005

The University of Sydney

PARTICIPANT CONSENT FORM

I, ...……..............., give consent to my participation in the research project
Name (please print)

TITLE: The use of video for increasing student understanding and appreciation of quantum tunnelling and related phenomena

In giving my consent I acknowledge that:

1. The procedures required for the project and the time involved have been explained to me, and any questions I have about the project have been answered to my satisfaction.
2. I have read the Participant Information Sheet and have been given the opportunity to discuss the information and my involvement in the project with the researcher/s.
3. I understand that I can withdraw from the study at any time, without affecting my relationship with the researcher(s) now or in the future.
4. I understand that my involvement is strictly confidential and no information about me will be used in any way that reveals my identity. My SID may be used to correlate results with my exam performance.

Signed: ....

Name: ....

Date: ....

A.3 Quantum tunneling student information sheet, 2005

The University of Sydney

Information Sheet

SID:______________________

Please circle the appropriate answer.

1. I am enrolled in PHYS: **2012 / 2912**
2. Sex: **M / F**
3. English is my first language: **Yes / No**
4. I plan on doing this many more years of study in physics: **0 / 1 / 2 / 3+**

	Strongly Disagree 1	2	3	Neutral 4	5	6	Strongly Agree 7
5. I am interested in quantum mechanics:	1	2	3	4	5	6	7
6. I feel I understand quantum mechanics:	1	2	3	4	5	6	7
7. I have revised my notes from this class:	1	2	3	4	5	6	7
8. I like learning from videos:	1	2	3	4	5	6	7
9. I like watching documentaries:	1	2	3	4	5	6	7
10. I like it when there are discussions in class:	1	2	3	4	5	6	7
11. I like just being told the right answers:	1	2	3	4	5	6	7

A.4 Newtonian mechanics information sheet, 2006

The University of Sydney

MECHANICS CONCEPTS 2006

To kick off your first year in physics at university, test your mechanics knowledge and learn some cool conceptual physics online. You will receive credit towards your next assignment mark – this does not depend on how well you do, it is just a completion mark. Only you will know your post-quiz results which indicate how well you're going in mechanics early in semester.

By Friday, March 31:

1. **Go to http://www.physics.usyd.edu.au/~muller/mechanics**
2. **Complete the pre-quiz (Password is 'funphysics', without the quotations)**
3. **Watch a mechanics video using Media Player or QuickTime**
4. **Complete the post-quiz**

Please do this individually without consulting physics resources. The whole thing should take 30 to 45 minutes maximum. Email Derek at muller@physics.usyd.edu.au or Jamie at jbewes@physics.usyd.edu.au with any technical concerns.

What it's all about: A research group in the School of Physics has an ongoing project to understand how students learn physics. In 2006, we are conducting an online experiment to explore your ideas about mechanics and to see how you learn best.

What it involves: As part of your first assignment, please go to http://www.physics.usyd.edu.au/~muller/mechanics, **password: funphysics**. Here you will fill out a brief mechanics pre-quiz (no numbers, only concepts), watch a video and fill out a post-quiz. You will be awarded a completion mark for participating in the experiment but your score on the quizzes will be known only to you (it will not be used in your assessment). You will need broadband internet, speakers or head phones, and either Windows Media Player or QuickTime. If you don't have one of these requirements, see *Access labs* below.

The details: The quizzes must be completed *individually* without any help from textbooks or online resources. This will make it easier and quicker for you to complete. Again, how well you do will not affect your mark, we just want to see how you learn. The whole experiment should take 30 to 45 minutes.

Results: Only you will receive your quiz results after completing the test online. All aspects of the study, including results, will be strictly confidential and only the researchers will have access to information on participants. A report of the study may be submitted for publication, but individual participants will not be identifiable in such a report.

Access labs: There are computer labs around campus that are available for your use. If you don't have access to broadband, it will be useful to familiarize yourself with the locations on the back of this page. The closest lab to physics is on level 2 of the Education Building. You'll need head phones to listen to sound (the computers also have a default mute setting you'll have to change).

Being in this study is completely voluntary - you are under no obligation to consent. Instructions for how to receive completion marks without participating are on the website.

The study is being conducted by Derek Muller and James Bewes under the supervision of Dr. Manjula Sharma, senior lecturer of physics and head of the Sydney University Physics Education Research group. If you would like to know more at any stage, please feel free to contact Derek Muller or James Bewes by email at muller@physics.usyd.edu.au or jbewes@physics.usyd.edu.au, or by phone on (02) 9351 2553. They are also available for face to face discussions and can be found in rooms 233A or 426 in the Physics Building, A28.

Any person with concerns or complaints about the conduct of a research study can contact the Manager, Ethics Administration, University of Sydney on (02) 9351 4811.

This information sheet is for you to keep

A.5 Newtonian mechanics information sheet, 2007

The University of Sydney

MECHANICS CONCEPTS 2007

To kick off your first year in physics at university, test your mechanics knowledge and learn some cool conceptual physics online. You will receive credit towards your assignment marks – this does not depend on how well you do, it is just a completion mark. Only you will know your post-quiz results which indicate how well you're going in mechanics early in semester.

By 5 pm, Friday, March 16:

1. **Go to http://www.physics.usyd.edu.au/~muller/advmechanics**
2. **Sign in with course number: 1901 (Password is 'newton', without the quotations)**
3. **Sit a pre-quiz**
4. **Watch a mechanics video using Media Player or QuickTime**
5. **Complete the post-quiz**

Please do this individually and without consulting physics resources. The whole thing should take 30 to 45 minutes maximum. Email Derek at muller@physics.usyd.edu.au with any technical concerns.

What it's all about: A research group in the School of Physics has an ongoing project to understand how students learn physics. In 2007, we are conducting an online experiment to explore your ideas about mechanics and to see how you learn best.

What it involves: As part of your first assignment, please go to http://www.physics.usyd.edu.au/~muller/advmechanics, **password: newton**. Here you will watch a video and fill out a quiz. You will be awarded a completion mark for participating in the experiment but your score on the quizzes will be known only to you (it will not be used in your assessment). You will need broadband internet, speakers or head phones, and either Windows Media Player or QuickTime. If you don't have one of these requirements, see *Access labs* below.

The details: The quizzes must be completed *individually* without any help from textbooks or online resources. This will make it easier and quicker for you to complete. Again, how well you do will not affect your mark, we just want to see how you learn. The whole experiment should take 30 to 45 minutes.

Results: Only you will receive your quiz results after completing the test online. All aspects of the study, including results, will be strictly confidential and only the researchers will have access to information on participants. A report of the study may be submitted for publication, but individual participants will not be identifiable in such a report.

Access labs: There are computer labs around campus that are available for your use. If you don't have access to broadband, it will be useful to familiarize yourself with the locations on the back of this page. The closest lab to physics is on level 2 of the Education Building. You'll need head phones to listen to sound (the computers also have a default mute setting you'll have to change).

Being in this study is completely voluntary - you are under no obligation to consent. Instructions for how to receive completion marks without participating are on the website.

The study is being conducted by Derek Muller under the supervision of Dr. Manjula Sharma, senior lecturer of physics and head of the Sydney University Physics Education Research group. If you would like to know more at any stage, please feel free to contact Derek Muller by email at muller@physics.usyd.edu.au, or by phone on (02) 9351 2553. He is also available for face to face discussions and can be found in rooms 233A in the Physics Building, A28.

Any person with concerns or complaints about the conduct of a research study can contact the Manager, Ethics Administration, University of Sydney on (02) 9351 4811.

This information sheet is for you to keep

A.6 Newtonian mechanics consent form, 2006 & 2007

The University of Sydney

PARTICIPANT CONSENT FORM

I, ..., give consent to my participation in the research project
Name (please print)

TITLE: The use of video for increasing student understanding and appreciation of quantum tunnelling and related phenomena

In giving my consent I acknowledge that:

1. The procedures required for the project and the time involved have been explained to me, and any questions I have about the project have been answered to my satisfaction.
2. I have read the Participant Information Sheet and have been given the opportunity to discuss the information and my involvement in the project with the researcher/s.
3. I understand that I can withdraw from the study at any time, without affecting my relationship with the researcher(s) now or in the future.
4. I understand that my involvement is strictly confidential and no information about me will be used in any way that reveals my identity. My SID may be used to correlate results with my exam performance.

SID ________________________

Password ___________________

By completing this form and continuing, I give my consent to be involved with this study concordant with the terms outlined above.

Appendix B

Physics questionnaires and tests

B.1 *Falling Cats* questionnaire, 2004

Dr. Karl: Falling Cats Survey

The University of Sydney

Please circle the appropriate choice

My current year of Study: High School / 1^{st} Year / 2^{nd} Year / 3^{rd} Year / Hon / Mast / PhD

	Very Low		Moderate		Very High	
I think my interest in Physics is:	1	2	3	4	5	Sex: M/F

Have you seen this video before? Yes/No   Age:_____________

The highest level or course in physics I have completed is:_________________________

1. What was your overall impression of the video?
-
-
-

2. What did you like the most about the video?
-
-
-

3. What did you enjoy the least?
-
-
-

4. List some things the video showed that you already knew.
-
-
-

5. List some things you learned, or better understood, from viewing the video.
-
-
-

6. What techniques employed in the video helped you learn or understand better?
-
-
-

In your own words, describe what is meant by “terminal velocity.” Use diagrams if you like.

B.2 Quantum tunneling questionnaire, 2004

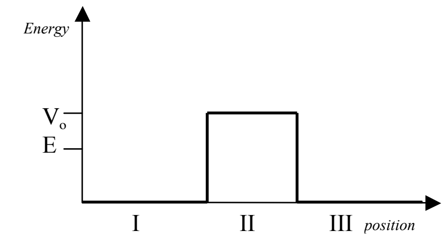
The University of Sydney
Quantum Physics Review Survey

SID*:______________
Please circle the appropriate answer.
I am enrolled in PHYS: 2002 / 2902

	Not at all		Some		A lot	
So far, for this course I have studied:	1	2	3	4	5	Sex: M/F

**Optional – no part of this survey will be used in your evaluation, all answers are confidential*

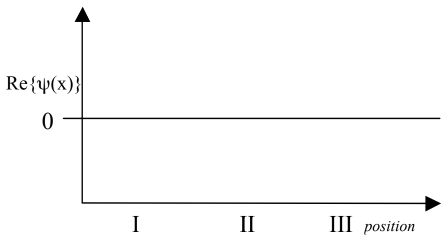
Consider a beam of mono-energetic electrons, with energy E, incident from the left on the step potential shown below.



a) Do the electrons in the different regions have different kinetic energies? If so, rank the energies (E_I, E_{II}, E_{III}) in order from highest to lowest.

____

____

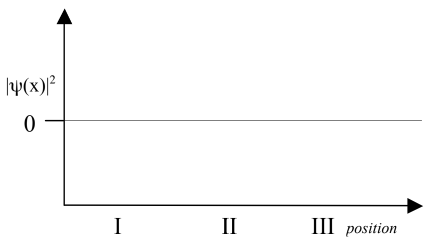


b) Make a detailed sketch of the real part of the wave function in all three regions. Comment on wavelength and amplitude.

____

____

____



c) Make a detailed sketch of the probability density $|\psi(x)|^2$

d) If the barrier were doubled in **height**, what would happen to the transmission probability and the energy of the transmitted electrons?

____

e) If the barrier were doubled in **width**, what would happen to the transmission probability and the energy of the transmitted electrons?

____

f) Some radioactive elements with widely varying atomic masses emit alpha particles with similar energies but with vastly different decay constants. How would you explain this phenomenon?

____

____

____

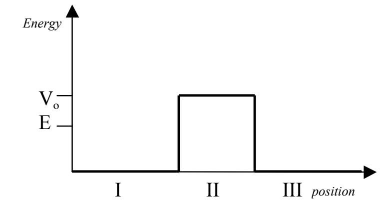
____

B.3 Quantum tunneling pre- and post-test, 2005

The University of Sydney

Quantum Physics Survey

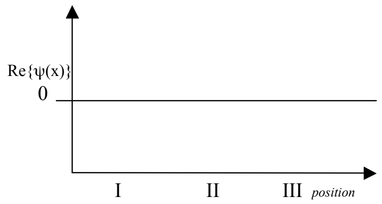
1. Consider a beam of mono-energetic electrons, with energy E, incident from the left on the step potential shown below.



a) Do the electrons in the different regions have different kinetic energies? If so, rank the energies (E_I, E_{II}, E_{III}) in order from highest to lowest.

____

____

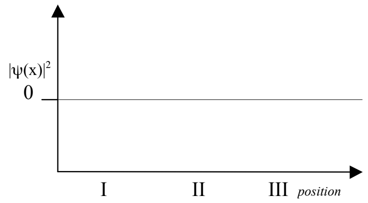


b) Make a detailed sketch of the real part of the wave function in all three regions. Comment on wavelength and amplitude.

____

____

____



c) Make a detailed sketch of the probability density $|\psi(x)|^2$

d) If the barrier were doubled in **height**, what would happen to the transmission probability? What would happen to the energy of the transmitted electrons?

____

e) If the barrier were doubled in **width**, what would happen to the transmission probability? What would happen to the energy of the transmitted electrons?

____

f) Define the amplitude of the wave function.

____

g) Is the slope of the amplitude of the wave function always continuous? Why or why not?

____

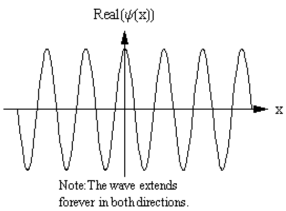
2. Please circle the correct answer

a) In the absence of external forces, electrons move along sinusoidal paths: True / False

b) You see an electron and a neutron moving by you at the same speed. How do their wavelengths λ compare?

A.	λ neutron > λ electron	C.	λ neutron = λ electron
B.	λ neutron < λ electron	D.	I have no idea

c) A particle with the spatial wave function $\psi(x) = e^{ikx}$ can be thought of as a plane wave traveling along the x-axis. Its real part is a cosine wave, as shown in the figure at right. Which of the following statements most accurately describes the probability of finding the particle at any location along the x-axis?



A. It is equally likely to find the particle anywhere along the x-axis.
B. It is most likely to be found in the peaks of the wave.
C. It is most likely to be found in the peaks or the troughs of the wave.
D. The particle is actually located in one particular place, independent of the wave function, and that is the only place you can find it.
E. I have no idea how to answer this question.

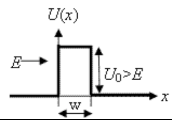
d) Some radioactive elements with widely varying atomic masses emit alpha particles with similar energies but with vastly different decay constants. How would you explain this phenomenon?

____

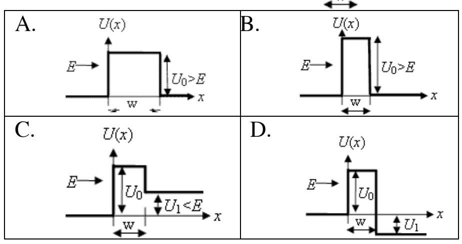
____

____

e) Suppose that protons are incident on a potential barrier as shown to the right. You would like to decrease the speed of the protons coming out on the right side. Which of the following changes to the experimental set-up would decrease this speed?



A. Increase the width w of the gap:
B. Increase U_o, the potential energy of the gap:
C. Increase the potential energy to the right of the gap:
D. Decrease the potential energy to the right of the gap:
E. More than one of the changes above would decrease the speed of the electron.



B.4 Quantum tunneling multimedia opinion form, 2005

The University of Sydney

SID: ____________________

Please provide us with some feedback about the video

	Strongly Disagree			Neutral			Strongly Agree
1. I learned something from the video:	1	2	3	4	5	6	7
2. I stopped concentrating after a while:	1	2	3	4	5	6	7
3. I enjoyed watching the video:	1	2	3	4	5	6	7
4. I knew everything that the video showed:	1	2	3	4	5	6	7
5. The video kept my attention:	1	2	3	4	5	6	7
6. I think this is a good learning tool:	1	2	3	4	5	6	7
7. I found the video dull:	1	2	3	4	5	6	7
8. I could follow the explanations:	1	2	3	4	5	6	7
9. The video was no fun for me:	1	2	3	4	5	6	7
10. I couldn't focus on the information:	1	2	3	4	5	6	7
11. I'd enjoy seeing stuff like this in lectures:	1	2	3	4	5	6	7
12. I found the video easy to follow:	1	2	3	4	5	6	7
13. I got something out of the video:	1	2	3	4	5	6	7
14. The video was too long:	1	2	3	4	5	6	7
15. I want to know more after seeing that:	1	2	3	4	5	6	7

16. If you have any questions about material in the video (things that were not clear), please list them below:

-
-

17. If the video triggered some questions for you (that you'd like to know more about), please list them below:

-
-

Any other comments you have are appreciated:

B.5 Quantum tunneling multimedia opinion form results, 2005

Treatment	Dialogue		Exposition	
	M	SD	M	SD
I learned something from the video:	5.7	0.7	5.6	1.3
I stopped concentrating after a while:	3.3	1.2	3.3	1.8
I enjoyed watching the video:	4.9	1.1	5.2	1.1
I knew everything that the video showed:	2.9	1.1	2.7	1.3
The video kept my attention:	4.8	0.9	5.1	1.1
I think this is a good learning tool:	5.1	0.9	5.3	1.2
I found the video dull:	3.4^a	1.1	2.6^a	1.3
I could follow the explanations:	5.4	1.0	5.6	1.2
The video was no fun for me:	3.1	1.0	2.5	1.3
I couldn't focus on the information:	2.9	0.9	2.8	1.3
I'd enjoy seeing stuff like this in lectures:	4.8^b	1.3	5.5^b	1.1
I found the video easy to follow:	5.2	1.0	5.6	0.8
I got something out of the video:	5.5	0.8	5.7	1.1
The video was too long:	3.2^a	1.2	2.4^a	1.0
I want to know more after seeing that:	4.9	1.1	5.1	1.1

[a] Differences significant at the $p < .01$ level.

[b] Differences significant at the $p < .05$ level.

Table B.1: Results from the quantum tunneling multimedia opinion form.

B.6 Quantum tunneling multimedia interview worksheet, 2005

Quantum Tunneling Video

The University of Sydney

SID:________________

1. What was your overall impression of the video?
-
-
-

2. What did you like the most about the video?
-
-
-

3. What did you enjoy the least?
-
-
-

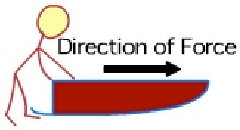
4. What techniques employed in the video helped you learn or understand better?
-
-
-

5. What resources do you find most helpful for learning and why?
-
-
-

B.7 Newtonian mechanics pre- and post-test, 2006 & 2007

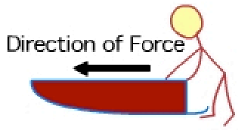
A sled on ice moves in the ways described in questions 1-5 below. Friction is so small that it can be ignored. A person wearing spiked shoes standing on the ice can apply a force to the sled and push it along the ice. Choose the one force (A through G) which would keep the sled moving as described in each statement below.
You may use a choice more than once or not at all but choose only one answer for each blank. If you think that none is correct, answer choice J.

A. The force is toward the **right** and is **increasing** in strength (magnitude).
B. The force is toward the **right** and is of **constant** strength (magnitude).
C. The force is toward the **right** and is **decreasing** in strength (magnitude).



D. No applied force is needed.

E. The force is toward the **left** and is **increasing** in strength (magnitude).
F. The force is toward the **left** and is of **constant** strength (magnitude).
G. The force is toward the **left** and is **decreasing** in strength (magnitude).



	A	B	C	D	E	F	G	J
1. Which force would keep the sled moving toward the right and speeding up at a steady rate (constant acceleration)?	○	○	○	○	○	○	○	◉
2. Which force would keep the sled moving toward the right at a steady (constant) velocity?	○	○	○	○	○	○	○	◉
3. The sled is moving toward the right. Which force would slow it down at a steady rate (constant acceleration)?	○	○	○	○	○	○	○	◉
4. Which force would keep the sled moving toward the left and speeding up at a steady rate (constant acceleration)?	○	○	○	○	○	○	○	◉
5. The sled is moving toward the left. Which force would slow it down at a steady rate (constant acceleration)?	○	○	○	○	○	○	○	◉

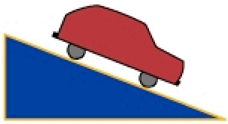
Please indicate how confident you are about your answers.

uncertain						certain
1	2	3	4	5	6	7
○	○	○	○	○	○	○

Submit Answers

Page 1 of 10

Questions 6-8 refer to a toy car which is given a quick push so that it rolls up an inclined ramp. After it is released, it rolls up, reaches its highest point and rolls back down again. *Friction is so small it can be ignored.*



Using one of the following choices (A through G) indicate the net force acting on the car for each of the cases described below. Answer choice J is you think that none is correct.

A. Net **constant** force down ramp
B. Net **increasing** force down ramp
C. Net **decreasing** force down ramp
D. Net force zero
E. Net **constant** force up ramp
F. Net **increasing** force up ramp
G. Net **decreasing** force up ramp

	A	B	C	D	E	F	G	J
6. The car is moving up the ramp after it is released.	○	○	○	○	○	○	○	●
7. The car is at its highest point.	○	○	○	○	○	○	○	●
8. The car is moving down the ramp.	○	○	○	○	○	○	○	●

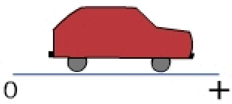
Please indicate how confident you are about your answers.

uncertain						certain
1	2	3	4	5	6	7
○	○	○	○	○	○	○

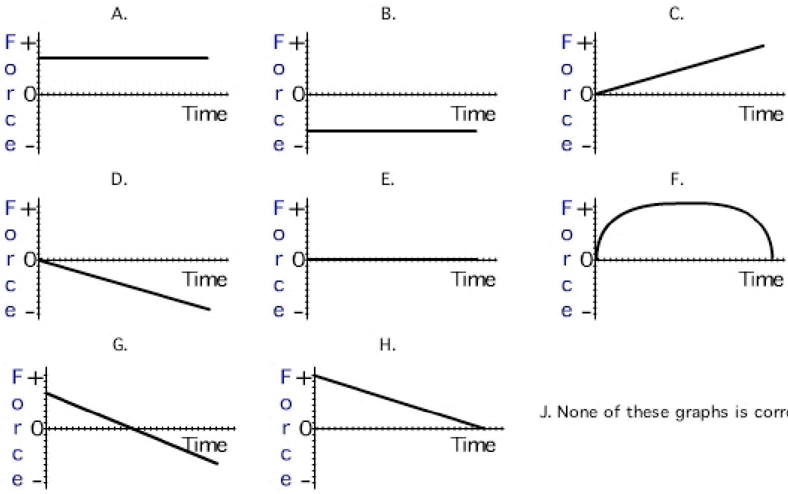
Submit Answers

Page 2 of 10

Questions 9-13 refer to a toy car which can move to the right or left along a horizontal line (the positive portion of the distance axis). Assume that friction is so small that it can be ignored.



A force is applied to the car. Choose the <u>one</u> force graph (A through H) for each statement below which could allow the described motion of the car to continue. You may use a choice more than once or not at all. If you think that none is correct, answer choice J.



J. None of these graphs is correct.

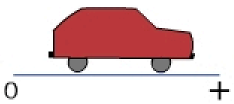
	A	B	C	D	E	F	G	H	J
9. The car moves toward the right (away from the origin) with a steady (constant) velocity.	○	○	○	○	○	○	○	○	●
10. The car moves toward the right and is slowing down at a steady rate (constant acceleration).	○	○	○	○	○	○	○	○	●
11. The car moves toward the left and is speeding up at a steady rate (constant acceleration).	○	○	○	○	○	○	○	○	●
12. The car moves toward the right, speeds up and then slows down.	○	○	○	○	○	○	○	○	●
13. The car was pushed toward the right and then released. Which graph describes the force after the car is released?	○	○	○	○	○	○	○	○	●

Please indicate how confident you are about your answers.

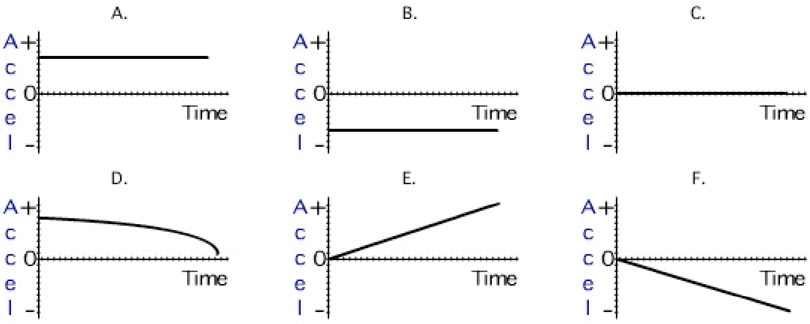
uncertain — certain

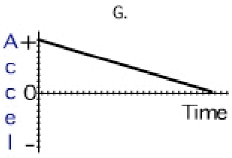
1 2 3 4 5 6 7

Questions 14-16 refer to a toy car which can move to the right or left on a horizontal surface along a straight line (the positive portion of the distance axis). The positive direction is to the right.



Different motions of the car are described below. Choose the letter (A to G) of the acceleration-time graph which corresponds to the motion of the car described in each statement.





J. None of these graphs is correct.

	A	B	C	D	E	F	G	J
14. The car moves toward the right, slowing down at a steady rate.	○	○	○	○	○	○	○	●
15. The car moves toward the left (toward the origin) at a constant velocity.	○	○	○	○	○	○	○	●
16. The car moves toward the left, speeding up at a steady rate.	○	○	○	○	○	○	○	●

Please indicate how confident you are about your answers.

uncertain						certain
1	2	3	4	5	6	7
○	○	○	○	○	○	○

Submit Answers

Page 4 of 10

Questions 17-19 refer to a coin which is tossed straight up into the air. After it is released it moves upward, reaches its highest point and falls back down again. Use one of the following choices (A through G) to indicate the acceleration of the coin during each of the three stages of the coin's motion described below. Take **up** to be the **positive** direction. Answer choice J if you think that none is correct.

A. The acceleration is in the negative direction and constant.
B. The acceleration is in the negative direction and increasing.
C. The acceleration is in the negative direction and decreasing.
D. The acceleration is zero.
E. The acceleration is in the positive direction and constant.
F. The acceleration is in the positive direction and increasing.
G. The acceleration is in the positive direction and decreasing.

	A	B	C	D	E	F	G	J
17. The coin is moving upward after it is released.	○	○	○	○	○	○	○	◉
18. The coin is at its highest point.	○	○	○	○	○	○	○	◉
19. The coin is moving downward.	○	○	○	○	○	○	○	◉

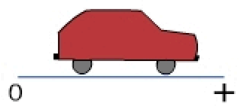
Please indicate how confident you are about your answers.

uncertain						certain
1	2	3	4	5	6	7
○	○	○	○	○	○	○

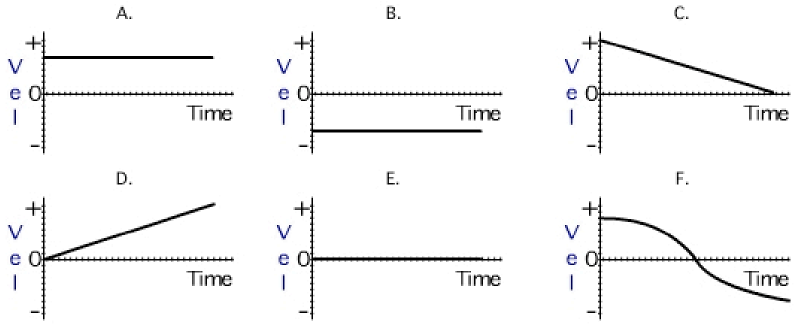
Submit Answers

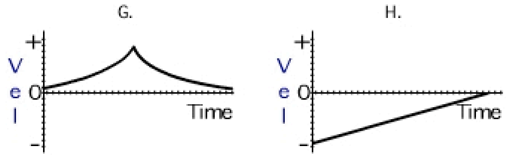
Page 5 of 10

Questions 20-22 refer to a toy car which can move to the right or left on a horizontal surface along a straight line (the positive portion of the distance axis). The positive direction is to the right.



Choose the correct **velocity-time** graph (A to G) for each of the following questions. You may use a choice more than once or not at all. If you think that none is correct, answer choice J.





J. None of these graphs is correct.

	A	B	C	D	E	F	G	H	J
20. Which velocity graph shows the car reversing direction?	○	○	○	○	○	○	○	○	●
21. Which velocity graph shows the car moving toward the left (toward the origin) at a steady (constant) velocity?	○	○	○	○	○	○	○	○	●
22. Which velocity graph shows the car increasing its *speed* at a steady (constant) rate?	○	○	○	○	○	○	○	○	●

Please indicate how confident you are about your answers.

uncertain — certain

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Submit Answers

Page 6 of 10

A boy throws a steel ball straight up. Consider the motion of the ball only after it has left the boy's hand but before it touches the ground, and assume that forces exerted by the air are negligible. For these conditions, the force(s) acting on the ball is (are):

A. A downward force of gravity along with a steadily decreasing upward force.

B. A steadily decreasing upward force from the moment it leaves the boy's hand until it reaches its highest point; on the way down there is a steadily increasing downward force of gravity as the object gets closer to the earth.

C. An almost constant downward force of gravity along with an upward force that steadily decreases until the ball reaches its highest point; on the way down there is only a constant downward force of gravity.

D. An almost constant downward force of gravity only.

E. None of the above, the ball falls back to ground because of its natural tendency to rest on the surface of the earth.

	A	B	C	D	E
23.	○	○	○	○	◉

Please indicate how confident you are about your answer.

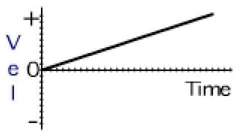
uncertain						certain
1	2	3	4	5	6	7
○	○	○	○	○	○	○

Submit Answers

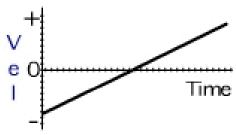
Page 7 of 10

A gymnast is jumping on a trampoline. *While she is in contact with the trampoline* a velocity-time graph for her motion looks most like (taking up to be positive):

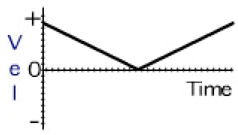
A.



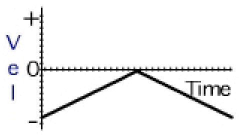
B.



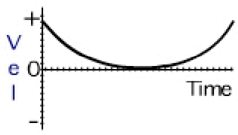
C.



D.



E.



F.
None of these graphs is correct.

	A	B	C	D	E	F
24.	○	○	○	○	○	◉

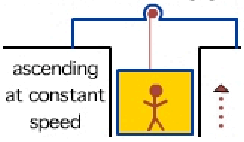
Please indicate how confident you are about your answer.

uncertain						certain
1	2	3	4	5	6	7
○	○	○	○	○	○	○

Submit Answers

Page 8 of 10

An elevator is being lifted up an elevator shaft at a constant speed by a steel cable as shown in the figure below. All frictional effects are negligible. In this situation, forces on the elevator are such that:



A. The upward force by the cable is greater than the downward force of gravity.
B. The upward force by the cable is equal to the downward force of gravity.
C. The upward force by the cable is smaller than the downward force of gravity.
D. The upward force by the cable is greater than the sum of the downward force of gravity and a downward force due to the air.
E. None of the above. (The elevator goes up because the cable is being shortened, not because an upward force is exerted on the elevator by the cable).

	A	B	C	D	E
25.					

Please indicate how confident you are about your answer.

uncertain						certain
1	2	3	4	5	6	7

Submit Answers

Page 9 of 10

A woman exerts a constant horizontal force on a large box. As a result, the box moves across a horizontal floor at a constant speed "v_0".
The constant horizontal force applied by the woman:

A. has the same magnitude as the weight of the box.
B. is greater than the weight of the box.
C. has the same magnitude as the total force which resists the motion of the box.
D. is greater than the total force which resists the motion of the box.
E. is greater than either the weight of the box or the total force which resists its motion.

	A	B	C	D	E
26.					

Please indicate how confident you are about your answer.

uncertain						certain
1	2	3	4	5	6	7

Submit Answers

Page 10 of 10

B.8 Newtonian mechanics focus group questionnaire, 2007

The University of Sydney

Mechanics Concepts 2007

SID:______________________

While I was watching the multimedia, I invested:

extremely low mental effort	very low mental effort	low mental effort	rather low mental effort	neither low nor high mental effort	rather high mental effort	high mental effort	very high mental effort	extremely high mental effort
○	○	○	○	○	○	○	○	○

1. What did you like the most about the video?

-
-
-

2. What did you enjoy the least?

-
-
-

3. Please list some things the video showed that you already knew.

-
-
-

4. List some things you learned, or better understood, from viewing the video.

-
-
-

5. What techniques employed in the video helped you learn or understand better?

-
-
-

Appendix C

Scripts

C.1 Quantum tunneling

C.1.1 Exposition script

Quantum tunneling is one of the most intriguing phenomena of quantum mechanics. It's when particles are able to pass through a potential barrier without the energy classically required to overcome it.

I wish I could draw you a picture of what the whole process looks like. But unfortunately, that's impossible since the quantum world can't be adequately described by our everyday experiences. So, the best thing we can do is look at the mathematical representations of tunneling as they contain all the information available about the system. Hopefully by taking you through the different representations of tunneling, you can solidify your understandings of the process and quantum mechanics in general.

The first thing to look at is the energy landscape. On a potential energy diagram, the quantum tunneling scenario is often represented as a square bump. The bump has some height, which we'll call V_0. Particles coming from the left have energy, E, which is less than V_0.

The gradient of a potential energy diagram represents the force in the opposite direction. The vertical lines do not represent infinite force, they are simply idealizations of a steep but finite slope.

To make this all a bit clearer, we can make an analogy with a classical potential barrier, namely a hill. The slope determines the force experienced by an object on the hill. When an object like a tennis ball is on top of the hill, the system has gravitational potential energy mgh. This is like V_0. If you want to roll a ball up the hill and down the other side, you need to give it enough kinetic energy to overcome this potential barrier. The initial energy E has to be greater than mgh.

With quantum mechanics, however, the particle doesn't need to have the energy V_0 to make it to the other side of the barrier. The reason for this is that in quantum mechanics, we can no longer think of particles as objects with definite size, position, momentum or energy.

The wave function of a particle is continuous and therefore there is a finite probability that it can exist inside and on the far side of the barrier.

But if a particle doesn't need to have energy V_0 to get past the barrier, what happens to its energy once it's on the other side? Well, here the tennis ball analogy is useful to think about. The kinetic energy that was changed into potential energy when the ball was on the hill is converted back into kinetic energy once the ball rolls down the other side.

So, ideally, no energy is lost in this process, just like in quantum tunneling. The particles that start with energy E on the left and manage to tunnel through the barrier, end up with energy E on the right.

It's important to realize the limitations of this analogy. The quantum barrier doesn't look like a hill or a wall - it looks like empty space. A way of setting up a quantum tunneling experiment would be to fire electrons down a pipe where a certain segment is maintained at a greater potential than the energy of the electrons.

The reason we don't see balls tunneling through hills or people tunneling through doors is because these things are made up of so many particles. Even if an electron in one atom tunnelled through one potential barrier, it is incredibly unlikely that the rest of the zillions of electrons and nuclei would manage this all at the same time.

Another limitation of the hill analogy is that we can predict beforehand whether a ball will make it over the hill or not depending on its initial energy. In the quantum world, we can't say which particles will be reflected by the barrier and which ones will tunnel through. We can only give probabilities. We can work these out by using the Schrödinger equation.

For a free particle, the Schrödinger equation yields complex exponential solutions. These are sinusoids in the real and imaginary planes. The wavelengths of these sinusoids are inversely proportional to momentum by the de Broglie relation and hence are also inversely related to energy. The longer the wavenlength, the less the energy—just like in electromagnetic waves. In the barrier region, the solutions are exponentials and the wavelength is undefined.

So, if we were to build a wave function for a beam of tunneling particles it would be sinusoidal on either side of the barrier with equal wavelengths, with a decaying exponential in the barrier region. The wave function and its slope both have to be continuous.

What does the wave function mean, though? It doesn't look like the particles, and it doesn't represent the path that they take; it is just a mathematical construction that is extremely useful in determining particular observables like position and momentum.

There are other ways of visualizing tunneling. Taking the square modulus of the wave function, that is multiplying the wave function by its complex conjugate, gives you the probability density. Multiplied by a small interval dx, this is the probability that you'll find the particle at any given position x. It's a nice way to think about quantum mechanics because the probability of finding the particle is a very tangible idea. If you take the square root of the probability density, you get the amplitude of the wave function. This can be a useful idea too, and it's often used to give a more visually appealing view of the wave function. The CUPS simulations make use of this representation.

Tunneling is not just another example of the strangeness of quantum mechanics, it's also extremely important. Tunneling takes place all the time in the sun, allowing protons to overcome Coulombic barriers and fuse at relatively low temperatures. It also explains how alpha particles can tunnel out of radioactive elements. We would otherwise expect the barrier created by the strong force to be too great for particles to escape. Furthermore, tunneling allows us to see things on the atomic scale through the use of the scanning tunneling microscope. Finally, quantum tunneling highlights a lot of key ideas of quantum mechanics, so it allows us to check how well we understand the things we know.

C.1.2 Dialogue script

A student walks into a tutor's office as he is working at his computer.

Student Hey, do you have a minute? I have a question from my quantum mechanics class.

Tutor Sure, what is it?

Student It's this problem on quantum tunneling—you know, where some particles are able to penetrate a barrier that they wouldn't classically have enough energy to get over. On the potential energy diagram, there's this square barrier, with height, V_0 and the energy of the electrons coming from the left is E, there?

Tutor OK, that's the standard potential energy diagram for tunneling.

Student So my question is this: why don't they draw the energy of the electrons the whole way across the graph, You know, like this.

Tutor Why would the energy of the electrons do that?

Student Because, they're tunneling through the barrier... So, look [region I] here they've got the incident energy E, and then they dissipate some energy in the barrier [region II], so the ones that get through have less energy [region III].

Tutor Why is energy dissipated in the barrier?

Student When something goes through a barrier, it loses energy—to heat and stuff.

Tutor How do you mean?

Student Like if I threw a tennis ball through a wall of water, it would lose some energy.

Tutor But in quantum tunneling, what kind of barrier is it?

Student Uh, a potential barrier?

Tutor Not a physical barrier?

Student No... It doesn't look like a wall or anything. I know the quantum world isn't like everyday stuff—the barrier is just empty space—but there's some potential there... How could we do a tunneling experiment?

Tutor Well, we could fire electrons down a pipe where one segment is maintained at a potential higher than the energy of the electrons. Does that help? How do you like thinking about potential barriers?

Student Um, I'm not sure. I guess I like to think of it as though... if I put a particle there [indicates the barrier region] then if it went into one of the other regions it would have that much kinetic energy.

Tutor So an electron with no energy relative to the barrier then has V_0 kinetic energy if it goes into the other two regions?

Student Yeah, I mean if I put a tennis ball on top of a hill—so it's got this stored potential energy, mgh, right, that's kinda like V_0 and if it rolled down, all that energy would be converted into kinetic energy by the time it got to the bottom of the hill.

Tutor Yeah, I think that analogy works.

Student What about force? You know how the steepness of the hill determines the force on an object? Is it the same way with the potential energy diagram?

Tutor Yes it is. The gradient of a potential energy diagram represents the force, but in the opposite direction to the gradient.

Student But in the tunneling diagram, the lines are vertical, so what's that like infinite force right?!

Tutor Actually that picture they draw is just an idealization. In reality it's just a steep slope, so the force isn't actually infinite.

Student Oh, right.

Tutor Ok so if you had a tennis ball on one side of the hill and you rolled it up and over the hill, would it lose energy?

Student Ideally? You mean like without air friction or slipping or anything?

Tutor Yes.

Student Then no, it wouldn't lose any energy. The energy is just converted from kinetic to potential and then back to kinetic. So on the other side, it would have the same energy I gave it to start with.

Tutor Well it's like that in quantum mechanics too.

Student Really?

Tutor Really.

Student So then, we don't draw the energy line across the graph because it's the same the whole way along. The total energy of each electron never changes [draws straight line]. You know this is really funny. With the tennis ball and stuff (I know we can't think about electrons like tennis balls but) you could say before the ball went over the hill if it would make it or not, depending on its initial energy. But with quantum we can't say that any more. They all have the same energy, so we can only give probabilities that they're going to make it over. The total energy doesn't change, ever.

Tutor But does any type of energy actually change?

Student Well, I guess like we were saying - the kinetic energy gets turned into potential energy and then back into kinetic, so kinetic energy of each electron can't stay constant.

Tutor Right.

Student But wait, if the energy E is less than the height of the barrier, then when the electron's in the barrier, it's got negative kinetic energy.

Tutor Not quite.

Student What do you mean? E minus V_0 is negative?

Tutor It seems to be but this is one of the weird consequences of quantum mechanics. We can't say anymore that electrons have a certain size and definite position and momentum or energy. You can't think of an electron like a tennis ball any more—it's not like a ball, it's more like a wave.

Student So conservation of energy isn't violated because of the uncertainty principle. I guess it also makes sense because the wave function is continuous everywhere so there's gotta be some probability of finding it on the other side of the barrier.

Tutor Right.

Student You know, I used to think that the wave function represented kinetic energy or something. So it was positive in the incident and transmitted areas but negative in the barrier.

Tutor Well the wave function does represent kinetic energy in a way, but not like this, not by its value.

Student Oh, so how then?

Tutor You tell me. How do you find the wave function?

Student Well you have to solve the Schrödinger equation in the three regions. I know that in scattering states—the free particle, like when the particle has higher energy than its surroundings, the solutions of the Schrödinger equation are complex exponentials.

Tutor What do they look like?

Student Sine waves in the real and imaginary planes.

Tutor What about in the barrier?

Student Well the energy is less than the potential there, so um I think the solution is a decaying exponential. You can't say what the wavelength is because it's undefined.

Tutor So how do you get the whole picture of the wave function?

Student You match the solutions in the three regions so the boundary conditions are right.

Tutor What are the boundary conditions?

Student You've got to make sure the function is continuous and the slope has to be continuous as well.

Tutor And what does it all mean?

Student Well I don't know. I don't think it really means anything. It's just a mathematical representation that we can do operations on and find stuff out. Like position, or momentum, or the probability that a particle makes it past a barrier.

Tutor What about the wavelength of the wave function?

Student I'm not sure.

Tutor Do you remember de Broglie?

Student He said that matter also has wave properties and that the wavelength is inversely proportional to momentum...So the wavelength of the wave function is also inversely related to kinetic energy—the longer the wavelength, the less the energy—just like in electromagnetic waves.

Tutor What does that mean for the wave function for a beam of tunneling particles?

Student The wavelength of the wave function must be the same before and after the barrier because the energy is the same.

Tutor Right.

Student But wait, in all this, I thought the wave function looked liked this [axis shift drawing]. Doesn't this height represent the number of particles? Shouldn't it be higher on the left than on the right?

Tutor Well there are more particles on the left than on the right, but I think you're confusing the pictures of the wave function with the probability density.

Student Doesn't this offset relate to the probability of finding a particle in this space?

Tutor It does in the probability density picture but not with the wave function. The wave function doesn't have an offset. [picture changes]

Student Ok so for the wave function of a beam of tunneling particles, you've got sine waves with equal wavelengths on either side of the barrier, with a decaying exponential in the barrier. Then you have to make the whole thing continuous, and make the slope continuous. There.

Tutor Cool.

Student Yeah, but I don't like thinking about the wave function.

Tutor Why not?

Student Because it doesn't mean much to me. I used to think the wave function somehow looked like the particles or was like, the path that they followed. But now I know that's not right. I really prefer thinking about the probability density, ψ squared.

Tutor How do you go from the wave function to the probability density?

Student You just square it.

Tutor That would work if the wave function were purely real, but since it's generally complex, you have to multiply the wave function by its complex conjugate in order to get the actual square modulus, which is the probability density.

Student Yeah that's what I meant, I just always forget how the wave function is complex. Anyway, so the probability density is better because I can think about it in terms of the probability of finding a particle at a particular spot. If I multiply the probability density by a small interval dx, that's the probability that I'll find the particle there at a position x. I like that because it's a really tangible idea—the chance of finding something somewhere.

Tutor So if we had a beam of tunneling particles, at what spot would you most likely find them?

Student Well, if I just think about it, probably in the barrier because that's where the particles would be going slowest. And if they're going the same speed before and after then it would probably look like this ['well' misconception].

Tutor To find out if you're right, why don't you draw the square modulus of the wave function?

Student Oh, yeah that's a good idea, I think it should look like this [rectified sinusoid conception].

Student Woah, that looks nothing like what I thought.

Tutor That looks a lot better than your first answer, but why are there minima and maxima in the transmitted beam? Is there a greater chance of finding the particle here or here than over here? Does that make sense?

Student No, it doesn't. The chance of finding a particle should be the same everywhere after the barrier so this should just be a straight line. Wait, I know, I just squared the wave function. I didn't actually multiply it by its complex conjugate. If I do that, then I get a horizontal line, meaning there's equal probability of finding the particle [standard minus interference conception].

Tutor OK, that's better, but would the reflected beam interfere with the incident beam here?

Student Oh right, you add the wave functions before you take the square modulus. So, in places they cancel out and then you get the interference pattern [standard conception].

Tutor Excellent, that's the probability density for a beam of tunneling particles.

Student You know that looks similar, but not exactly the same as the pictures we see in lectures from that computer simulation CUPS or something. They look more like this [modulus picture]. Do you know why that is?

Tutor Well yeah, that looks like the modulus of the wave function. That's what you get when you take the square root of the probability density. Depending on what you want to do, any of these representations can be useful. Some people like looking at the amplitude of the wave function rather than the real or imaginary parts. After all, mathematical representations contain all the information about the system.

Student Cool, so now I know how to do some really obscure physics.

Tutor But tunneling's not obscure at all. It's vital to your survival.

Student Yeah right. I mean does this actually happen? Not just in the lab?

Tutor Absolutely, I can give you some examples.

Student So why don't we tunnel through doors or see balls tunnel through hills?

Tutor That's because people, doors, balls and hills are all made up of so many particles. Even if one electron in one atom managed to tunnel through a potential barrier, it is so incredibly unlikely that all of the other electrons and nuclei could tunnel at exactly the same time.

Student Yeah, I guess that makes sense. So where does tunneling happen? How is it important to my survival?

Tutor For one, it's happening in the sun all the time. Otherwise the protons wouldn't have enough energy to overcome the Coulombic repulsions between them so fusion reactions couldn't take place and the sun would shut off. Plus, tunneling explains how alpha particles can be emitted from radioactive nuclei. The strong force would otherwise create a barrier that is much too strong for alpha particles to escape. And, a very direct use of tunneling, the scanning tunneling microscope, allows us to image things on the atomic scale!

C.2 Newtonian mechanics

C.2.1 Exposition script

Understanding how objects move is one of the greatest insights provided by physics, first conceived of by Newton over 300 years ago. Newton's laws of motion form the basis for all introductory courses in physics. In the next few minutes, I'm going to take you through Newton's first and second laws with some examples to help illustrate the concepts. For some of you this may be review while for others it may be fairly new.

First off, it's important that we establish some definitions because clear definitions are required to understand the more complicated bits. Speed, how fast something is moving, is the rate of change in distance traveled with time. So if I'm walking through one meter every second, I have a constant speed of one meter per second. Velocity is just speed with a direction attached to it, like one meter per second North or two meters per second up. If I start to speed up so initially I'm going one meter per second, then the next second I'm going two meters per second, and then the second after that I'm going three meters per second and so on, then I have a constant acceleration of 1 meter per second, every second, or one meter per second, per second.

One of the interesting insights Newton had was that all objects like to keep their motion unchanged, going the same direction with the same speed. This is Newton's first law of motion. Stated more formally, it goes: an object will continue with uniform velocity unless acted on by an unbalanced force. Of course, a special, pretty important case of this is that if an object is not moving, it won't move unless an unbalanced force acts on it.

I didn't define force before because it's somewhat of a tricky idea. Essentially it's just a push or a pull—like right now I'm putting a force on this wall—but there are subtleties to the idea of force. Hopefully these will become a bit clearer when we introduce Newton's second law and go through some examples.

You've probably heard of Newton's second law. It describes how an unbalanced force affects the motion of an object. An unbalanced force acting on an object makes it accelerate in the same direction as the force. The bigger the mass of the object, the smaller the resulting acceleration for the same unbalanced force. Or put in its usual form, unbalanced force equals mass times acceleration. This makes a fair deal of sense since it takes a lot less force to accelerate a tennis ball than a lead weight of the same size.

Book moving with constant velocity

Now let's look at an example of a book being pushed at constant velocity across a table. Four forces are acting on the book: there's the force of my hand, the force of gravity downward, the force from the table upward that balances the gravitational force, and friction backwards. While the book is moving at constant velocity, I know that the force from my hand pushing the book forward is exactly equal to the force of friction from the table backwards on the book. The unbalanced force is zero and so the acceleration is zero; the book moves with constant speed.

A position time graph for this situation would look like this [position-time graph]. A velocity-time graph would look like this [velocity-time graph].

Of course at the beginning, when I first start pushing the book and its velocity goes from zero to some final value, the force exerted by my hand must be greater than the force of friction and so the book accelerates.

Newton's laws are easier to see if we take friction out of the equation. If I put a slider on this air track (which has such little friction that we can ignore it) and give it a push, then with no unbalanced force acting on it, the slider continues at constant velocity.

Juggling ball

Now we'll consider a case where gravitational force is the only force acting on an object.

The role of a juggler is to catch the falling balls and throw them back into the air. To catch a ball, the juggler's hand applies an upward force on the ball greater than the downward gravitational force. So there is an unbalanced upward force that accelerates the ball in the upward direction. While in contact with the juggler's hand, the ball first slows down, is momentarily stationary, then it speeds up in the upward direction. The juggler then lets go and the ball leaves his hand with upward velocity.

While the ball is in the air (we will ignore air friction because it is so small) only one force acts on the ball throughout its flight. This is the force of gravity which is constant and downward, accelerating the ball in the downward direction. After being thrown up, a ball travels slower and slower upward, stopping momentarily before speeding up in the downward direction. Then it meets with the juggler's hand again and the process repeats.

Car on ramp

Next, let's consider the case of a toy car rolling down a ramp. This is a bit more complicated than what we have seen before. Now, in addition to the gravitational force downward on the car, there is a force from the ramp pushing on the car. This means that the unbalanced force the car will feel is down the ramp and is only a fraction of the gravitational force. How large or small this fraction is depends on the steepness of the ramp. Again, we'll ignore air and wheel resistance in this example.

If the steepness of the ramp is uniform, the unbalanced force is constant causing the car to accelerate at a constant rate. That is, the velocity of the car down the ramp increases linearly with time as shown in this velocity-time graph taking down the ramp to be positive [velocity-time graph].

If we give the car a push up the ramp, it initially starts out with some velocity up the ramp. As before, the fraction of gravitational force that is not cancelled by the force of the ramp is constant and down the ramp. So there is a constant acceleration in the 'down ramp' direction. That acceleration means that the car's velocity up the ramp gets smaller, goes through zero and then increases down the ramp. You can see what this looks like on a velocity-time graph again taking down the ramp to be positive [velocity-time graph].

C.2.2 Extended exposition script

Understanding how objects move is one of the greatest insights provided by physics, first conceived of by Newton over 300 years ago. Newton is said to have come up with the ideas of universal gravitation and the laws of motion after being struck in the head with a falling apple in an orchard. Although this is most likely an exaggeration, many people believe there is a certain amount of truth to the story. Newton apparently enjoyed looking out the window of his country home and would likely have considered the falling of fruits from trees. The fruit always fell downward, perpendicular to the earth below, never sideways or upward. Seeing as the apple was a good distance from the ground before it fell, Newton likely got the idea that the same force governing the motion of apples could extend over great distances, potentially as far away as the moon. This was the great leap that led ultimately to a coherent view of gravitation and the movement of celestial bodies in the solar system.

These findings are still so relevant and striking that Newton's laws of motion form the basis for all introductory courses in physics. In the next few minutes I'm going to take you through Newton's first and second laws with some examples to help illustrate the concepts. For some of you this may be review while for others it may be fairly new.

First off, it's important that we establish some definitions because clear definitions are required to understand the more complicated bits. Speed, how fast something is moving, is the rate of change in distance traveled with time. So if I'm walking through one meter every second, I have a constant speed of one meter per second. Velocity is just speed with a direction attached to it, like one meter per second North or two meters per second up. If I start to speed up so initially I'm going one meter per second, then the next second I'm going two meters per second, and then the second after that I'm going three meters per second and so on, then I have a constant acceleration of 1 meter per second, every second. Or one meter per second, per second.

If you're looking to buy a really fast car, you'll probably hear about the amount of time it takes for the model to go from zero to one hundred kilometers an hour. The Jaguar XJR-15 can go from zero to a hundred in 3.1 seconds. Assuming constant acceleration that's equal to 8.6 m/s^2. A 2000 Chevy Camero can achieve the same feat in 2.6 seconds, with an average acceleration of 10.3 m/s^2 which is greater than the acceleration due to gravity!

One of the interesting insights Newton had was that all objects like to keep their motion unchanged, going the same direction with the same speed. This is Newton's first law of motion. Stated more formally, it goes: an object will continue at a uniform velocity unless acted on by an unbalanced force. Of course, a special, pretty important case of this is that if an object is not moving, it will not move unless an unbalanced force acts on it.

I didn't define force before because it's somewhat of a tricky idea. Essentially it's just a push or a pull, like right now I'm putting a force on this wall, but there are subtleties to the idea of force. Hopefully these will become a bit clearer when we introduce Newton's second law and go through some examples.

You've probably heard of Newton's second law. It describes how an unbalanced force affects the motion of an object. An unbalanced force acting on an object makes it accelerate in the same direction as the force. The bigger the mass of the object, the smaller the resulting acceleration for the same unbalanced force. Or put in its usual form, unbalanced force equals mass times acceleration. This makes a fair deal of sense since it takes a lot less force to accelerate a tennis ball than a lead weight of the same size.

Newton published his findings in the three volume *Principia* in 1687. This is widely regarded as one of the most important scientific documents of all time. The Fisher Library here at the University of Sydney has a first edition copy of the *Principia* annotated by an unknown mathematician with corresponding notes from Newton himself.

Book moving with constant velocity

Now let's look at an example of a book being pushed at constant velocity across a table. Four forces are acting on the book: there's the force of my hand, the force of gravity downward, the force from the table upward that balances the gravitational force, and friction backwards. When the book is moving at constant velocity, I know that the force from my hand pushing the book forward is exactly equal to the force of friction from the table backwards on the book. The unbalanced force is zero and so the acceleration is zero; the book moves with constant speed.

A position time graph for this situation would look like this [position-time graph]. A velocity-time graph would look like this [velocity-time graph].

Of course at the beginning, when I first start pushing the book and its velocity goes from zero to some final value, the force exerted by my hand must be greater than the force of friction and so the book accelerates.

Newton's laws are easier to see if we take friction out of the equation. If I put a slider on this air track (which has such little friction that we can ignore it) and give it a push, then with no unbalanced force acting on it, the slider continues at constant velocity.

Juggling ball

Now we'll consider a case where gravitational force is the only force acting on an object.

The role of a juggler is to catch the falling balls and throw them back into the air. To catch a ball, the juggler's hand applies an upward force on the ball greater than the downward gravitational force. So there is an unbalanced upward force that accelerates the ball in the upward direction. While in contact with the juggler's hand, the ball first slows down, is momentarily stationary, then speeds up in the upward direction. The juggler then lets go and the ball leaves his hand with upward velocity.

While the ball is in the air (we will ignore air friction because it is so small) only one force acts on the ball throughout its flight. This is the force of gravity which is constant and downward, accelerating the ball in the downward direction. After being thrown up, a ball travels slower and slower upward, stopping momentarily before speeding up in the downward direction. Then it meets with the juggler's hand again and the process repeats.

If you are learning to juggle it might be nice to have the balls fall a bit slower to give you more time to coordinate your efforts catching and throwing the balls. Unfortunately if you use lighter balls, they won't fall any slower than heavy ones. Even though the force of gravity on them is less, it takes proportionately less force to accelerate them by Newton's second law, so there is no net effect and the balls accelerate at the same rate whether they are heavy or light. The only advantage of using light balls is that you won't expend as much energy throwing them into the air. Something you might try to make learning to juggle easier would be juggling tissues or scarves. These items have significant air resistance so they don't accelerate downward at the same rate as balls. Most beginners start out this way and work up to more aerodynamic and even dangerous objects later.

Car on ramp

Next, let's consider the case of a car rolling down a ramp. This is a bit more complicated than what we have seen before. Now, in addition to the gravitational force downward on the car, there is a force from the ramp pushing on the car. This means that the unbalanced force the car will feel is down the ramp and is only a fraction of the gravitational force. How large or small this fraction is depends on the steepness of the ramp. Again, we'll ignore air and wheel resistance in this example.

If the steepness of the ramp is uniform, the unbalanced force is constant causing the car to accelerate at a constant rate. That is, the velocity of the car down the ramp increases linearly with time as shown in this velocity-time graph taking down the ramp to be positive [velocity-time graph].

If we give the car a push up the ramp, it initially starts out with some velocity up the ramp. As before, the fraction of gravitational force that is not cancelled by the force of the ramp is constant and down the ramp. So there is a constant acceleration in the 'down ramp' direction. That acceleration means that the car's velocity up the ramp gets smaller, goes through zero and then increases down the ramp. You can see what this looks like on a velocity-time graph again taking down the ramp to be positive [velocity-time graph].

We could speculate about Newton's fondness for cars or juggling, but what is well-documented are his interests in alchemy, history, and religion.

In addition to his scientific treatises, Newton wrote works on religion such as An historical account of two notable corruptions of scripture, and Observations upon the prophecies of Daniel, and the apocalypse of St. John, though he never gained as much recognition for these works as for his investigations in physics.

Newton got into debates with contemporary scientists like Hooke, who criticized his ideas on optics. By most accounts, Newton was not a very social man. He was engaged once, at the age of nineteen to Anne Storer. He then got busy with his studies, however, and she married someone else. No records of other romantic pursuits exist.

Samples of Newton's hair that were saved for posterity at Cambridge show high concentrations of heavy metals like lead and mercury. These findings have led many historians to believe Newton's eccentricities and abrasive social tendencies, especially later in life, may have been due to metal poisoning. Despite poisoning himself with heavy metals and staring at the sun so long he almost went blind, Newton lived to be 84.

C.2.3 Refutation script

Understanding how objects move is one of the greatest insights provided by physics, first conceived of by Newton over 300 years ago. Newton's laws of motion form the basis for all introductory courses in physics. In the next few minutes I'm going to take you through Newton's first and second laws with some examples to help illustrate the concepts. I will also outline some common misconceptions that exist concerning Newton's first and second laws. For some of you this may be a review while for others it may be fairly new.

It's important that we establish some definitions because clear definitions are required to understand the more complicated bits. Speed, how fast something is moving, is the rate of change in distance traveled with time. So if I'm walking through one meter every second, I have a constant speed of one meter per second. Velocity is just speed with a direction attached to it, like one meter per second North or two meters per second up. If I start to speed up so initially I'm going one meter per second, then the next second I'm going two meters per second, and then the second after that I'm going three meters per second and so on, then I have a constant acceleration of 1 meter per second, every second. Or one meter per second, per second.

One of the interesting insights Newton had was that all objects like to keep their motion unchanged, going the same direction with the same speed. This is Newton's first law of motion. Stated more formally, it goes: an object will continue at a uniform velocity unless acted on by an unbalanced force. Of course, a special, pretty important case of this is that if an object is not moving, it will not move unless an unbalanced force acts on it.

I didn't define force before because it's somewhat of a tricky idea. Essentially it's just a push or a pull, like right now I'm putting a force on this wall, but there are subtleties to the idea of force. Hopefully these will become a bit clearer when we introduce Newton's second law and go through some examples.

You've probably heard of Newton's second law. It describes how an unbalanced force affects the motion of an object. An unbalanced force acting on an object makes it accelerate in the same direction as the force. The bigger the mass of the object, the smaller the resulting acceleration for the same unbalanced force. Or put in its usual form, unbalanced force equals mass times acceleration. This makes a fair deal of sense since it takes a lot less force to accelerate a tennis ball than a lead weight of the same size.

Book moving with constant velocity

Now let's look at an example of a book being pushed at constant velocity across a table. Four forces are acting on the book: there's the force of my hand, the force of gravity downward, the force from the table upward that balances the gravitational force, and friction backwards. When the book is moving at constant velocity, I know that the force from my hand pushing the book forwards is exactly equal to the force of friction from the table backwards on the book. The unbalanced force is zero and so the acceleration is zero; the book moves with constant speed.

A misconception is to think that if the book is moving there is a force inside the book that keeps it moving. Similarly, another misconception is that if no force is applied it will simply come to a stop. This is because we encounter friction when we move objects around, and a force is needed to balance this friction. According to Newton's First Law, once an object is moving with constant velocity, it will remain at that velocity unless acted upon by an unbalanced force. It will not slow down and stop by itself. Newton's laws are easier to see if we take friction out of the equation. If I put a slider on this air track (which has such little friction that we can ignore it) and give it a push, then with no unbalanced force acting on it, the slider continues at constant velocity.

Of course at the beginning, when I first start pushing the book and its velocity goes from zero to some final value, the force exerted by my hand must be greater than the force of friction and so the book accelerates.

A position time graph for this situation would look like this [position-time graph]. A velocity-time graph would look like this [velocity-time graph].

It should be noted that it is easy to confuse the velocity graph with the position graph, thinking that a constant increase in position, is instead increasing speed.

Juggling ball

Now we'll consider a case where gravitational force is the only force acting on an object.

The role of a juggler is to catch the falling balls and throw them back into the air. To catch a ball, the juggler's hand applies an upward force on the ball greater than the downward gravitational force. So there is an unbalanced upward force that accelerates the ball in the upward direction. While in contact with the juggler's hand, the ball first slows down, is momentarily stationary, then speeds up in the upward direction. The juggler then lets go and the ball leaves his hand with upward velocity.

A misconception is that as the ball travels upward, there is an upward force from the juggler's hand that stays with the ball even after it has lost contact with the juggler's hand—a force in the ball to keep it moving. This force gradually dies away until it balances gravitational force at the peak. Then gravity takes over and pulls the ball downward. However, there is no upward force on the ball after it has left the juggler's hand and gravity is acting all the time. In this misconception we are simply confusing velocity with force.

While the ball is in the air (we will ignore air friction because it is so small) only one force acts on the ball throughout its flight. This is the force of gravity which is constant and downward, accelerating the ball in the downward direction. After being thrown up a ball travels slower and slower upward, stopping momentarily before speeding up in the downward direction. Then it meets with the juggler's hand again and the process repeats.

Sometimes we may think that the unbalanced force varies over a ball's flight, but in reality the force of gravity is constant throughout, regardless of where the ball is in its trajectory. For example, a common mistake is believing that a ball has zero acceleration at the top of its flight, when it actually has zero velocity. The ball still experiences the same downward force and therefore, by Newton's second law, has the same downward acceleration as it does at all other points of its path.

Car on ramp

Next, let's consider the case of a car rolling down a ramp. This is a bit more complicated than what we have seen before. Now, in addition to the gravitational force downward on the car, there is a force from the ramp pushing on the car. This means that the unbalanced force the car will feel is down the ramp and is only a fraction of the gravitational force. How large or small this fraction is depends on the steepness of the ramp. Again, we'll ignore air and wheel resistance in this example.

If the steepness of the ramp is uniform, the unbalanced force is constant causing the car to accelerate at a constant rate. That is, the velocity of the car down the ramp increases linearly with time as shown in this velocity-time graph, taking down the ramp to be positive [velocity-time graph].

If we give the car a push up the ramp, it initially starts out with some velocity up the ramp. As before, the fraction of gravitational force that is not cancelled by the force of the ramp is constant and down the ramp. So there is a constant acceleration in the 'down ramp' direction. That acceleration means that the car's velocity up the ramp gets smaller, goes through zero and then increases down the ramp. You can see what this looks like on a velocity-time graph again taking down the ramp to be positive [velocity-time graph].

When drawing this graph, a common mistake is to make a picture of what the phenomena looks like rather than graph the velocities. A common misconception would be to have increasing speed up to a point, then decreasing down to zero, [inverted V] but this doesn't correspond to the car slowing down while going up the ramp, stopping momentarily, and then speeding up down the ramp [correct velocity-time graph].

People who hold this misconception think that an increasing force is required to increase velocity steadily. That is a fallacy, as actually, according to Newton's Second Law, a constant force will result in constant acceleration. Constant acceleration in the direction of motion means a constantly increasing velocity. It is not necessary to have an increasing force on an object, to achieve constantly increasing speed.

C.2.4 Dialogue script

Student Hey, can you help me out with mechanics? I think I get it, kind of, but I have a lot of difficulty when I try to solve problems.

Tutor OK, sure. Let's do an example. If I take this textbook and push it across the table at constant speed, can you describe what's happening in terms of forces? What forces are there?

Student Well, there's the force of your hand, obviously, and there's the force of gravity down, because everything is attracted towards the earth. The table counteracts that; it puts an equal force up. And then there's friction, pushing back on the book.

Tutor So why does the book move at constant speed when I push it across the table?

Student Well, like I said, gravity and the upward force from the table balance out, so there's just the force of your hand, which is slightly greater than friction so it keeps the book moving at a constant speed.

Tutor Hmm, what would Newton's first law of motion say about this?

Student It says an object moving with constant velocity will keep going at that constant velocity unless acted on by an unbalanced force. That's Newton's first law!

Tutor So what kind of motion is the book in when I move it across the table?

Student It's going the same speed... and in the same direction... It's moving with constant velocity.

Tutor So are any unbalanced forces acting on it?

Student No, I guess not... but that means friction must be exactly equal to the force from your hand, otherwise it wouldn't remain in that uniform motion.

Tutor Right.

Student But the forces weren't always balanced were they?

Tutor What do you mean?

Student When you first started pushing the book, it went from rest to some constant speed. That means it accelerated. So you must have put a greater force on it than the force of friction at the start, to accelerate it.

Tutor Yeah, you're right.

Student It's just weird that you don't have to have an unbalanced force to keep something moving, that's what I always thought. Hey, when you're not touching the book, there are no unbalanced forces acting on it then too, right?

Tutor What do you think?

Student Well I know there aren't any unbalanced forces, but the book just sits there. Why doesn't it 'keep going with constant velocity'?

Tutor Well zero is a constant velocity isn't it?

Student Yeah I guess so. If you gave it a push and let go, it would slow down and stop though.

Tutor That's because of friction. It's easier to see if we take friction out of the equation. If I put this slider on an air track, you can see that whatever its motion, without any unbalanced forces it continues at constant speed.

Student Ah, yeah. I just thought things always tend to lose energy, slow down and go to rest.

Tutor Well the things we deal with everyday encounter lots of friction. But sometimes we can assume that friction is so small it doesn't matter. Like in this case, the slider will just keep travelling at the same speed. Can you show me what a velocity-time graph would look like for this slider, when it's moving to the right?

Student Hang on, can you explain velocity to me? You've been using the word speed a lot right? 'It's moving at constant speed' I know what that is. Like if I walk through one meter every second, then I'm going a constant speed of one meter per second. But how is that different from velocity?

Tutor Well velocity is just a speed with a direction attached to it, like one meter per second North or two meters per second up.

Student Oh, ok. That's alright then.

Tutor Now can you draw me a velocity-time graph for the slider?

Student OK, I think that would look like this [position-time graph].

Tutor But if I read off that graph, the speed is different here than at some later time.

Student Oh, right that's just a silly mistake. This is the position-time graph, so the actual velocity-time graph looks like this [velocity-time graph].

Tutor Good.

Student OK, that was a really simple example, but I don't think my exam is going to be all books moving at constant speed across tables. What about something else?

Tutor Hey, Luke's juggling outside, let's see if you can explain that to me. Can you tell me what happens when a single ball goes around once?

Student Well Luke's hand gives the ball a force that drives it upward against gravity, but as it goes up that force dies away, right? So at the top then, it perfectly balances gravity. Then gravity wins and the ball falls downward.

Tutor Hmm. . . you said that the force from his hand and gravity are equal at the top.

Student Yeah.

Tutor Then why doesn't the ball keep doing what it's doing? Like in the book example, no unbalanced force means it remains at constant velocity.

Student I don't know. . . Maybe air resistance. No. I mean they're only balanced for a split second—so then gravity wins. . . I don't know, I must be missing something.

Tutor Does it make sense that Luke's hand can put a force on the ball after it leaves his hand?

Student No. . . not really. But the ball's still going up, isn't it? Doesn't that mean there's a force?

Tutor I think now's a good time for Newton's second law.

Student You mean F equals ma?

Tutor That's the one, but tell me what it means?

Student Well it states that an unbalanced force on an object causes it to accelerate in the same direction as the force. And the bigger the mass of the object, the smaller the resulting acceleration for the same unbalanced force.

Tutor Does that make sense to you?

Student I guess, like it takes a lot less effort to accelerate a tennis ball than a lead weight of the same size.

Tutor What does Newton's second law say about velocity though?

Student Well it doesn't really say anything about velocity. It only tells you what happens to acceleration—you know, the change in velocity.

Tutor So when the ball's going up, is its velocity changing?

Student Yes, it's slowing down.

Tutor Which direction is the acceleration?

Student Down, opposite the direction of motion.

Tutor And when the ball's falling down again, is its velocity changing?

Student Yes, it's speeding up. The acceleration is in the same direction as motion this time: down.

Tutor So, by Newton's second law, which way is the force on the ball.

Student Down. It's just the force of gravity.

Tutor Good. What about at the top of its flight - what forces are acting on the ball?

Student Um, none. It's not going anywhere.

Tutor Just before it stops, which direction is it going?

Student Up.

Tutor And just after it stops, which direction is it going?

Student Down.

Tutor So when it's stopped, could it still be accelerating?

Student Yes - it's accelerating down! That's just weird to think about—that velocity can be changing when velocity is zero. But I guess, like I said, Newton's second law doesn't say anything about velocity, it only talks about acceleration.

Tutor So when the ball's in the air, what forces are acting on it?

Student Just the gravitational force downward. It has the same force acting on it, and so the same constant acceleration at all times.

Tutor Good. Now what about when the ball hits Luke's hand?

Student Well first he catches the ball, right? Like his hand is moving down.

Tutor Which way is he applying a force?

Student He's applying downward force, obviously, when he's catching it.

Tutor But if he's applying downward force, which way would the ball accelerate?

Student Down. Oh, and that doesn't happen does it? So even though his hand is going down, the force is upward, and the force has to be greater than the gravitational force on the ball so the unbalanced force is upward and the ball accelerates in the upward direction.

Tutor Right.

Student So he applies this force and the ball slows down, then its velocity goes through zero and, as he continues to apply the upward force, the ball's velocity increases in the upward direction... and then he lets go.

Tutor Good.

Student OK give me another example.

Tutor How about a car on a ramp?

Student OK, that's a bit more complicated, because now in addition to the force of gravity down, there's the force from the ramp up on the car. But that only cancels some of the gravitational force, depending on the steepness of the ramp, so there's a net force down the ramp.

Tutor So what happens to the car?

Student Well, ignoring air and rolling friction, the car accelerates at a constant rate down the ramp.

Tutor Does the unbalanced force change?

Student No. I mean that depends on the steepness of the hill, so if it stays the same steepness, the unbalanced force is constant and down the hill. The velocity increases linearly with time.

Tutor Can you graph what you mean by that?

Student OK it's like this [velocity-time graph].

Tutor Right. Now, what if we give the car a push at the bottom of the hill and wait for it to come back down?

Student OK so you're saying it starts out with some speed up the ramp?

Tutor Yes, and what about the forces?

Student The unbalanced force is the same as before, constant and down the ramp, so the car slows down as it goes up.

Tutor What about at its highest point? Is the force zero?

Student No, even though the car's velocity goes to zero, it's still accelerating in the downward direction. That's why it starts to speed up down the ramp.

Tutor Absolutely. Now could you make me a velocity-time graph for the car?

Student OK [inverted V picture] I think it would look like this. The car goes up the ramp, turns around and comes back down.

Tutor That kind of looks like the path the car traced out, rather than its velocity.

Student Oh. Is there something wrong with this picture?

Tutor You tell me. When is the car going at zero velocity?

Student Well, according to my graph, it's at the beginning and end of the trip. Wait, that's not right. The car started with some speed up the ramp, and ended with some speed down the ramp. And since we agreed that the acceleration was constant the whole time, the correct velocity time graph must look like this [correct velocity-time graph].

Tutor Right, I think you've got Newton's first and second laws down.

Student Yeah, I feel a bit better about mechanics now, thanks.

Bibliography

Agarwal, A. (2003), 'The real deal on 8.02 TEAL', *The Tech Online Edition* **123**(16).

Altschuler, G. C. (1999), 'Let me edutain you', *The New York Times* p. 50.

Ambrose, B. (1999), Investigation of student understanding of the wave-like properties of light and matter, PhD thesis, University of Washington.

Andriessen, J., Baker, M. & Suthers, D. (2003), *Arguing to Learn*, Kluwer Academic, Boston.

Ausubel, D. P. (1968), *Educational Psychology: A Cognitive View*, Holt, Rinehart and Winston, New York.

Baddeley, A. (1997), *Human Memory: Theory and Practice*, Psychology Press, Hove, U.K.

Badger, R., White, G., Sutherland, P. & Haggis, T. (2001), 'Note perfect: An investigation of how students view taking notes in lectures', *System* **29**(3), 405–17.

Bandura, A. (1986), *Social Foundations of Thought and Action*, Prentice-Hall, Englewood Cliffs, NJ.

Bao, L. & Redish (2002), 'Understanding probabilistic interpretations of physical systems: A prerequisite to learning quantum physics', *American Journal of Physics* **70**(3), 210–17.

Barford, J. & Weston, C. (1997), 'The use of video as a teaching resource in a new university', *British Journal of Educational Technology* **28**(1), 40–50.

Bartlett, F. C. (1932), *Remembering: An Experimental and Social Study*, Cambridge University Press, Cambridge, U.K.

BBC (1970), 'Between two stools'.

BBC News, (2006), 'Youtube hits 100m videos per day'.

Belcher, J. W. (2003), 'Improving student understanding with TEAL', *The MIT Faculty Newsletter* **16**(2), 1–5.

Blinn, J. (1989), The making of The Mechanical Universe, *in* S. R. Ellis, M. Kaiser & A. Grunwald, eds, 'Spatial displays and spacial instruments: Proceedings of a conference sponsored by NASA Ames Research Center, and the School of Optometry, University of California', U.S. Department of Commerce, National Technical Information Service, Asilomar, California, pp. 45–1–45–18.

Borzyskowski, G. (2004), Animated text: More than meets the eye?, *in* R. Atkinson, C. McBeath, D. Jonas-Dwyer & R. Phillips, eds, 'Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference', Perth, pp. 141–144.

Brown, A. L. (1992), 'Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings', *Journal of the Learning Sciences* **2**(2), 141–178.

Burns, T., O'Conner, D. & Stocklmayer, S. (2003), 'Science communication: A contemporary definition', *Public Understanding of Science* **12**, 183–202.

Carey, S. (1986), 'Cognitive science and science education', *American Psychologist* **41**(10), 11123–30.

Carlsen, W. S. (2007), Language and science learning, *in* S. K. Abell & N. G. Lederman, eds, 'Handbook of Research on Science Education', Lawrence Erlbaum Associates, Mahwah, NJ, pp. 57–74.

Chandler, P. & Sweller, J. (1991), 'Cognitive load theory and the format of instruction', *Cognition and Instruction* **8**(4), 293–332.

Chi, M. T. H. (1992), Conceptual change within and across ontological categories: Examples from learning and discovery in science, *in* R. Giere, ed., 'Cognitive Models of Science: Minnesota Studies in the Philosophy of Science', University of Minnesota Press, Minneapolis, MN, pp. 129–186.

Chi, M. T. H. (2005), 'Commonsense conceptions of emergent processes: Why some misconceptions are robust', *Journal of the Learning Sciences* **14**(4), 161–199.

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P. & Glaser, R. (1989), 'Self-explanations: How students study and use examples in learning to solve problems', *Cognitive Science* **13**, 145–182.

Chi, M. T. H. & Slotta, J. D. (1993), 'The ontological coherence of intuitive physics', *Cognition and Instruction* **10**(2/3), 249–260.

Chi, M. T. H., Slotta, J. D. & De Leeuw, N. (1994), 'From things to processes: A theory of conceptual change for learning science concepts', *Learning and Instruction* **4**(1), 27–43.

Clark, A. (1997), *Being There: Putting Brain, Body, and World Together Again*, MIT Press, Boston.

Clark, R. C. & Mayer, R. E. (2003), *E-Learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*, Jossey-Bass/Pfeiffer, San Francisco.

Clark, R. E. (1983), 'Reconsidering research on learning from media', *Review of Educational Research* **53**(4), 445–59.

Clark, R. E. (1988), The future of technology in educational psychology, *in* M. Wittrock & F. Farley, eds, 'The future of educational psychology', Laurence Erlbaum Associates, Hillsdale, NJ.

Clark, R. E. (1994*a*), 'Media and method', *Educational Technology Research and Development* **42**(3), 7–10.

Clark, R. E. (1994*b*), 'Media will never influence learning', *Educational Technology, Research and Development* **42**(2), 21–29.

Clark, R. E. & Estes, F. (1998), 'Technology or craft: What are we doing?', *Educational Technology* **38**(5), 5–11.

Clark, R. E. & Estes, F. (1999), 'The development of authentic educational technologies', *Educational Technology* **39**(2), 5–16.

Clark, R. E. & Salomon, G. (1986), Media in teaching, *in* M. C. Mittrock, ed., 'Handbook of Research on Teaching', 3rd edn, Macmillan, New York, pp. 464–478.

Clement, J. (1982), 'Students' preconceptions in introductory mechanics', *American Journal of Physics* **50**(1), 66.

Collins, A., Brown, J. S. & Newman, S. E. (1987), Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics, Technical report, Centre for the Study of Reading.

Condon, E. U. (1978), 'Tunneling—how it all started', *American Journal of Physics* **46**(4), 319–323.

Cook, M. P. (2006), 'Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles', *Science Education* **90**, 1073–1091.

Cox, R., McKendree, J., Tobin, R., Lee, J. & Mayes, T. (1999), 'Vicarious learning from dialogue and discourse', *Instructional Science* **27**(6), 431–58.

Craig, S. D., Driscoll, D. M. & Gholson, B. (2004), 'Constructing knowledge from dialogue in an intelligent tutoring system: Interactive learning, vicarious learning, and pedagogical agents', *Journal of Educational Multimedia and Hypermedia* **13**(2), 163(21).

Cromer, A. (1997), *Connected Knowledge*, Oxford University Press, New York.

Crouch, C. H., Fagen, A. P., Callan, P. & Mazur, E. (2004), 'Classroom demonstrations: Learning tools or entertainment?', *American Journal of Physics* **72**, 835–838.

Cuban, L. (1986), *Teachers and Machines: The Classroom Use of Technology Since 1920*, Teachers College Press, New York.

Cuban, L., Kirkpatrick, H. & Peck, C. (2001), 'High access and low use of technologies in high school classrooms: Explaining an apparent paradox', *American Educational Research Journal* **38**(4), 813–834.

Dawkins, R. (2006), *The Selfish Gene*, 30th anniversary edn, Oxford University Press, Oxford.

Dean, D. & Kuhn, D. (2007), 'Direct instruction vs. discovery: The long view', *Science Education* **91**(3), 384–397.

Diakidoy, I. N., Kendeou, P. & Ioannides, C. (2003), 'Reading about energy: The effects of text structure in science learning and conceptual change', *Contemporary Educational Psychology* **28**, 335–356.

Diamond, J. (1988), 'Why cats have nine lives', *Nature* **332**, 586–587.

Diamond, J. (1989), 'How cats survive falls from New York skyscrapers', *Natural History* **8**, 20–26.

diSessa, A. A. (1982), 'Unlearning Aristotelian physics: A study of knowledge-based learning', *Cognitive Science* **6**(1), 37–75.

diSessa, A. A. (1996), What do "just plain folk" know about physics?, *in* D. R. Olson, ed., 'Handbook of Education and Human Development: New Models of Learning, Teaching and Schooling', Blackwell Publishers, Cambridge, MA, pp. 705–730.

diSessa, A. A. (2006), A history of conceptual change research: Threads and fault lines, *in* K. Sawyer, ed., 'Cambridge handbook of the learning sciences', Cambridge University Press, Cambridge, U.K., pp. 265–282.

diSessa, A. A. & Sherin, B. L. (1998), 'What changes in conceptual change?', *International Journal of Science Education* **20**(10), 1155–91.

Domert, D., Linder, C. & Ingerman, A. (2005), 'Probability as a conceptual hurdle to understanding one-dimensional quantum scattering and tunneling', *Eur. J. Phys.* **26**, 47–59.

Duit, R. (2007), 'Bibliography: Students' and teacher's conceptions and science education'.

Duit, R. & Treagust, D. F. (2003), 'Conceptual change: A powerful framework for improving science teaching and learning', *International Journal of Science Education* **25**(6), 671–688.

Edelson, D. C. (1996), 'Learning from cases and questions: The Socratic case-based teaching architecture', *Journal of the Learning Sciences* **5**(4), 357–410.

Efron, R. (1990), *The Decline and Fall of Hemispheric Specialization*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Embassy of France in Australia, (1997), 'Images of ccience: Falling cats and australian museums'.

Estes, F. & Clark, R. E. (1999), 'Authentic educational technology: The lynchpin between theory and practice', *Educational Technology* **39**(6), 5–13.

Eysenck, M. W. & Keane, M. T. (2005), *Cognitive Psychology: A Student's Handbook*, 5th edn, Psychology Press, New York.

Falk, J. (2004), Developing a quantum mechanics concept inventory, MSc thesis, Uppsala University.

Falk, J., Linder, C. & Kung, R. L. (in press), 'Review of empirical studies into students' depictions of quantum mechanics', *Educational Research Review* .

Fletcher, P. R. (2004), How tertiary level physics students learn and conceptualize quantum mechanics, PhD thesis, University of Sydney.

Fraser, B. J. & Tobin, K. G., eds (1998), *International Handbook of Science Education*, Vol. 1, Kluwer, Dordrecht.

Galileo, G. (1954), *Dialogues Concerning Two New Sciences*, Dover, New York.

Gorard, S., Roberts, K. & Taylor, C. (2004), 'What kind of creature is a design experiment?', *British Educational Research Journal* **30**(4), 577–590.

Graesser, A. C. & Person, N. K. (1994), 'Question asking during tutoring', *American Educational Research Journal* **31**, 104–137.

Greenbaum, T. L. (1998), *The handbook for focus group research*, 2nd edn, Sage Publications, Thousand Oaks, CA.

Guzzetti, B. J. & Hynd, C., eds (1998), *Perspectives on Conceptual Change*, Lawrence Erlbaum Associates, Mahwah, NJ.

Guzzetti, B. J., Snyder, T. E., Glass, G. V. & Gamas, W. S. (1993), 'Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education', *Reading Research Quarterly* **28**(2), 116–159.

Guzzetti, B. J., Williams, W. O., Skeels, S. A. & Wu, S. M. (1997), 'Influence of text structure on learning counterintuitive physics concepts', *Journal of Research in Science Teaching* **34**(7), 701–19.

Hake, R. R. (1992), 'Socratic pedagogy in the introductory physics laboratory', *Physics Teacher* **30**(9), 546–52.

Hake, R. R. (1998), 'Interactive-engagement vs. traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses', *American Journal of Physics* **66**(1), 64–74.

Halliday, D., Resnick, R. & Walker, J. (2003), *Fundamentals of Physics*, 6th edn, Wiley, New York.

Halloun, I. A. & Hestenes, D. (1985), 'The initial knowledge state of college physics students', *American Journal of Physics* **53**(11), 1043–1055.

Hegarty, M., Carpenter, P. A. & Just, M. A. (1991), Diagrams in the comprehension of scientific text, *in* R. Barr, M. L. Kamil, P. B. Mosenthal & P. D. Pearson, eds, 'Handbook of Reading Research', Vol. 2, International Reading Association, New York.

Hellige, J. B. (1993), *Hemispheric Asymmetry: What's Right and What's Left*, Harvard University Press, Cambridge, MA.

Henderson, C. R. (2002), 'Common concerns about the force concept inventory', *The Physics Teacher* **40**, 542–547.

Hestenes, D., Wells, M. & Swackhamer, G. (1992), 'Force concept inventory', *Physics Teacher* **30**(3), 141–58.

Hewitt, P. G. (1997), *Conceptual Physics*, 8th edn, Addison Wesley, Reading, MA.

Hiller, J. R., Johnston, I. D. & Styer, D. F. (1995), *Quantum mechanics simulations: The consortium for upper-level physics software*, Wiley, New York.

Holland, J. H., Holyoak, K. J., Nisbett, R. E. & Thagard, P. R. (1986), *Induction: Processes of Interference, Learning and Discovery*, MIT Press, Cambridge, MA.

Huffman, D. & Heller, P. (1995), 'What does the force concept inventory actually measure?', *The Physics Teacher* **33**, 138–143.

Johnston, I. D., Crawford, K. & Fletcher, P. (1998), 'Student difficulties in learning quantum mechanics', *International Journal of Science Education* **20**(4), 427–46.

Johnston, I. & The Consortium for Upper Level Physics Software (1996), *Solid State Physics Simulations: The Consortium for Upper-Level Physics Software*, John Wiley & Sons, New York.

Kalyuga, S. (1999), 'Managing split-attention and redundancy in multimedia instruction', *Applied Cognitive Psychology* **13**(4), 351–371.

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

Kalyuga, S., Chandler, P. & Sweller, J. (1998), 'Levels of expertise and instructional design', *Human Factors* **40**, 1–17.

Kane, M. J. & Engle, R. W. (2000), 'Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval', *Journal of Experimental Psychology: Learning, Memory, and Cognition* **26**(2), 336–358.

Kearney, M. & Treagust, D. F. (2001), 'Constructivism as a referent in the design and development of a computer program using interactive digital video to enhance learning in physics', *Australasian Journal of Educational Technology* **17**(1), 64–79.

Kelly, E. & Lesh, R. (2002), 'Understanding and explicating the design experiment methodology', *Building Research Capacity* (3), 1–3.

Kim, S., Yoon, M., Whang, S. M., Tversky, B. & Morrison, J. B. (2007), 'The effect of animation on comprehension and interest', *Journal of Computer Assisted Learning* **23**(3), 260–270.

Kirkpatrick, D. L. (1994), *Evaluating Training Programs: The Four Levels*, 1st edn, Publishers Group West, San Francisco.

Kirschner, P. A., Sweller, J. & Clark, R. E. (2006), 'Why minimal guidance during instruction does not work: An analysis of the fall of constructivist, discovery, problem-based, experiential, and inquiry-based teaching', *Educational Psychologist* **41**(2), 75–86.

Klahr, D. & Li, J. (2005), 'Cognitive research and elementary science instruction: From the laboratory, to the classroom, and back', *Journal of Science Education and Technology* **14**(2), 217–238.

Klein, P. D. (2006), 'The challenges of scientific literacy: From the viewpoint of second-generation cognitive science', *International Journal of Science Education* **28**(2-3), 143–178.

Kleitman, S. & Stankov, L. (2001), 'Ecological and person-oriented aspects of metacognitive processes in test-taking', *Applied Cognitive Psychology* **15**(3), 321–341.

Kozma, R. (2000), 'Reflections on the state of educational technology research and development', *Educational Technology Research and Development* **48**(1), 5–15.

Kozma, R. B. (1991), 'Learning with media', *Review of Educational Research* **61**(2), 179–211.

Kozma, R. B. (1994*a*), 'A reply: Media and methods', *Educational Technology Research and Development* **42**(3), 1–14.

Kozma, R. B. (1994*b*), 'Will media influence learning? reframing the debate', *Educational Technology, Research and Development* **42**(2), 7–19.

Kozma, R. B. & Anderson, R. E. (2002), 'Qualitative case studies of innovative pedagogical practices using ict', *Journal of Computer Assisted Learning* **18**(4), 387–94.

Krendl, K. A. (1986), 'Media influence on learning: Examining the role of preconceptions', *Educational Communication and Technology* **34**(4), 223–34.

Kuhn, T. S. (1996), *The structure of scientific revolutions*, 3rd edn, University of Chicago Press, Chicago, IL.

Lagemann, E. C. & Shulman, L. S. (1999), *Issues in Education Research: Problems and Possibilities*, Jossey-Bass Publishers, San Fransisco.

Langley, P. & Simon, H. A. (1995), 'Applications of machine learning and rule induction', *Communications of the ACM* **38**, 54–64.

Lasry, N. (2007), 'Peer instruction', *The Physics Teacher* **in press**.

Lawrence, I. (1996), 'Quantum physics in school', *Physics Education* **31**(5), 278–86.

LeBon, L. E. (2003), 'Students petition against TEAL', *The Tech Online Edition* **123**(14).

Lee, J., Dineen, F., McKendree, J. & Mayes, T. (1999), Vicarious learning: Cognitive and linguistic effects of observing peer discussion, *in* 'American Educational Research Association', Montreal.

Lewis, R. A. (1995), 'Video introductions to laboratory: Students positive, grades unchanged', *American Journal of Physics* **63**(5), 468–70.

Limon, M. (2001), 'On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal', *Learning and Instruction* **11**(4-5), 357–380.

Lowe, R. (2004), 'Interrogation of a dynamic visualization during learning', *Learning and Instruction* **14**(3), 257–274.

Lowe, R. K. (2003), 'Animation and learning: Selective processing of information in dynamic graphics', *Learning and Instruction* **13**(2), 157–176.

Mayer, R. E. (1997), 'Multimedia learning: Are we asking the right questions?', *Educational Psychologist* **32**(1), 1–19.

Mayer, R. E. (2001), *Multimedia Learning*, Cambridge University Press, Cambridge, U.K.

Mayer, R. E. (2003), ‘The promise of multimedia learning: Using the same instructional design methods across different media’, *Learning and Instruction* **13**(2), 125–139.

Mayer, R. E. (2004*a*), ‘Should there be a three-strikes rule against pure discovery learning?’, *American Psychologist* **59**(1), 14–19.

Mayer, R. E. (2004*b*), ‘Teaching of subject matter’, *Annual Review of Psychology* **55**, 715–744.

Mayer, R. E., Bove, W., Bryman, A., Mars, R. & Tapangco, L. (1996), ‘When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons’, *Journal of Educational Psychology* **88**(1), 64–73.

Mayer, R. E., ed. (2005), *The Cambridge Handbook of Multimedia Learning*, University of Cambridge, Cambridge, U.K.

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.

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.

Mayer, R. E. & Jackson, J. (2005), ‘The case for coherence in scientific explanations: Quantitative details can hurt qualitative understanding’, *Journal of Experimental Psychology Applied* **11**(1), 13–18.

Mazur, E. (1997), *Peer Instruction: A User’s Manual*, Prentice Hall Series in Educational Innovation, Prentice Hall, Upper Saddle River, NJ.

McClosky, M. (1983), ‘Intuitive physics’, *Scientific American* **248**, 122–130.

McDermott, L. C. (1991), ‘What we teach and what is learned - closing the gap’, *American Journal of Physics* **59**(4), 301–315.

McDermott, L. C. & Shaffer, P. S. (2001), *Tutorials in Introductory Physics*, Prentice Hall, Upper Saddle River, NJ.

McInnis, C., James, R. & Hartley, R. (2000), 'Trends in the first year experience in australian universities'.

McKagan, S. B. & Wieman, C. E. (2005), Exploring student understanding of energy through the quantum mechanics conceptual survey, *in* P. Heron, L. McCullough & J. Marx, eds, 'American Institute of Physics Conference', Vol. 818, Salt Lake City, pp. 65–68.

McKendree, J., Lee, J., Dineen, F. & Mayes, T. (1998), The vicarious learner: Helping students "listen in" to learn, Technical report.

McLuhan, M. (1964), *Understanding Media: The Extensions of Man*, Routledge & Kegan Paul, London.

Melton, A. W. (1963), 'Implications of short-term memory for a general theory of memory', *Journal of Verbal Learning and Verbal Behavior* **2**, 1–21.

Millar, R. (1989), 'Constructive criticisms', *International Journal of Science Education* **11**(5), 587–596.

Miller, G. A. (1956), 'The magical number seven, plus or minus two: Some limits on our capacity for processing information', *The Psychological Review* **63**, 81–97.

Mitchell, M. (1993), 'Situational interest: Its multifaceted structure in the secondary school mathematics classroom', *Journal of Educational Psychology* **85**, 424–436.

Moore, D. M., Burton, J. K. & Myers, R. J. (2004), Multiple-channel communication: The theoretical and research foundations of multimedia, *in* D. H. Jonassen, ed., 'Handbook of research on educational communications and technology', Lawrence Erlbaum, Mahwah, N.J., pp. 979–1005.

Moreno, R., Mayer, R. E., Spires, H. A. & Lester, J. C. (2001), 'The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents', *Cognition and Instruction* **19**(2), 177–213.

Morgan, J. T., Wittmann, M. C. & Thompson, J. R. (2004), Student understanding of tunneling in quantum mechanics: Examining interview and survey results for clues to student reasoning, *in* S. Franklin, K. Cummings & J. Marx, eds, 'Physics Education Research Conference Proceedings 2003', Vol. 720, AIP Proceedings, pp. 97–100.

Muller, D. A. (2005), Inside the quantum mechanics lecture: Changing practices, *in* 'Higher Education Research and Development Society of Australasia', Sydney.

Muller, D. A., Lee, K. J. & Sharma, M. D. (2007), 'Coherence or interest: Which is most important in online multimedia learning?', *Australasian Journal of Educational Technology* **in press**.

Muller, D. A. & Sharma, M. D. (2005), Student conceptions of quantum tunneling, *in* 'International Conference on Physics Education', World Scientific Publishing, New Delhi.

Muller, R. & Wiesner, H. (2002), 'Teaching quantum mechanics on an introductory level', *American Journal of Physics* **70**(3), 200–09.

National Research Council, (2002), *Scientific Research in Education*, National Academies Press, New York.

Nelson, D. L., Reed, U. S. & Walling, J. R. (1976), 'Picture superiority effect', *Journal of Experimental Psychology: Human Learning and Memory* **2**, 523–528.

Norman, D. A. (1976), *Memory and Attention: An Introduction to Human Information Processing*, Wiley, New York.

Norman, D. A. & Rumelhart, D. E. (1975), *Explorations in Cognition*, W. H. Freeman, San Fransisco.

Ogborn, J. (1994), Theoretical and empirical investigations of the nature of scientific and commonsense knowledge, PhD thesis, King's College.

Oliver, R., Grant, M. & Younger, G. (1994), 'The perth educational television project', *Australian Journal of Educational Technology* **10**(1), 27–40.

Olsen, R. V. (2002), 'Introducing quantum mechanics in the upper secondary school: A study in Norway', *International Journal of Science Education* **24**(6), 565–74.

Osborne, J. F. (1996), 'Beyond constructivism', *Science Education* **80**(1), 53–82.

Osborne, R. J. & Wittrock, M. C. (1983), 'Learning science: A generative process', *Science Education* **67**(4), 489–508.

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.

Paas, F., Van Merriënboer, J. J. G. & Adam, J. J. (1994), 'Measurement of cognitive load in instructional research', *Perceptual and Motor Skills* **79**, 419–430.

Paivio, A. (1971), *Imagery and verbal processes*, Holt, New York,.

Paivio, A. (1986), *Mental Representations: A Dual Coding Approach*, Oxford Psychology Series, Oxford University Press; Clarendon Press, New York.

Paivio, A. (1991), *Images in Mind: The Evolution of a Theory*, Harvester Wheatsheaf, New York.

Phillips, D. C. (1995), 'The good, the bad, and the ugly: The many faces of constructivism', *Educational Researcher* **24**(7), 5–12.

Piaget, J. (1970), *Genetic Epistemology*, Columbia University Press, New York.

Pines, A. L. & West, L. (1983), A framework for conceptual change with special reference to misconceptions, *in* H. Helm & J. D. Novak, eds, 'Proceedings of the International Seminar on Misconceptions in Science and Mathematics', Cornell University, Ithaca, NY, pp. 47–52.

Pollock, E., Chandler, P. & Sweller, J. (2002), 'Assimilating complex information', *Learning and Instruction* **12**, 61–86.

Posner, G. J., Strike, K. A., Hewson, P. W. & Gertzog, W. A. (1982), 'Accomodation of a scientific concept: Toward a theory of conceptual change', *Science Education* **66**(2), 211–227.

Prosser, M. & Trigwell, K. (1999), *Understanding Learning and Teaching*, The Society for Research into Higher Education, Philadelphia.

Redish, E., Wittmann, M. & Steinberg, R. (2000), Affecting student reasoning in the context of quantum tunneling, *in* 'AAPT Summer Meeting', Guelph, Ontario.

Reiner, M., Slotta, J. D., Chi, M. T. H. & Resnick, L. B. (2000), 'Naïve physics reasoning: A commitment to substance-based conceptions', *Cognition and Instruction* **18**(1), 1–34.

Research on teaching and learning quantum mechanics (1999), *in* D. Zollman, ed., 'Annual Meeting National Association For Research in Science Teaching'.

Rieber, L. P. (1990), 'Animation in computer-based instruction', *Educational technology Research and Development* **38**, 77–86.

Rieber, L. P. (1994), *Computers, graphics, and learning*, Brown & Benchmark, Madison, Wisconsin.

Rieber, L. P., Tzeng, S.-C. & Tribble, K. (2004), 'Discovery learning, representation, and explanation within a computer-based simulation: Finding the right mix', *Learning and Instruction* **14**(3), 307–323.

Rivard, L. P. (1994), 'A review of writing to learn in science: Implications for practice and research', *Journal of Research in Science Teaching* **31**(9), 969–983.

Ronen, M. & Eliahu, M. (2000), 'Simulation—a bridge between theory and reality: The case of electric cicuits', *Journal of Computer Assisted Learning* **16**, 14–26.

Russell, A. L. (1985), 'How media materials are used in university instruction: Reports of selected university faculty members', *Australian Journal of Educational Technology* **1**(1), 47–58.

Ryan, M. P. (2001), 'Conceptual models of lecture learning: Guiding metaphors and model-appropriate notetaking practices', *Reading Psychology* **22**(4), 289–312.

Salomon, G. (1994), *Interaction of Media, Cognition, and Learning*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Schank, R. C. (1979), 'Interestingness: Controlling inferences', *Artificial Intelligence* **12**, 273–297.

Schober, M. F. & Clark, H. H. (1989), 'Understanding by addressees and observers', *Cognitive Psychology* **21**, 211–232.

Schraw, G. & Lehman, S. (2001), 'Situational interest: A review of the literature and directions for future research', *Educational Psychology Review* **13**(1), 23–52.

Schunk, D. H. & Hanson, A. R. (1985), 'Peer models: Influence on children's self-efficacy and achievement', *Journal of Educational Psychology* **77**(3), 313–22.

Schunk, D. H., Hanson, A. R. & Cox, P. D. (1987), 'Peer-model attributes and children's achievement behaviors', *Journal of Educational Psychology* **79**(1), 54–61.

Scott, P., Asoko, H. & Leach, J. (2007), Student conceptions and conceptual learning in science, *in* S. K. Abell & N. G. Lederman, eds, 'Handbook of Research on Science Education', Lawrence Erlbaum Associates, Mahwah, NJ, pp. 31–56.

Semrau, P. & Boyer, B. A. (1994), *Using interactive video in education*, Allyn and Bacon, Boston.

Sharma, M. D., Mendez, A. & O'Byrne, J. W. (2005), 'The relationship between attendance in student-centred physics tutorials and performance in university examinations', *International Journal of Science Education* **27**(11), 1375–1389.

Sharma, M. D., Millar, R. M., Smith, A. & Sefton, I. M. (2004), 'Students' understandings of gravity in an orbiting space-ship', *Research in Science Education* **34**, 267–289.

Sharma, M. D., Millar, R. & Seth, S. (1999), 'Workshop tutorials: Accommodating student-centered learning in large first year university physics courses', *International Journal of Science Education* **21**(8), 839–53.

Slotta, J. D., Chi, M. T. H. & Joram, E. (1995), 'Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change', *Cognition and Instruction* **3**, 373–400.

Smith, J. P. I., diSessa, A. & Roschelle, J. (1993), 'Misconceptions reconceived: A constructivist analysis of knowledge in transition', *Journal of the Learning Sciences* **3**(2), 115–163.

Sokoloff, D. R. & Thornton, R. K. (1997), 'Using interactive lecture demonstrations to create an active learning environment', *The Physics Teacher* **35**(10), 340–347.

Solomon, J. (1994), 'The rise and fall of constructivism', *Studies in Science Education* **23**, 1–19.

Stevenson, A. (2005), 'Empty lecture halls in the online campus', *Sydney Morning Herald* p. 4.

Strauss, A. L. & Corbin, J. M. (1998), *Basics of qualitative research: Techniques and procedures for developing grounded theory*, 2nd edn, Sage Publications, Thousand Oaks, CA.

Strike, K. A. & Posner, G. J. (1992), A revisionist theory of conceptual change, *in* R. Duschl & R. Hamilton, eds, 'Philosophy of Science, Cognitive Psychology, and Educational Theory and Practice', State University of New York, Albany, NY.

Styer, D. F. (1996), 'Common misconceptions regarding quantum mechanics', *American Journal of Physics* **64**(1), 31–34.

Swartz, C. (1999), 'Editorial: Demise of a shibboleth', *The Physics Teacher* **37**, 330.

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

Sweller, J. (1994), 'Cognitive load theory, learning difficulty, and instructional design', *Learning and Instruction* **4**, 295–312.

Sweller, J. (2004), 'Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture', *Instructional Science* **32**(1-2), 9–31.

Sweller, J. & Cooper, G. A. (1985), 'The use of worked examples as a substitute for problem solving in learning algebra', *Cognition and Instruction* **2**, 59–89.

Sweller, J. & Levine, M. (1982), 'Effects of goal specificity on means-ends analysis and learning', *Journal of Experimental Psychology: Learning, Memory, and Cognition* **8**(5), 463–74.

Sweller, J., Mawer, R. & Howe, W. (1983), 'Development of expertise in mathematical problem solving', *Journal of Experimental Psychology: General* **112**, 639–661.

Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. (1998), 'Cognitive architecture and instructional design', *Educational Psychology Review* **10**(3), 251–296.

Tabbers, H. K., Martens, R. L. & van Merriënboer, J. J. G. (2004), 'Multimedia instruction and cognitive load theory: Effects of modality and cueing', *British Journal of Educational Psychology* **74**, 71–81.

Taylor, R., ed. (1980), *The Computer in the School: Tutor, Tool, Tutee*, Teachers College Press, New York.

The Design-Based Research Collective (2003), 'Design-based research: An emerging paradigm for educational inquiry', *Educational Researcher* **32**(1), 5–8.

Thornton, R. K. & Sokoloff, D. R. (1998), 'Assessing student learning of Newton's laws: The force and motion conceptual evaluation of active learning laboratory and lecture curricula', *American Journal of Physics* **66**(4), 338–352.

Treagust, D. F. (1987), 'The development and use of diagnostic instruments to evaluate students' misconceptions in science', *International Journal of Science Education* **10**, 159–169.

Trowbridge, D. E. & McDermott, L. C. (1980), 'Investigation of student understanding of the concept of velocity in one dimension', *American Journal of Physics* **48**(12), 1020–1028.

Trowbridge, D. E. & McDermott, L. C. (1981), 'Investigation of student understanding of the concept of acceleration in one dimension', *American Journal of Physics* **49**(3), 242–253.

Tyack, D. & Cuban, L. (1995), *Tinkering Toward Utopia*, Harvard University Press, Cambridge, MA.

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G. & Schmidt, H. G. (2004), 'Memory load and task-evolved pupillary responses in aging', *Psychophysiology* **41**(2), 167–174.

van Joolingen, W. R., de Jong, T. & Dimitrakopoulou, A. (2007), ‘Issues in computer supported inquiry learning in science’, *Journal of Computer Assisted Learning* **23**(2), 111–119.

van Merriënboer, J. J. G. & Sweller, J. (2005), ‘Cognitive load theory and complex learning: Recent developments and future directions’, *Educational Psychology Review* **17**(2), 147–177.

Vosniadou, S. (1994), ‘Capturing and modeling the process of conceptual change’, *Learning and Instruction* **4**, 45–69.

Vosniadou, S., Ioannides, C., Dimitrakopoulou, A. & Papademetriou, E. (2001), ‘Designing learning environments to promote conceptual change in science’, *Learning and Instruction* **11**(4-5), 381–419.

Vosniadou, S. & Verschaffel, L. (2004), ‘Extending the conceptual change approach to mathematics learning and teaching’, *Learning and Instruction The Conceptual Change Approach to Mathematics Learning and Teaching* **14**(5), 445–451.

Vygotsky, L. S. (1978), *Mind in Society: The Development of Higher Psychological Processes*, Harvard University Press, Cambridge, MA.

Weaver, G. C. (1998), ‘Strategies in K–12 science instruction to promote conceptual change’, *Science Education* **82**, 455–472.

Wendt, A. (1992), ‘Anarky is what the states make of it: The social construction of power politics’, *International Organization* **46**(2), 391–425.

Wetzel, C. D., Radtke, P. H. & Stern, H. W. (1994), *Instructional Effectiveness of Video Media*, Lawrence Erlbaum Associates, Hillsdale, N.J.

Wittmann, M. C. (2003), Understanding coordinated sets of resources: An example from quantum tunneling, *in* ‘Proceedings of the Fermi School in Physics Education Research’, Varenna, Italy.

Wittmann, M. C. & Morgan, J. T. (2004), Understanding data analysis from multiple viewpoints: An example from quantum tunneling, *in* S. Franklin, K. Cummings & J. Marx, eds, 'Physics Education Research Conference Proceedings 2003', Vol. 720, AIP Conference Proceedings, pp. 3–6.

Wittmann, M. C., Morgan, J. T. & Bao, L. (2005), 'Addressing student models of energy loss in quantum tunneling', *European Journal of Physics* **26**, 939–950.

Wittmann, M. C., Steinberg, R. N. & Redish (2002), 'Investigating student understanding of quantum physics: Spontaneous models of conductivity', *American Journal of Physics* **70**(3), 218–26.

Yeo, S., Loss, R., Zadnik, M., Harrison, A. & Treagust, D. F. (2004), 'What do students really learn from interactive multimedia? A physics case study', *American Journal of Physics* **72**(10), 1351–1358.

Yeo, S. & Zadnik, M. (2001), 'Introductory thermal concept evaluation: Assessing students' understanding', *The Physics Teacher* **39**, 496–504.

Young, H. D. & Freedman, R. A. (2000), *University Physics: With Modern Physics*, World student series., 10th edn, Addison-Wesley, San Francisco.

Zacharia, Z. C. (2007), 'Comparing and combining real and virtual experimentation: An effort to enhance students' conceptual understanding of electric circuits', *Journal of Computer Assisted Learning* **23**(2), 120–132.

Zillmann, D., Masland, J. L., Weaver, J. B., Lacey, L. A., Jacobs, N. E., Dow, J. H., Klein, C. A. & Banker, S. R. (1984), 'Effects of humorous distortions on children's learning from educational television', *Journal of Educational Psychology* **76**(5), 802–12.

Zollman, D., Rebello, N. & Hogg, K. (2002), 'Quantum mechanics for everyone: Hands-on activities integrated with technology', *American Journal of Physics* **70**(3), 252–259.